Lie Isomorphisms of Triangular and Block-Triangular Matrix Algebras over Commutative Rings

by

Anthony John Cecil
MA, University of Oxford, 1997

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Mathematics and Statistics

Lie Isomorphisms of
Triangular and Block-Triangular
Matrix Algebras over Commutative Rings

by

Anthony John Cecil
MA, University of Oxford, 1997

Supervisory Committee

Dr. Ahmed R. Sourour, Supervisor
(Department of Mathematics and Statistics)

Dr. John Phillips, Departmental Member
(Department of Mathematics and Statistics)

Dr. Venkatesh Srinivasan, Outside Member
(Department of Computer Science)

Supervisory Committee

Dr. Ahmed R. Sourour, Supervisor
(Department of Mathematics and Statistics)

Dr. John Phillips, Departmental Member
(Department of Mathematics and Statistics)

Dr. Venkatesh Srinivasan, Outside Member
(Department of Computer Science)

ABSTRACT

For many matrix algebras, every associative automorphism is inner. We discuss results by Đoković that a non-associative Lie automorphism φ of a triangular matrix algebra T_n over a connected unital commutative ring, is of the form $\varphi(A) = SAS^{-1} + \tau(A)I$ or $\varphi(A) = -SJA^{\top}JS^{-1} + \tau(A)I$, where $S \in T_n$ is invertible, J is an antidiagonal permutation matrix, and τ is a generalized trace. We incorporate additional arguments by Cao that extended Đoković's result to unital commutative rings containing nontrivial idempotents.

Following this we develop new results for Lie isomorphisms of block upper-triangular matrix algebras over unique factorization domains. We build on an approach used by Marcoux and Sourour to characterize Lie isomorphisms of nest algebras over separable Hilbert spaces.

We find that these Lie isomorphisms generally follow the form $\varphi = \sigma + \tau$ where σ is either an associative isomorphism or the negative of an associative anti-isomorphism, and τ is an additive mapping into the center, which maps commutators to zero. This echoes established results by Martindale for simple and prime rings.

Contents

ACKNOWLEDGEMENTS

I would like to thank my supervisor Ahmed R. Sourour for his support and patience. My gratitude is also due to Chris Bruce for many interesting conversations and brainstorming sessions. Finally, my deepest thanks go to Holly and Siena for standing by me through the slings and arrows of outrageous fortune.

The author's research was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC).

Chapter 1

Introduction

Matrix algebras and more general operator algebras appear throughout many branches of mathematics and physics. Characterizing their isomorphisms and automorphisms provides insight into the structure of these algebras.

Over the complex or real numbers, finite-dimensional matrix algebras are subsumed within the broader collection of operator algebras on Hilbert spaces. These algebras have been extensively studied and many deep results are known. As a primary example, we have the classical Skolem-Noether theorem which states that any associative automorphism of a central simple algebra is inner. In particular this holds for $M_n(\mathbb{K})$ with $\mathbb{K}$ a field. However, matrix algebras over more general commutative rings have not been as extensively studied.

Using the standard associative multiplication in a matrix or operator algebra, we may also construct a non-associative Jordan algebra or Lie algebra. These algebras were initially motivated by questions in physics but were quickly discovered to have wide-ranging importance in other areas of mathematics.

In the case of Lie isomorphisms of simple rings, there is a strong result of Martindale [8]: Let $\mathcal{S}$ and $\mathcal{R}$ be simple unital rings, not of characteristic 2 or 3, such that $\mathcal{S}$ contains two nonzero orthogonal idempotents whose sum is 1. If $\varphi : \mathcal{S} \to \mathcal{R}$ is a Lie isomorphism, then $\varphi = \sigma + \tau$ where σ is either an associative isomorphism or the negative of an associative anti-isomorphism of $\mathcal{S}$ onto $\mathcal{R}$, and τ is an additive mapping of $\mathcal{S}$ into the center of $\mathcal{R}$ which maps commutators to zero.

Martindale also proved a modified version of this result for prime rings. In this paper we are interested in triangular or block triangular matrices over a commutative unital ring, and in these cases the algebra is no longer simple or prime. Hence, these established results cannot be applied directly.

In Chapter 2 we will summarize our notation and some fundamental results for triangular and block triangular matrix algebras over commutative rings. Chapter 3 examines previous work by Đoković [9] and Cao [1] characterizing Lie isomorphisms of upper-triangular matrix algebras over commutative rings. Finally, building on an approach by Marcoux and Sourour [7], Chapter 4 details new results for Lie isomorphisms of upper-block-triangular matrix algebras over unique factorization domains. In both the triangular and block-triangular cases the characterization echoes the form described by Martindale.

To make these results accessible to a wider audience we adopt an expository style throughout this thesis. However, several of the proofs require detailed calculations, so we thank the reader in advance for their patience reading these sections.

Chapter 2

Background

2.1 Notation

Let $\mathscr{R}$ be a non-trivial commutative ring with multiplicative identity 1 and let $\mathscr{R}^\times$ be the multiplicative group of invertible elements of $\mathscr{R}$. In general, we will refer to a ring with a multiplicative identity as a **unital ring** and, without further qualification, we will assume that rings in this work are unital. We may put additional restrictions on $\mathscr{R}$, such as it being an integral domain or field. For some of our results, $\mathscr{R}$ will be a unique factorization domain (UFD) such that $\mathscr{R}$ is not of characteristic 2 or 3.

Let $M_n(\mathscr{R})$ be the $\mathscr{R}$-algebra of $n \times n$ matrices with entries in $\mathscr{R}$, under the standard associative matrix multiplication. Unless stated explicitly, we assume $n \geq 2$. It is often useful to view these matrices as endomorphisms ($\mathscr{R}$-linear operators) of the free module $\mathscr{R}^n$, which we will express using the standard basis $\{e_i\}_{i=1}^n$. We use the standard notation $GL_n(\mathscr{R}) := M_n(\mathscr{R})^\times$ for the group of invertible elements in $M_n(\mathscr{R})$.

The rank of a free module is well-defined and, in our case, we will only be dealing with the standard basis of $\mathscr{R}^n$, since it is the matrix algebra itself that is of interest. For some arguments we will be using the fraction field $\mathbb{K}$ of $\mathscr{R}$, when the latter is an integral domain. The vector space $\mathbb{K}^n$ uses the same standard basis and its dimension is clearly equal to the rank of $\mathscr{R}^n$. As such we will abuse terminology and use the term *dimension* for both cases, rather than rank.

Idempotents

In a ring or algebra an element α is called **idempotent** if $\alpha^2 = \alpha$. The familiar idempotent elements of a matrix algebra are its projections, but the underlying ring itself may also contain idempotent elements. A ring with only 1 and 0 as idempotent is sometimes termed a **connected ring**.

Integral Domains

If $\mathscr{R}$ is an integral domain, the cancellation property states that for nonzero $\alpha \in \mathscr{R}$ and any $\beta, \gamma \in \mathscr{R}$ with $\alpha\beta = \alpha\gamma$, we have $\beta = \gamma$. This carries over to scalar multiples of matrices over integral domains. For $B, C \in M_n(\mathscr{R})$ then $\alpha B = \alpha C$ implies $B = C$, since the equality can be examined entry-wise. For the same reason, given $x, y \in \mathscr{R}^n$ with $\alpha x = \alpha y$, we have $x = y$.

Cancellation is often phrased as the absence of "zero divisors", so that if $\alpha\beta = 0$ then $\alpha = 0$ or $\beta = 0$. This shows that the only idempotent elements of an integral domain will be 1 and 0. If $\alpha^2 = \alpha$, this follows from the relation $\alpha(1 - \alpha) = 0$.

A **unique factorization domain** (UFD) is an integral domain such that every nonzero non-unit element can be uniquely written as a product of prime elements, up to order and multiplication by a unit. In this case, a greatest common divisor (gcd) of any two elements is defined up to multiplication by a unit. In a UFD, prime elements are equivalent to irreducible elements.

Interestingly, the polynomial ring $\mathscr{R}[x]$ over a UFD $\mathscr{R}$ is itself a UFD. As such, this extends to polynomials in an arbitrary number of unknowns. The conditions for a UFD are weaker than for a principal ideal domain (PID), but $\mathscr{R}[x]$ is a PID if and only if $\mathscr{R}$ is a field.

Lie Algebras

If we consider only the linear structure of $M_n(\mathscr{R})$ and introduce the bracket operation

$$[A, B] = AB - BA$$

for $A, B \in M_n(\mathscr{R})$, this forms a non-associative **Lie algebra**. The bracket product is bilinear, alternating, and anti-commutative. Instead of associativity, the Lie product satisfies the Jacobi identity. Explicitly, for all $A, B, C \in M_n(\mathscr{R})$ and $\alpha, \beta \in \mathscr{R}$, we have

$$\begin{aligned} \text{Linear on the RHS:} \quad & [\alpha A + \beta B, C] = \alpha[A, C] + \beta[B, C] \\ \text{Linear on the LHS:} \quad & [C, \alpha A + \beta B] = \alpha[C, A] + \beta[C, B] \\ \text{Alternating:} \quad & [A, A] = 0 \\ \text{Anti-commutative:} \quad & [A, B] = -[B, A] \\ \text{Jacobi identity:} \quad & [A, [B, C]] + [C, [A, B]] + [B, [C, A]] = 0. \end{aligned}$$

If $\mathcal{A} \subseteq M_n(\mathscr{R})$ and $\mathcal{B} \subseteq M_m(\mathscr{R})$ are Lie sub-algebras, we define a **Lie homomorphism** to be an $\mathscr{R}$-linear map $\varphi : \mathcal{A} \to \mathcal{B}$ that preserves the bracket product:

$$\varphi([A_1, A_2]) = [\varphi(A_1), \varphi(A_2)], \quad \text{or explicitly}$$
$$\varphi(A_1 A_2 - A_2 A_1) = \varphi(A_1)\varphi(A_2) - \varphi(A_2)\varphi(A_1),$$

for all $A_1, A_2 \in \mathcal{A}$.

A **Lie isomorphism** is a bijective Lie homomorphism, which is sufficient for φ^{-1} to also be a Lie homomorphism. For $B_1, B_2 \in \mathcal{B}$, there exists $A_1 = \varphi^{-1}(B_1)$ and $A_2 = \varphi^{-1}(B_2)$ in $\mathcal{A}$. Then,

$$\begin{aligned} \varphi^{-1}[B_1, B_2] &= \varphi^{-1}[\varphi(A_1), \varphi(A_2)] \\ &= \varphi^{-1}(\varphi[A_1, A_2]) = [A_1, A_2] \\ &= [\varphi^{-1}(B_1), \varphi^{-1}(B_2)]. \end{aligned}$$

Since Lie isomorphisms are the primary topic of this paper, we will use the term "associative" (ideal, isomorphism, etc.) when distinguishing the usual matrix multiplication (as an associative $\mathscr{R}$-algebra) from the Lie bracket operation (as a Lie $\mathscr{R}$-algebra).

We recall that an **automorphism** is an isomorphism from an algebra back to itself, and the standard notation for the group of associative $\mathscr{R}$-algebra automorphisms of an algebra $\mathcal{A}$ is $\mathrm{Aut}(\mathcal{A})$. For the group of Lie $\mathscr{R}$-algebra automorphisms we use $\mathrm{Aut}_L(\mathcal{A})$. In both cases, we denote the identity automorphism as id.

Matrix Notation

We denote the identity matrix as I and use $\mathtt{E}_{ij}$ (in typewriter font) for the matrix with 1 in the ij^{th} entry and zeros everywhere else. These elements provide a basis for $M_n(\mathscr{R})$ when viewed as an $\mathscr{R}$-module itself. While not part of this basis, we define $\mathtt{E}_{00} = [0]$ to be the zero matrix. For $\alpha \in \mathscr{R}$ we call αI a **scalar matrix** or, by abuse of terminology, just a "scalar" when the context is reasonably unambiguous.

If we wish to refer explicitly to the entries of a matrix $A \in M_n(\mathscr{R})$, we will write $A = [a_{ij}] \coloneqq \sum_{i,j=1}^{n} a_{ij}\mathtt{E}_{ij}$ where $a_{ij} \in \mathscr{R}$. Using the Kronecker delta notation,

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases},$$

then $I = [\delta_{ij}]$. We will use this matrix notation sparingly since it overloads the brackets also used for the Lie product.

Let $A^\top$ denote the usual matrix transpose $[a_{ij}]^\top = [a_{ji}]$. A particularly useful permutation matrix is defined to have ones on the antidiagonal and zeros elsewhere:

$$J = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & & 1 & 0 \\ & & \iddots & & \\ 0 & 1 & & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} = \sum_{i=1}^{n} \mathtt{E}_{i,n+1-i} = [\delta_{i,n+1-i}] \,.$$

We will reserve the letter J to represent this matrix throughout this work. We see that $J^2 = I$ and $J^\top = J$. For $A \in M_n(\mathscr{R})$, taking the product $JA^\top J$ reflects the matrix across the antidiagonal. Intuitively, we might think of this operation as an alternate kind of transpose.

Structure Equations

If $\{e_k\}_{k=1}^{n}$ is an explicit basis for an arbitrary $\mathscr{R}$-algebra $(\mathcal{A}, +, \cdot)$, then we may write the product of basis elements as $e_i \cdot e_j = \sum_{k=1}^{n} \lambda_{ijk} e_k$ for some $\lambda_{ijk} \in \mathscr{R}$, which are called *structure constants* or *structure coefficients*. These determine the multiplication on the whole algebra by linearity. If there exists a convenient form for writing each $e_i \cdot e_j$, they are called the **structure equations** for the algebra.

As an associative algebra, the structure equations of $M_n(\mathscr{R})$ in terms of the basis $\mathtt{E}_{ij}$ are

$$\mathtt{E}_{ij}\mathtt{E}_{rs} = \delta_{jr}\mathtt{E}_{is} \ ,$$

for i, j, r, s from 1 to n. From this, we see that the structure equations of $M_n(\mathscr{R})$ as Lie algebra are

$$[\mathtt{E}_{ij}, \mathtt{E}_{rs}] = \mathtt{E}_{ij}\mathtt{E}_{rs} - \mathtt{E}_{rs}\mathtt{E}_{ij} = \delta_{jr}\mathtt{E}_{is} - \delta_{si}\mathtt{E}_{rj} \ .$$

The Standard Bilinear Form and Duality

If $\mathscr{R}$ is a field then $\mathscr{R}^n$ becomes a vector space. However, for commutative rings in general it is rare to be able to define an inner product (including over fields with non-zero characteristic). Hence, we cannot rely on the very powerful results available for inner product spaces. However, for $x = \sum x_i e_i$ and $y = \sum y_i e_i$ expressed using the standard basis in $\mathscr{R}^n$, we will denote the "standard" symmetric **bilinear form** as $(x|y) = \sum x_i y_i$. When $\mathscr{R} = \mathbb{R}$, this coincides with the inner product, but this bilinear form is not positive definite for more general rings.

Using the standard bilinear form, we can show that a finite-dimensional free module over a commutative ring is self-dual, so $\mathscr{R}^n \cong (\mathscr{R}^n)^*$. For a standard proof, see Hungerford [3], Theorem IV.4.11. In this spirit, we provide a short constructive separation lemma.

Lemma 2.1.1 (Separation Lemma)**.** *Let $\mathscr{R}$ be a commutative ring. For any linearly independent $y, x \in \mathscr{R}^n$, there exists $z \in \mathscr{R}^n$ such that $(z|y) = 0$ and $(z|x) \neq 0$.*

This is equivalent to the existence of $\psi_z = (z|\cdot) \in (\mathscr{R}^n)^*$ that separates y and x.

Proof. We write $x = \sum x_i e_i$ and $y = \sum y_i e_i$ using the standard basis. Linear independence gives $\alpha y \neq \beta x$ for all $\alpha, \beta \in \mathscr{R}$ unless both α and β are zero. For some j, we have $y_j \neq 0$. There then exists some $k \neq j$ such that $y_j x_k \neq x_j y_k$, otherwise we would have $y_j x = x_j y$, contradicting independence. We use this j and k to define $z = y_j e_k - y_k e_j$.

Then, $(z|y) = y_j y_k - y_k y_j = 0$ and $(z|x) = y_j x_k - y_k x_j \neq 0$. □

If we represent elements $x, y \in \mathscr{R}^n$ as column vectors ($n \times 1$ matrices), by abuse of notation we can equate the bilinear form $(x|y)$ with the 1×1 matrix $x^\top y$. Then, for $A \in M_n(\mathscr{R})$, we have $(Ax|y) = (Ax)^\top y = x^\top A^\top y = (x|A^\top y)$, as we would expect from the duality. On the other hand, $xy^\top$ is a rank-1 $n \times n$ matrix, which is akin to the notation $x \otimes y^*$ used in a Hilbert space context, or the bra-ket notation $|x\rangle\langle y|$ used by physicists. In several places we will use the fact that $e_i e_j^\top = \mathtt{E}_{ij}$.

2.2 Triangular Algebras

For $n \geq 2$, we define $T_n(\mathscr{R}) \subseteq M_n(\mathscr{R})$ to be the algebra of upper-triangular matrices; i.e., those that are zero below the main diagonal; i.e., if $A \in T_n(\mathscr{R})$ is written $A = [a_{ij}]$, then $a_{ij} = 0$ for $i > j$.

The set of strictly upper triangular matrices, with zeros on the main diagonal, will be denoted $\mathcal{U}$ and we shall see that it is both an associative and a Lie ideal of $T_n(\mathscr{R})$. Explicitly, if $A \in \mathcal{U} \subseteq T_n(\mathscr{R})$ is written $A = [a_{ij}]$, then $a_{ij} = 0$ for $i \geq j$.

2.3 Block-Triangular Algebras

We next define the notation for an upper block-triangular algebra, $\mathcal{T} \subseteq M_n(\mathscr{R})$. Given $n > 1$, let $n_1, \ldots, n_k \in \mathbb{Z}^+$ be such that $\sum_{i=1}^k n_i = n$.

For $\mathcal{T}$ we may then use the more descriptive notation $T(n_1, \ldots, n_k)(\mathscr{R})$ to mean the algebra of elements $[A_{ij}]_{i,j=1}^k$, where A_{ij} is a $n_i \times n_j$ block matrix over $\mathscr{R}$ if $i \leq j$ and $A_{ij} = [0]$ if $i > j$. For example, for $a_{ij}, b_{ij}, c_{ij}, m_{ij} \in \mathscr{R}$,

$$S = \left[\begin{array}{cc:cc:c} a_{11} & a_{12} & m_{13} & m_{14} & m_{15} \\ a_{21} & a_{22} & m_{23} & m_{24} & m_{25} \\ \hdashline 0 & 0 & b_{33} & b_{34} & m_{35} \\ 0 & 0 & b_{43} & b_{44} & m_{45} \\ \hdashline 0 & 0 & 0 & 0 & c_{55} \end{array}\right]$$

represents a general element in $T(2,2,1)(\mathscr{R})$.

We note that there are k blocks along the diagonal in $T(n_1, \ldots, n_k)(\mathscr{R})$ of size $n_i \times n_i$ for i from 1 to k. For convenience, we may break the identity matrix into these same blocks so that $I = \sum_{i=1}^k I_{ii}$.

Nests

Viewing elements of $\mathcal{T}$ as endomorphisms of $\mathscr{R}^n$, we observe that there is a nest of subspaces that remain invariant under the action of these matrix operators. In addition to the block indices n_i, we will find it useful to have an index that tracks the invariant subspaces. Thus, we define $m_0 = 0$ and $m_t = \sum_{i=1}^{t} n_i$ for $1 \le t \le k$. We see that $m_k = n$ and $n_t = m_t - m_{t-1}$.

Using this index, let $N_t = \text{span}\{e_i\}_{i=1}^{m_t}$ and define the whole **nest** of subspaces as $\mathcal{N} = \{N_t\}_{t=0}^{k}$. Here, we have defined N_0 to be the zero subspace and we note that $N_k = \mathscr{R}^n$. The terminology "nest" comes from the fact that $N_{t-1} \subseteq N_t$ for $t \in \{1, \dots, k\}$. We will use the notation $\mathcal{N}^0 = \mathcal{N} \setminus \{\{0\}, \mathscr{R}^n\}$, for the nontrivial subspaces in the nest. In Chapter 4 we will rely on this m_t to reference the dimension of the nest subspace N_t.

If we are only dealing with a single block-triangular algebra $\mathcal{T}$ we will assume that the corresponding nest of invariant subspaces is labelled $\mathcal{N}$. If we have more than one algebra, we may use the corresponding nests to distinguish between them. For example $\mathcal{T}(\mathcal{N})$ versus $\mathcal{T}(\mathcal{M})$. Where the block structure is explicit, the notation $\mathcal{T}(\mathcal{N}) := T(n_1, \dots, n_k)(\mathscr{R})$ gives the required information about $\mathcal{N}$.

When dealing with block-triangular algebras, we will reserve n to be the size of the matrices in $\mathcal{T}(\mathcal{N}) \subseteq M_n(\mathscr{R})$, and k to be the number of blocks on the diagonal of the matrices in the algebra.

Continuing the example of $T(2, 2, 1)(\mathscr{R})$, we have:

$$
\begin{aligned}
N_0 &= \{0\} \\
N_1 &= \text{span}\{e_i\}_{i=1}^{m_1} = \text{span}\{e_1, e_2\} \\
N_2 &= \text{span}\{e_i\}_{i=1}^{m_2} = \text{span}\{e_1, e_2, e_3, e_4\} \\
N_3 &= \text{span}\{e_i\}_{i=1}^{m_3} = \text{span}\{e_1, e_2, e_3, e_4, e_5\} = \mathscr{R}^5
\end{aligned}
$$

It is worth emphasizing that, although the nest structure is very useful for our proofs, the main results apply to the matrix algebras themselves, without reference to being endomorphisms (operators) of a free module.

”Orthogonal” and “Reflected” Nests

Despite there being no general notion of orthogonality in $\mathscr{R}^n$, we will use the notation $N_t^\perp = \operatorname{span}\{e_i\}_{i=m_t+1}^n$ for the subspace “perpendicular” to N_t. Continuing our previous example, we have $N_1^\perp = \operatorname{span}\{e_3, e_4, e_5\}$.

Borrowing Hilbert-space notation, define $N_t \ominus N_{t-1} = \operatorname{span}\{e_i\}_{i=m_{t-1}+1}^{m_t} = N_t \cap N_{t-1}^\perp$. Then, for $t \in \{1, \dots, k\}$ we have $\dim(N_t \ominus N_{t-1}) = m_t - m_{t-1} = n_t$. However, in the absence of a general notion of orthogonality, we avoid using “$\ominus$” for any subspaces that are not in the nest.

The set $\mathcal{N}^\perp := \{N_t^\perp : N_t \in \mathcal{N}\}$ is also a nest, but when ordered by the standard basis, the subspace inclusions are descending rather than ascending. For an upper block-triangular algebra $\mathcal{T}$, the corresponding nest $\mathcal{N}$ will have $e_1 \in N_t$ for all $1 \leq t \leq k$. For $\mathcal{N}^\perp$ we have $e_n \in N_t^\perp$ for all $1 \leq t \leq k$, so the corresponding algebra would be *lower* block-triangular.

The map $A \mapsto -A^\top$ is a Lie isomorphism from $\mathcal{T}(\mathcal{N})$ onto $\mathcal{T}(\mathcal{N}^\perp)$ since it is the negative of an anti-isomorphism. However, to turn it into a map between upper block-triangular algebras, we can conjugate with the matrix J. Define

$$\omega : T(n_1, \dots, n_k)(\mathscr{R}) \longrightarrow T(n_k, \dots, n_1)(\mathscr{R}),$$
$$\omega(A) = -JA^\top J.$$

Conjugating $A^\top$ with J flips the matrix A across its anti-diagonal, so the block order is reversed. The composite map ω is then a Lie isomorphism between upper block-triangular algebras. Clearly, this continues to be true for the special case of the triangular algebra $T(1, 1, \dots, 1)(\mathscr{R}) = T_n(\mathscr{R})$.

The map ω is self-inverse, preserves the bracket, and may be expressed by its explicit action on basis matrices:

$$\omega^2(A) = \omega(-JA^\top J) = -J(-JAJ)J = A \ ,$$

$$\begin{aligned}\omega([A,B]) &= -J(AB - BA)^\top J \\ &= -J(B^\top A^\top)J + J(A^\top B^\top)J \\ &= -(-JB^\top J)(-JA^\top J) + (-JA^\top J)(-JB^\top J) \\ &= [\omega(A), \omega(B)] \ ,\end{aligned}$$

$$\omega(\mathtt{E}_{ij}) = -\mathtt{E}_{(n+1-j),(n+1-i)} \ .$$

Let $\mathcal{N}$ be the nest associated with $\mathcal{T}(\mathcal{N}) = T(n_1, \ldots, n_k)(\mathscr{R})$, for the subspaces $N_t = \text{span}\{e_i\}_{i=1}^{m_t}$. We then define $N_t^\angle = \text{span}\{e_i\}_{i=1}^{n-m_{k-t}}$ and denote the "reflected" nest as $\mathcal{N}^\angle = \{N_t^\angle : N_t \in \mathcal{N}\}$, which is associated with $T(n_k, \ldots, n_1)(\mathscr{R}) =: \mathcal{T}(\mathcal{N}^\angle)$.

We may then write this Lie isomorphism as $\omega : \mathcal{T}(\mathcal{N}) \to \mathcal{T}(\mathcal{N}^\angle)$. We note that in the $T_n(\mathscr{R})$ case, ω is a Lie automorphism. This will be true in any case where the block structure is "symmetric" so that $n_t = n_{k-(t-1)}$ for $1 \le t \le k$ (or using the other index, $m_t = n - m_{k-t}$).

Using the ongoing example, if $\mathcal{T}(\mathcal{N}) = T(2,2,1)(\mathscr{R})$ then $\mathcal{T}(\mathcal{N}^\angle) = T(1,2,2)(\mathscr{R})$.

For the rest of the paper, we will assume that all the nests are "ascending" so that the corresponding algebras are *upper* block-triangular.

Projections

We define the projection P_t to act as the identity on N_t and zero on $N_t^\perp$. In other words, P_t will have ones along the diagonal of the first t diagonal blocks. If $\mathcal{N} = \{N_t\}_{t=0}^k$, we use the notation $\mathcal{P}(\mathcal{N}) = \{P_t\}_{t=0}^k$ for the collection of corresponding projections. In explicit terms, let $P_0 = 0$ and

$$P_t = \sum_{i=1}^{m_t} \mathtt{E}_{ii} = \sum_{i=1}^{t} I_{ii} \ .$$

Since $m_k = n$, we see that $P_k = I$ and $I - P_t = \sum_{i=m_t+1}^{n} \mathtt{E}_{ii} = \sum_{i=t+1}^{k} I_{ii}$. This latter projection acts as the identity on $N_t^{\perp}$.

Continuing with the example $T(2,2,1)(\mathscr{R})$, we have $m_2 = n_1 + n_2 = 2 + 2 = 4$, so

$$P_2 = \left[\begin{array}{cc|cc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Along with these standard projections, which would be called "orthogonal" over an inner product space, more general idempotent elements will be important to our arguments. We will use the notation $\mathcal{E}(\mathcal{N}) := \{A \in \mathcal{T}(\mathcal{N}) : A^2 = A\}$ to denote this larger set of idempotents, and define the nontrivial idempotents as $\mathcal{E}^0(\mathcal{N}) := \mathcal{E}(\mathcal{N}) \setminus \{0, I\}$. We note that, unlike the projections $\mathcal{P}(\mathcal{N})$, the range of a general $A \in \mathcal{E}(\mathcal{N})$ may not be one of the nest subspaces in $\mathcal{N}$. A simple example of this would be $I - P$ for some $P \in \mathcal{P}(\mathcal{N})$.

Traces

The standard trace on a matrix is the sum of its diagonal entries. For $A = [a_{ij}] \in M_n(\mathscr{R})$ we use the notation $\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}$. The trace is an $\mathscr{R}$-linear map $\mathrm{tr} : M_n(\mathscr{R}) \to \mathscr{R}$ and, following vector space terminology, we refer to such maps as functionals. One of the defining characteristics of the trace is that it sends commutators to zero: $\mathrm{tr}([A,B]) = 0$ or $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ for all $A, B \in M_n(\mathscr{R})$. This makes the trace invariant under cyclic permutations, for example $\mathrm{tr}(ABC) = \mathrm{tr}(CAB)$.

Lemma 2.3.1 (Generalized Traces)**.** *Let $\mathcal{T} = T(n_1, \ldots, n_k)(\mathscr{R})$ be an upper block-triangular algebra over a commutative ring $\mathscr{R}$. If $\tau : \mathcal{T} \to \mathscr{R}$ is an arbitrary $\mathscr{R}$-linear functional that sends commutators to zero, then for all $A \in \mathcal{T}$ we have $\tau(A) = \mathrm{tr}(D_\beta A)$ where $D_\beta = \sum_{i=1}^{k} \beta_i I_{ii}$ is a diagonal matrix that is constant on each block, with $\beta_i \in \mathscr{R}$ for $1 \le i \le k$.*

In particular, when there is only one block and $\mathcal{T} = T(n)(\mathscr{R}) = M_n(\mathscr{R})$, then D_β is a scalar matrix βI for some $\beta \in \mathscr{R}$. At the other extreme, when each block is one dimensional so that $\mathcal{T} = T(1, 1, \ldots, 1)(\mathscr{R}) = T_n(\mathscr{R})$ is the upper-triangular algebra, then $D_\beta = \sum_{i=1}^n \beta_i \mathtt{E}_{ii}$ may have different entries at each diagonal position.

Proof. Off-diagonal basis matrices can be written as commutators, $\mathtt{E}_{ij} = [\mathtt{E}_{ii}, \mathtt{E}_{ij}]$ for $i \neq j$, so we see that $\tau(\mathtt{E}_{ij}) = 0$. If we define $\beta_i := \tau(\mathtt{E}_{ii})$ for $1 \leq i \leq n$ and denote $A = [a_{ij}]$, then by linearity we have $\tau(A) = \sum_{i=1}^n \beta_i a_{ii}$. Defining the diagonal matrix $D_\beta := \mathrm{diag}(\beta_1, \ldots, \beta_n)$, we may write $\tau(A) = \mathrm{tr}(D_\beta A)$.

We examine the conditions on D_β so that $\tau([A, B]) = \mathrm{tr}(D_\beta[A, B]) = 0$ for all $A, B \in \mathcal{T}$. For $i < j$ we have

$$0 = \tau([\mathtt{E}_{ij}, \mathtt{E}_{ji}]) = \mathrm{tr}(D_\beta[\mathtt{E}_{ij}, \mathtt{E}_{ji}]) = \mathrm{tr}(D_\beta(\mathtt{E}_{ii} - \mathtt{E}_{jj})) = \beta_i - \beta_j.$$

However, E_{ji} is in $T(n_1, \ldots, n_k)(\mathscr{R})$ if and only if E_{ij} has its nonzero entry in one of the diagonal blocks, so that E_{ji} has its nonzero entry in the same block.

This means that $\beta_i = \beta_j$ within each block $m_{t-1} \leq i < j \leq m_t$, as we range over the blocks $1 \leq t \leq k$. By the restriction on $\mathtt{E}_{ji}$ for $i < j$, the constant for each block is independent from the constants for the other blocks. By re-indexing we may write the diagonal matrix $D_\beta = \sum_{i=1}^k \beta_i I_{ii}$ for $\beta_i \in \mathscr{R}$ for $1 \leq i \leq k$, as desired. □

In the case of the upper triangular matrices $T_n(\mathscr{R})$, we have another useful property of the trace.

Lemma 2.3.2. *Let $T_n(\mathscr{R})$ be an upper triangular algebra over a commutative ring $\mathscr{R}$. For all $A, B, C \in T_n(\mathscr{R})$ we have*

$$\mathrm{tr}(ABC) = \mathrm{tr}(ACB) = \mathrm{tr}(BAC) = \mathrm{tr}(BCA) = \mathrm{tr}(CAB) = \mathrm{tr}(CBA).$$

In other words, in $T_n(\mathscr{R})$ the trace is invariant under any permutation of matrices, not just cyclic ones.

Proof. Denote $A = [a_{ij}]$, $B = [b_{ij}]$ and $C = [c_{ij}]$ in $T_n(\mathscr{R})$ and define $AB = [s_{ij}]$ so that $s_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$. In $T_n(\mathscr{R})$ all entries are zero except for $i \leq k$ and $k \leq j$. This means $s_{ij} = \sum_{k=i}^j a_{ik} b_{kj}$, and so the diagonal elements take the form $s_{ii} = \sum_{k=i}^i a_{ik} b_{ki}$, or simply $s_{ii} = a_{ii} b_{ii}$. Extending this observation, the diagonal entries of ABC will be $a_{ii}b_{ii}c_{ii}$. The trace only "sees" these diagonal elements, so it is "blind" to any permutation of elements in $T_n(\mathscr{R})$. ☐

This gives another way to show that in $T_n(\mathscr{R})$ we have $\tau(AB) = \mathrm{tr}(D_\beta AB) = \mathrm{tr}(D_\beta BA) = \tau(BA)$ for any $D_\beta := \mathrm{diag}(\beta_1, \ldots, \beta_n)$. In all these cases, we call an $\mathscr{R}$-linear functional that sends commutators to zero a **generalized trace**.

In the proof of Lemma 2.3.2, the diagonal $a_{ii}b_{ii}$ of the matrix product $AB \in T_n(\mathscr{R})$ also shows that the set of strictly upper triangular matrices $\mathcal{U}$ is both an associative ideal and a Lie ideal in $T_n(\mathscr{R})$.

The Center

Both the Lie product and the trace are intimately connected to questions of commutativity. The **center** $Z(\mathcal{A})$ of an algebra $\mathcal{A}$ is the set of all elements in the algebra that commute with all other elements.

Lemma 2.3.3. *A matrix $A = [a_{ij}] \in M_n(\mathscr{R})$ commutes with $\mathtt{E}_{rs}$ if and only if $a_{ir} = 0$ for $i \neq r$, $a_{sj} = 0$ for $j \neq s$ and $a_{rr} = a_{ss}$.*

Proof. Intuitively, multiplication on the right by $\mathtt{E}_{rs}$ takes the r column of A and puts it into the s column, leaving zeros everywhere else. Similarly, multiplication on the left by $\mathtt{E}_{rs}$ takes the s row of A and puts it into the r row, leaving zeros everywhere else. Hence, $a_{rr} = a_{ss}$ and A must have zeros elsewhere in the r column and s row. Explicitly,

$$\begin{aligned} A\mathtt{E}_{rs} &= \mathtt{E}_{rs}A, \\ \left(\sum_{ij} a_{ij}\mathtt{E}_{ij}\right)\mathtt{E}_{rs} &= \mathtt{E}_{rs}\left(\sum_{ij} a_{ij}\mathtt{E}_{ij}\right), \\ \sum_i a_{ir}\mathtt{E}_{is} &= \sum_j a_{sj}\mathtt{E}_{rj}\,, \end{aligned}$$

which holds if and only if $a_{ir} = 0$ for $i \neq r$, $a_{sj} = 0$ for $j \neq s$ and $a_{rr} = a_{ss}$. ☐

Corollary 2.3.4. *If a sub-algebra of $M_n(\mathscr{R})$ contains the upper (or lower) triangular matrices, the center consists of scalar multiples of the identity. Symbolically,*

$$T_n \subseteq \mathcal{A} \subseteq M_n(\mathscr{R}) \quad \Rightarrow \quad Z(\mathcal{A}) = \{\lambda I : \lambda \in \mathscr{R}\}.$$

Proof. This follows directly from the previous Lemma and the fact that $\mathtt{E}_{ij} \in T_n \subseteq \mathcal{A}$ for all $i \le j$ (or $i \ge j$ for lower triangular matrices). □

2.4 Isomorphism and Automorphism Theorems

For an associative $\mathscr{R}$-algebra $\mathcal{A}$, if $\varphi \in \mathrm{Aut}(\mathcal{A})$ is an automorphism and there exists an invertible $S \in \mathcal{A}$ such that $\varphi(A) = S^{-1}AS$ for all $A \in \mathcal{A}$, then we call φ an **inner automorphism**. We note that some authors will write inner automorphisms with the inverse notated on the right of A, so that $\varphi(A) = SAS^{-1}$. Clearly both conventions are equivalent.

For any algebraic structure, this type of explicit characterization of automorphisms can be a powerful tool. One of the most well-known isomorphism theorems is the Skolem-Noether Theorem, for simple rings (Hungerford [3], Theorem IX.6.7):

Theorem 2.4.1 (Skolem-Noether for Rings)**.** *Let $\mathcal{R}$ be a simple left Artinian ring and let $\mathcal{K}$ be the center of $\mathcal{R}$ (so that $\mathcal{R}$ is a $\mathcal{K}$-algebra). Let $\mathcal{A}$ and $\mathcal{B}$ be finite dimensional simple $\mathcal{K}$-subalgebras of $\mathcal{R}$ that contain $\mathcal{K}$. If $\varphi : \mathcal{A} \to \mathcal{B}$ is an associative $\mathcal{K}$-algebra isomorphism that leaves $\mathcal{K}$ fixed elementwise, then φ extends to an inner automorphism of $\mathcal{R}$.*

Our primary application of this theorem is to associative automorphisms of the full matrix algebra, with $\mathcal{A} = \mathcal{B} = \mathcal{R} = M_n(\mathbb{K})$ for a field $\mathbb{K}$.

Corollary 2.4.2 (Skolem-Noether for $M_n(\mathbb{K})$)**.** *Let $\mathbb{K}$ be a field. Then every associative $\mathbb{K}$-algebra automorphism of $M_n(\mathbb{K})$ is inner.*

Proof. In this case, $M_n(\mathbb{K})$ is Artinian ([3], Corollary VIII.1.12) and it is a straightforward exercise to show that $M_n(\mathbb{K})$ is a simple ring ([3], Exercise III.2.9). As shown in Corollary 2.3.4, the center of $M_n(\mathbb{K})$ is the set of scalar matrices, which is isomorphic to $\mathbb{K}$. Any unital $\mathbb{K}$-algebra automorphism $\varphi : M_n(\mathbb{K}) \to M_n(\mathbb{K})$ must fix the identity matrix and so φ leaves the center fixed elementwise. □

Generalizing this theorem, Isaacs [4] examined automorphisms of full matrix algebras $M_n(\mathscr{R})$ where $\mathscr{R}$ is a general commutative unital ring. Among other results, in the case where $\mathscr{R}$ is a UFD, a similar theorem is obtained ([4], Corollary 15):

Theorem 2.4.3 (Isaacs)**.** *Let $\mathscr{R}$ be a UFD. Then every associative $\mathscr{R}$-algebra automorphism of $M_n(\mathscr{R})$ is inner.*

Along with other results, Cheung [2] then extended this characterization to block-triangular matrix algebras over a UFD ([2] Corollary 5.4.10):

Theorem 2.4.4 (Cheung)**.** *Let $\mathscr{R}$ be a UFD. Then every associative $\mathscr{R}$-algebra automorphism of $T(n_1, \ldots, n_k)(\mathscr{R})$ is inner.*

When dealing with upper-triangular matrices, Kezlan [5] showed a similar result where the only restriction on the ring $\mathscr{R}$ is that it is commutative and unital.

Theorem 2.4.5 (Kezlan)**.** *Let $\mathscr{R}$ be a commutative unital ring. Then every associative $\mathscr{R}$-algebra automorphism of $T_n(\mathscr{R})$ is inner.*

All these theorems are important and we will rely on the last two results in particular.

Lie Isomorphisms

From the construction of the bracket product, any inner automorphism of a matrix algebra $\mathcal{A}$ will also be a Lie automorphism. However, preservation of the Lie product is not as restrictive a condition as preservation of the associative multiplication. As such, we might expect that there are Lie automorphisms that are not inner.

Martindale [8] characterized Lie isomorphisms for simple and prime rings. We adapt a version for simple rings here:

Theorem 2.4.6 (Martindale)**.** *Let $\mathcal{S}$ and $\mathcal{R}$ be simple unital rings, not of characteristic 2 or 3, such that $\mathcal{S}$ contains two nonzero orthogonal idempotents whose sum is 1. If $\varphi : \mathcal{S} \to \mathcal{R}$ is a Lie isomorphism, then $\varphi = \sigma + \tau$ where σ is either an associative isomorphism or the negative of an associative anti-isomorphism of $\mathcal{S}$ onto $\mathcal{R}$, and τ is an additive mapping of $\mathcal{S}$ into the center of $\mathcal{R}$ which maps commutators to zero.*

As was mentioned, the full matrix algebra $M_n(\mathbb{K})$ is simple over a field $\mathbb{K}$. However, for triangular or block-triangular matrices over a general commutative ring, the algebra is not a simple or prime ring. Hence, Martindale's results cannot be applied directly. With additional work, however, we can show that the building blocks of the isomorphisms remain broadly the same.

For a matrix version of the additive mapping τ mentioned above, we will use a generalized trace. By definition, the Lie product vanishes under a generalized trace, so if we define $\psi(A) = A + \tau(A)I$ for $A \in \mathcal{A}$, then ψ is a Lie homomorphism,

$$\begin{aligned}[\psi(A), \psi(B)] &= [A + \tau(A)I, B + \tau(B)I] = [A, B] \\ &= [A, B] + \tau([A, B])I \\ &= \psi([A, B]).\end{aligned}$$

If we impose the condition that $1+\tau(I) \in \mathscr{R}^{\times}$, then $\psi(I) = (1+\tau(I))I$ is still invertible in $\mathcal{A}$. In Section 3.2 we will show that this makes ψ a Lie automorphism of $\mathcal{A}$, which we shall call a **trace automorphism**.

An (associative) anti-homomorphism of algebras is a linear map that reverses the order of the multiplication. So, if $\varphi : \mathcal{A} \to \mathcal{B}$ is an anti-homomorphism then $\varphi(AB) = \varphi(B)\varphi(A)$ for all $A, B \in \mathcal{A}$. The transpose map $\varphi(A) = A^{\top}$ is a familiar example for matrix algebras.

Due to the construction of the Lie product, the negative of an anti-homomorphism is a Lie homomorphism: Composing with a sign-change, define $\varphi : \mathcal{A} \to \mathcal{B}$ so that $\varphi(AB) = -\varphi(B)\varphi(A)$ for all $A, B \in \mathcal{A}$. Then,

$$\varphi([A, B]) = \varphi(AB) - \varphi(BA) = -\varphi(B)\varphi(A) + \varphi(A)\varphi(B) = [\varphi(A), \varphi(B)].$$

In the block-triangular case, we saw that negative transpose map $A \mapsto -A^{\top}$ is a Lie isomorphism from $\mathcal{T}(\mathcal{N})$ onto $\mathcal{T}(\mathcal{N}^{\perp})$ and the negative anti-diagonal reflection map $\omega(A) = -JA^{\top}J$ is a Lie isomorphism $\omega : \mathcal{T}(\mathcal{N}) \to \mathcal{T}(\mathcal{N}^{\angle})$, or a Lie automorphism when $\mathcal{T}(\mathcal{N}) = T_n(\mathscr{R})$. In Section 3.3 we will term this a **reflection isomorphism**.

The primary goal of this work is to show that, over appropriate rings, Lie isomorphisms of triangular and block-triangular matrix algebras are composed of these three building blocks: inner automorphisms, negative anti-isomorphisms, and maps involving a generalized trace.

Lie automorphisms of $T_n(\mathscr{R})$

With the restriction that $\mathscr{R}$ is a connected ring, Đoković [9] characterized the Lie automorphisms φ of $T_n(\mathscr{R})$ as a group, rather than explicitly writing the form they take. In [1], Cao was able to remove Đoković's restriction that $\mathscr{R}$ is connected.

We will show in Theorem 3.6.1 that for all $A \in T_n(\mathscr{R})$, in concrete terms this means φ can be written as

$$\varepsilon(SAS^{-1} + \operatorname{tr}(D_\alpha A)I) - (1-\varepsilon)(SJA^\top JS^{-1} + \operatorname{tr}(D_\beta A)I).$$

where $S \in \mathfrak{T}_n(\mathscr{R})$ is invertible, $\varepsilon \in \mathscr{R}$ is an idempotent, and the generalized traces have the condition $D_\beta = JD_\alpha J$ and $1 + \operatorname{tr}(D_\beta) = 1 + \operatorname{tr}(D_\alpha) \in \mathscr{R}^\times$.

Independently in [6], Marcoux and Sourour used a different approach to show a similar result for $T_n(\mathbb{K})$ where $\mathbb{K}$ is a field. They also classified more general commutativity-preserving linear maps on such algebras.

In Chapter 3 we integrate the arguments of Đoković and Cao to give a combined proof of the general result for $T_n(\mathscr{R})$ where $\mathscr{R}$ is a commutative unital ring.

Lie isomorphisms of $T(n_1, \ldots, n_k)(\mathscr{R})$

In [7], Marcoux and Sourour show a very similar characterization for Lie isomorphisms of nest algebras on a complex separable Hilbert space.

In Chapter 4 we adapt the broad approach used by these authors to characterize Lie isomorphisms of finite-dimensional block-triangular matrix algebras $T(n_1, \ldots, n_k)(\mathscr{R})$ over a UFD $\mathscr{R}$, not of characteristic 2 or 3 (Theorem 4.1.1). Although we follow similar steps in this paper, many of the proofs are changed significantly, particularly where Marcoux and Sourour rely on Hilbert space results.

Chapter 3

Triangular Algebras

Let $\mathscr{R}$ be a (unital) commutative ring and let $T_n(\mathscr{R}) \subseteq M_n(\mathscr{R})$ denote the algebra of upper triangular matrices for $n \geq 1$. Since the ring will be assumed to not change, we will often shorten our notation from $T_n(\mathscr{R})$ to T_n.

In this chapter we combine the results of Đoković [9] and Cao [1] to characterize the Lie automorphisms of a triangular matrix algebra, $T_n(\mathscr{R})$. Đoković described the group of automorphisms of Lie algebras of upper triangular matrices over a connected commutative ring, while Cao was able to show that the connectedness condition on the ring may be removed. In both cases, the authors characterized the Lie automorphisms of $T_n(\mathscr{R})$ as a group, rather than explicitly writing the form they take. We will follow this approach, while also seeing how the maps are realized in concrete form.

Following the comments in Chapter 2, we seek to show that the building blocks of $\mathrm{Aut}_L(T_n)$ are inner automorphisms, negative anti-isomorphisms, and maps using a generalized trace. Hence, in this chapter our goal is to build three subgroups of $\mathrm{Aut}_L(T_n)$ and show that they generate all of the Lie automorphisms.

3.1 Inner Automorphisms

From Theorem 2.4.5 we have that every associative $\mathscr{R}$-algebra automorphism of T_n is inner. In other words, if $\varphi : T_n \to T_n$ is an associative automorphism, then there exists $S \in T_n^\times$ such that $\varphi(A) = SAS^{-1}$, for all $A \in T_n$. We add a subscript φ_S to show this relation. For inner automorphisms in this chapter we use the convention where the inverse is denoted on the right of A, to match the notation used in the papers by Đoković and Cao.

Since any associative automorphism of T_n will induce a Lie-automorphism of T_n, we define the subgroup G_0 of $\mathrm{Aut}_L(T_n)$ to be all such inner automorphisms:

$$G_0 = \{\varphi \in \mathrm{Aut}_L(T_n) : \exists S \in T_n^\times,\, \varphi(A) = SAS^{-1},\, \forall A \in T_n\}.$$

If we define the map $\Phi : T_n^\times \to G_0$ to be $\Phi(S) = \varphi_S$, we see that it is a surjective group homomorphism.

By Corollary 2.3.4, the center of T_n consists of scalar multiples of the identity. If the scalar is invertible, conjugation by the scalar matrix is trivial. It follows that the kernel of the group homomorphism $\Phi : T_n^\times \to G_0$ is

$$\ker(\Phi) = \{\lambda I : \lambda \in \mathscr{R}^\times\} \cong (\mathscr{R}^\times, \cdot).$$

3.2 Trace Automorphisms

We construct a second subgroup G_1 of $\mathrm{Aut}_L(T_n)$ using a generalized trace. In the following, we omit the summation bounds where they are unambiguous.

Let $\mathcal{S} \subseteq \mathscr{R}^n$ consist of all n-tuples $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ such that $1 + \sum \alpha_j \in \mathscr{R}^\times$. Using such an n-tuple α, we may define the diagonal matrix

$$D_\alpha := \mathrm{diag}(\alpha_1, \alpha_2, \ldots, \alpha_n),$$

and use this to build a generalized trace $\tau_\alpha : T_n \to \mathscr{R}$ where $\tau_\alpha(A) = \mathrm{tr}(D_\alpha A)$ for $A \in T_n$. For each such α, we then define the function $\psi_\alpha : T_n \to T_n$ to be

$$\begin{aligned}\psi_\alpha(A) &:= A + \tau_\alpha(A) I \\ &= A + \mathrm{tr}(D_\alpha A) I,\end{aligned}$$

for $A \in T_n$. This may also be written in terms of its action on the basis matrices,

$$\psi_\alpha(\mathtt{E}_{ij}) = \mathtt{E}_{ij} + \delta_{ij}\alpha_i I\,, \qquad i \le j\,.$$

We observe that if $1 + \sum \alpha_j \in \mathscr{R}^\times$ then $\psi_\alpha(I) = (1 + \tau_\alpha(I))I$ is invertible in T_n.

Define the set $G_1 = \{\psi_\alpha : \alpha \in \mathcal{S}\}$. We note that $0 = (0, 0, \ldots, 0) \in \mathcal{S}$ and $\psi_0(A) = A$, so in this notation ψ_0 is equal to the identity automorphism $\varphi_I \in G_0$.

We show that G_1 is closed under composition. Let $\alpha, \beta \in \mathcal{S}$ and $A \in T_n$. Since the trace is linear,

$$\begin{aligned}\psi_\beta \circ \psi_\alpha(A) &= \psi_\beta(A + \operatorname{tr}(D_\alpha A)I) \\ &= A + \operatorname{tr}(D_\alpha A)I + \operatorname{tr}(D_\beta[A + \operatorname{tr}(D_\alpha A)I]) \\ &= A + \operatorname{tr}(D_\beta A)I + (1 + \operatorname{tr}(D_\beta I))\operatorname{tr}(D_\alpha A)I \\ &= A + \operatorname{tr}(D_\beta A)I + (1 + \textstyle\sum \beta_j)\operatorname{tr}(D_\alpha A)I,\end{aligned}$$

so we may define $\gamma_i = \beta_i + (1 + \sum \beta_j)\alpha_i$ and $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n)$, to then write $\psi_\beta \circ \psi_\alpha(A) = \psi_\gamma(A)$. From the above, we see that

$$1 + \textstyle\sum \gamma_i = 1 + \sum \beta_i + (1 + \sum \beta_j)\sum \alpha_i = (1 + \sum \beta_i)(1 + \sum \alpha_i),$$

which is the product of two invertible elements and hence $(1 + \sum \gamma_i) \in \mathscr{R}^\times$ and $\gamma \in \mathcal{S}$. Additionally, we may define $\lambda_i = -\alpha_i(1 + \sum \alpha_j)^{-1}$ so that

$$1 + \textstyle\sum \lambda_i = 1 - \sum \alpha_j(1 + \sum \alpha_j)^{-1}.$$

Then,

$$(1 + \textstyle\sum \lambda_i)(1 + \sum \alpha_i) = (1 + \sum \alpha_i) - \sum \alpha_j(1 + \sum \alpha_j)^{-1}(1 + \sum \alpha_i) = 1.$$

Thus, $1 + \sum \lambda_i = (1 + \sum \alpha_i)^{-1}$ and so $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathcal{S}$. This shows that $G_1 = \{\psi_\alpha : \alpha \in \mathcal{S}\}$ is a group. Each $\psi_\alpha \in G_1$ also preserves the Lie product:

$$\begin{aligned}[\psi_\alpha(A), \psi_\alpha(B)] &= [A + \tau_\alpha(A)I, B + \tau_\alpha(B)I] = [A, B] \\ &= AB + \tau_\alpha(AB)I - BA - \tau_\alpha(BA)I \\ &= \psi_\alpha(AB) - \psi_\alpha(BA) \\ &= \psi_\alpha([A, B]),\end{aligned}$$

for any $A, B \in T_n$. Thus, G_1 is a subgroup of $\operatorname{Aut}_L(T_n)$.

While not directly applicable to the rest of our argument, Đoković also points out that G_1 is isomorphic to the semidirect product $\mathscr{R}^{n-1} \rtimes \mathscr{R}^{\times}$, where $\mathscr{R}^{\times}$ acts on $\mathscr{R}^{n-1}$ by multiplication.

For $v, w \in \mathscr{R}^{n-1}$ and $\lambda, \mu \in \mathscr{R}^{\times}$, define

$$(w, \mu) \cdot (v, \lambda) = (w + \mu v, \mu\lambda).$$

If $v = (\alpha_1, \ldots, \alpha_{n-1}) \in \mathscr{R}^{n-1}$ and $\lambda \in \mathscr{R}^{\times}$ then $\alpha = (\alpha_1, \ldots, \alpha_{n-1}, \lambda - 1 - \sum_{i=1}^{n-1} \alpha_i) \in \mathcal{S}$. We may then define the map $(v, \lambda) \mapsto \psi_\alpha$ to be the desired isomorphism.

3.3 Reflection Automorphisms

To construct the third subgroup G_2 of $\mathrm{Aut}_L(T_n)$ we first use the antidiagonal matrix J and the map $\omega : T_n \to T_n$ defined as

$$\omega(A) = -JA^{\top}J,$$

for all $A \in T_n$. As we noted in Chapter 2, $\omega^2 = id$ and ω preserves the Lie product, so $\omega \in \mathrm{Aut}_L(T_n)$. Since it is the negative of a reflection across the antidiagonal, ω preserves upper triangular matrices.

Let ε be an idempotent element of $\mathscr{R}$. Using ω, we define a new map $\omega_\varepsilon : T_n \to T_n$ such that

$$\omega_\varepsilon(A) = \varepsilon A + (1 - \varepsilon)\omega A,$$

for all $A \in T_n$. We see that $\omega_1 = id$ and $\omega_0 = \omega$. Also,

$$\begin{aligned}
\omega_\varepsilon^2(A) &= \omega_\varepsilon(\varepsilon A + (1 - \varepsilon)\omega A) \\
&= \varepsilon(\varepsilon A + (1 - \varepsilon)\omega A) + (1 - \varepsilon)\omega(\varepsilon A + (1 - \varepsilon)\omega A) \\
&= \varepsilon A + (1 - \varepsilon)A = A.
\end{aligned}$$

Hence, $\omega_\varepsilon^2 = id$.

We define the set

$$G_2 = \{\omega_\varepsilon : \varepsilon \in \mathscr{R} \text{ idempotent}\},$$

and show that G_2 is closed under composition. If $\varepsilon, \eta \in \mathscr{R}$ are idempotents, then

$$\begin{aligned}(\varepsilon - \eta)^4 &= \varepsilon^4 - 4\varepsilon^3\eta + 6\varepsilon^2\eta^2 - 4\varepsilon\eta^3 + \eta^4 \\ &= \varepsilon^2 - 2\varepsilon\eta + \eta^2 \\ &= (\varepsilon - \eta)^2,\end{aligned}$$

so $(\varepsilon - \eta)^2$ and $1 - (\varepsilon - \eta)^2$ are also idempotents. Then,

$$\begin{aligned}\omega_\eta \circ \omega_\varepsilon(A) &= \eta(\varepsilon A + (1-\varepsilon)\omega A) + (1-\eta)\omega(\varepsilon A + (1-\varepsilon)\omega A) \\ &= (1 - \eta + 2\varepsilon\eta - \varepsilon)A + (\eta - 2\varepsilon\eta + \varepsilon)\omega A \\ &= (1 - (\varepsilon - \eta)^2)A + (\varepsilon - \eta)^2\omega A \\ &= \omega_{1-(\varepsilon-\eta)^2} A.\end{aligned}$$

Along with showing closure under composition, this confirms again that $\omega_\varepsilon^2 = id$. Therefore, G_2 is a subgroup of $\mathrm{Aut}_L(T_n)$.

3.4 Relations between the Automorphism Subgroups

Gradually assembling the building blocks of $\mathrm{Aut}_L(T_n)$, we next show that the subgroup generated by G_0 and G_1 is their (internal) direct product $G_0 \times G_1$.

Proposition 3.4.1. *The groups G_0 and G_1 commute element-wise and $G_0 \cap G_1 = \{id\}$.*

Proof. By the trace result in Lemma 2.3.2 for T_n, the trace is invariant under any permutation of matrices. Hence, for any $S \in T_n^\times$ and $\alpha \in \mathcal{S}$, so that $\varphi_S \in G_0$ and $\psi_\alpha \in G_1$, we have

$$\begin{aligned}\psi_\alpha \circ \varphi_S(A) &= SAS^{-1} + \mathrm{tr}(D_\alpha SAS^{-1})I \\ &= SAS^{-1} + \mathrm{tr}(D_\alpha A)I \\ &= S(A + \mathrm{tr}(D_\alpha A)I)S^{-1} \\ &= \varphi_S \circ \psi_\alpha(A),\end{aligned}$$

for all $A \in T_n$, which proves the first assertion.

If $\psi_\alpha = \varphi_S$ for some $\alpha \in \mathcal{S}$ and $S \in T_n^\times$, then for any strictly upper triangular matrix $U \in \mathcal{U}$, we have $\mathrm{tr}(D_\alpha U) = 0$ Thus, for $i < j$,

$$\begin{aligned} \psi_\alpha(\mathtt{E}_{ij}) = \varphi_S(\mathtt{E}_{ij}) \quad &\Rightarrow \quad \mathtt{E}_{ij} + \mathrm{tr}(D_\alpha \mathtt{E}_{ij})I = S\mathtt{E}_{ij}S^{-1} \\ &\Rightarrow \quad \mathtt{E}_{ij}S = S\mathtt{E}_{ij}\,. \end{aligned}$$

The result on commuting matrices in Lemma 2.3.3, along with this extra restriction on the indices, gives that S is a scalar times the identity plus some upper right corner element; $S = \lambda I + \mu \mathtt{E}_{1n}$ for some $\lambda \in \mathscr{R}^\times$ and $\mu \in \mathscr{R}$.

In intuitive terms, for any $S \in T_n$ that commutes with every strictly upper triangular matrix, we can show that the diagonal elements must be equal and that almost all other entries must be zero. However, since we have excluded $\mathtt{E}_{11}$ and $\mathtt{E}_{nn}$ from our check, the intersection of first row and last column of S is not required to be zero.

Then,

$$\begin{aligned} \psi_\alpha(\mathtt{E}_{11}) = \varphi_S(\mathtt{E}_{11}) \quad &\Rightarrow \quad \mathtt{E}_{11} + \mathrm{tr}(D_\alpha \mathtt{E}_{11})I = S\mathtt{E}_{11}S^{-1} \\ &\Rightarrow \quad (\mathtt{E}_{11} + \alpha_1 I)S = S\mathtt{E}_{11} \\ &\Rightarrow \quad \lambda\alpha_1 I + \mu(1+\alpha_1)\mathtt{E}_{1n} = 0. \end{aligned}$$

This means that $\mu = \alpha_1 = 0$. Thus, φ_S is the identity automorphism, as asserted. □

We next show that G_2 normalizes G_0 and G_1:

Lemma 3.4.2.

(1) If $S \in T_n^\times$ and $\varepsilon \in \mathscr{R}$ is idempotent, then $\omega_\varepsilon \circ \varphi_S \circ \omega_\varepsilon = \varphi_C$ where $C = \varepsilon S + (1-\varepsilon)B$ and $B = J(S^{-1})^\top J \in T_n^\times$.

(2) If $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathcal{S}$ and $\varepsilon \in \mathscr{R}$ is idempotent, then $\omega_\varepsilon \circ \psi_\alpha \circ \omega_\varepsilon = \psi_\gamma$ where $\gamma = (\gamma_1, \ldots, \gamma_n) \in \mathcal{S}$ such that $\gamma_i = \varepsilon\alpha_i + (1-\varepsilon)\alpha_{n+1-i}$.

Proof. (1) We first show that for $S \in T_n^\times$ we have $\omega \circ \varphi_S \circ \omega = \varphi_B$ where $B = J(S^{-1})^\top J \in T_n^\times$. For $A \in T_n$, we have

$$\begin{aligned} \omega \circ \varphi_S \circ \omega(A) &= J(SJA^\top JS^{-1})^\top J \\ &= (J(S^{-1})^\top J)A(JS^\top J) \\ &= BAB^{-1}. \end{aligned}$$

As a product of idempotents we have $\varepsilon(1-\varepsilon) = 0$, so $C^{-1} = \varepsilon S^{-1} + (1-\varepsilon)B^{-1}$. Then, using $\varphi_S \circ \omega = \omega \circ \varphi_B$,

$$\begin{aligned} \omega_\varepsilon \circ \varphi_S \circ \omega_\varepsilon(A) &= \omega_\varepsilon(\varepsilon\varphi_S(A) + (1-\varepsilon)\varphi_S(\omega A)) \\ &= \omega_\varepsilon(\varepsilon\varphi_S(A) + (1-\varepsilon)\omega(\varphi_B(A))) \\ &= \varepsilon(\varepsilon\varphi_S(A) + (1-\varepsilon)\omega(\varphi_B(A))) + (1-\varepsilon)\omega(\varepsilon\varphi_S(A) + (1-\varepsilon)\omega(\varphi_B(A))) \\ &= \varepsilon\varphi_S(A) + (1-\varepsilon)\varphi_B(A), \quad \text{and} \end{aligned}$$

$$\begin{aligned} \varphi_C(A) &= (\varepsilon S + (1-\varepsilon)B)A(\varepsilon S^{-1} + (1-\varepsilon)B^{-1}) \\ &= \varepsilon SAS^{-1} + (1-\varepsilon)BAB^{-1} \\ &= \varepsilon\varphi_S(A) + (1-\varepsilon)\varphi_B(A), \end{aligned}$$

as desired.

(2) With γ defined where $\gamma_i = \varepsilon\alpha_i + (1-\varepsilon)\alpha_{n+1-i}$, we see that

$$\begin{aligned} (1 + \textstyle\sum\gamma_i) &= 1 + \varepsilon\textstyle\sum\alpha_i + (1-\varepsilon)\textstyle\sum\alpha_{n+1-i} \\ &= 1 + \varepsilon\textstyle\sum\alpha_i + (1-\varepsilon)\textstyle\sum\alpha_i \\ &= 1 + \textstyle\sum\alpha_i \, , \end{aligned}$$

so $\gamma \in \mathcal{S}$.

For $\alpha, \gamma \in \mathcal{S}$, the maps ψ_α and ψ_γ act trivially on strictly upper triangular matrices in $\mathcal{U}$. Hence, $\omega_\varepsilon \circ \psi_\alpha \circ \omega_\varepsilon$ also acts trivially on $\mathcal{U}$. Then, using the equation $\omega(\mathtt{E}_{ij}) =$

$-\mathtt{E}_{(n+1-j),(n+1-i)}$ gives,

$$\begin{aligned}
\omega_\varepsilon \circ \psi_\alpha \circ \omega_\varepsilon(\mathtt{E}_{ii}) &= \omega_\varepsilon \circ \psi_\alpha(\varepsilon \mathtt{E}_{ii} - (1-\varepsilon)\mathtt{E}_{n+1-i,n+1-i}) \\
&= \omega_\varepsilon(\varepsilon \mathtt{E}_{ii} - (1-\varepsilon)\mathtt{E}_{n+1-i,n+1-i} + (\varepsilon\alpha_i - (1-\varepsilon)\alpha_{n+1-i})I) \\
&= \varepsilon \mathtt{E}_{ii} + \varepsilon\alpha_i I + (1-\varepsilon)\mathtt{E}_{ii} + (1-\varepsilon)\alpha_{n+1-i} I \\
&= \mathtt{E}_{ii} + (\varepsilon\alpha_i + (1-\varepsilon)\alpha_{n+1-i})I \\
&= \psi_\gamma(\mathtt{E}_{ii}),
\end{aligned}$$

as desired. □

We now examine the relation between G_0, G_1 and G_2 depending on the dimension n.

Lemma 3.4.3.

(1) $G_0 = \{id\}$ *and* $G_2 \subseteq G_1$ *for* $n = 1$.

(2) $G_2 \subseteq G_0 \times G_1$ *for* $n = 2$.

(3) $(G_0 \times G_1) \cap G_2 = \{id\}$ *for* $n > 2$.

Proof. (1) For $n = 1$, the 'triangular' algebra is isomorphic to the underlying ring, $T_n \cong \mathscr{R}$. Conjugation by an invertible in $\mathscr{R}$ is trivial, so $G_0 = \{id\}$. For any idempotent $\varepsilon \in \mathscr{R}$, we have

$$\omega_\varepsilon(I) = \varepsilon I + (1-\varepsilon)\omega I = (2\varepsilon - 1)I.$$

We note that

$$(2\varepsilon - 1)^2 = 4\varepsilon^2 - 4\varepsilon + 1 = 1,$$

and we may write

$$2\varepsilon - 1 = 1 + 2(\varepsilon - 1).$$

Hence if we make $\alpha_1 = 2(\varepsilon - 1)$, then $\omega_\varepsilon = \psi_\alpha \in G_1$, where $\psi_\alpha(I) = I + \alpha_1 I = (1+\alpha_1)I$, and $1 + \alpha_1 \in \mathscr{R}^\times$. Thus, $G_2 \subseteq G_1$.

(2) Let $n = 2$ and take $\varepsilon \in \mathscr{R}$ to be any idempotent. Define $\alpha = (\varepsilon - 1, \varepsilon - 1)$ and $S = \mathtt{E}_{11} + (2\varepsilon - 1)\mathtt{E}_{22}$.

We note that $1 + \alpha_1 + \alpha_2 = 2\varepsilon - 1$ and so, by the comments above, $\alpha \in \mathcal{S}$. By a similar argument, we see that $S^2 = I$. We wish to show that $\omega_\varepsilon \in G_0 \times G_1$. Using S and α as defined,

$$\begin{aligned}
\psi_\alpha \circ \varphi_S \circ \omega_\varepsilon(\mathtt{E}_{11}) &= \psi_\alpha \circ \varphi_S(\varepsilon\mathtt{E}_{11} - (1-\varepsilon)\mathtt{E}_{22}) \\
&= \psi_\alpha(\varepsilon\mathtt{E}_{11} - (1-\varepsilon)\mathtt{E}_{22}) \\
&= (\varepsilon\mathtt{E}_{11} - (1-\varepsilon)\mathtt{E}_{22}) + ((\varepsilon-1)\varepsilon - (\varepsilon-1)(1-\varepsilon))I \\
&= \varepsilon\mathtt{E}_{11} - (1-\varepsilon)\mathtt{E}_{22} + (1-\varepsilon)I \\
&= \mathtt{E}_{11}.
\end{aligned}$$

By symmetry of the indices, a similar calculation shows

$$\psi_\alpha \circ \varphi_S \circ \omega_\varepsilon(\mathtt{E}_{22}) = \mathtt{E}_{22}.$$

Lastly,

$$\begin{aligned}
\psi_\alpha \circ \varphi_S \circ \omega_\varepsilon(\mathtt{E}_{12}) &= \psi_\alpha \circ \varphi_S(\varepsilon\mathtt{E}_{12} - (1-\varepsilon)\mathtt{E}_{12}) \\
&= \psi_\alpha \circ \varphi_S\ (2\varepsilon-1)\mathtt{E}_{12} \\
&= \psi_\alpha\ S(2\varepsilon-1)\mathtt{E}_{12}S \\
&= \psi_\alpha\ S(2\varepsilon-1)\mathtt{E}_{12}(\mathtt{E}_{11} + (2\varepsilon-1)\mathtt{E}_{22}) \\
&= \psi_\alpha(\mathtt{E}_{11} + (2\varepsilon-1)\mathtt{E}_{22})(2\varepsilon-1)^2\mathtt{E}_{12} \\
&= \psi_\alpha(\mathtt{E}_{12}) \\
&= \mathtt{E}_{12}\,.
\end{aligned}$$

Hence $\psi_\alpha \circ \varphi_S \circ \omega_\varepsilon = id$ and so $\omega_\varepsilon = (\psi_\alpha \circ \varphi_S)^{-1} \in G_0 \times G_1$.

(3) Let $n > 2$ and suppose $\omega_\varepsilon \in G_0 \times G_1$ for some idempotent $\varepsilon \in \mathscr{R}$. We may then write $\omega_\varepsilon = \psi_\alpha \circ \varphi_S$ for some $\alpha \in \mathcal{S}$ and $S \in T_n^\times$. By the construction of ω_ε, we have

$$\omega_\varepsilon(\mathtt{E}_{11}) = \varepsilon\mathtt{E}_{11} + (1-\varepsilon)\mathtt{E}_{nn}\,.$$

By the comments in the proof of Lemma 2.3.2 regarding multiplying diagonal entries of triangular matrices, $S\mathtt{E}_{11}S^{-1} = \mathtt{E}_{11} + U$ for some strictly upper triangular matrix $U \in \mathcal{U}$. Thus, we also have

$$\omega_\varepsilon(\mathtt{E}_{11}) = \psi_\alpha \circ \varphi_S(\mathtt{E}_{11}) = \psi_\alpha(\mathtt{E}_{11} + U) = \mathtt{E}_{11} + U + \alpha_1 I.$$

Since both these are equal, $U = 0$. In addition, since $n > 2$, we must have $\alpha_1 = 0$ and $\varepsilon = 1$. Hence $\omega_\varepsilon = \omega_1 = id$. □

3.5 Lie Ideals

The main theorem will rely on some basic results regarding the Lie ideal of strictly upper triangular matrices $\mathcal{U} \subseteq T_n$. Using the terminology from general Lie theory, the **derived algebra** of a Lie algebra $\mathcal{B}$, is the subalgebra $[\mathcal{B}, \mathcal{B}]$. The **lower central series** is then the sequence of subalgebras $\mathcal{B} \geq [\mathcal{B}, \mathcal{B}] \geq [\mathcal{B}, [\mathcal{B}, \mathcal{B}]] \geq \dots$. Since it is also an ideal, the derived algebra is sometimes called the derived ideal.

Proposition 3.5.1. *The derived ideal of T_n is the free $\mathscr{R}$-module $\mathcal{U}$ with basis $\{\mathtt{E}_{ij} : i < j\}$.*

In other words, $\mathcal{U} = [T_n, T_n] = \operatorname{span}\{[A, B] : A, B \in T_n\}$.

Proof. By the observation in the proof of 2.3.2, it follows that $[A, B]$ has zero diagonal for any $A, B \in T_n$. On the other hand, for $i < j$ we can write $\mathtt{E}_{ij} = \mathtt{E}_{ii}\mathtt{E}_{ij} = [\mathtt{E}_{ii}, \mathtt{E}_{ij}]$ and so any $U \in \mathcal{U}$ is a sum of commutators. □

Definition 3.5.2. Let $\mathcal{U}_1 = \mathcal{U}$, let $\mathcal{U}_2 = [\mathcal{U}, \mathcal{U}_1]$ and for $m > 2$, define

$$\mathcal{U}_m = [\mathcal{U}, \mathcal{U}_{m-1}] = \operatorname{span}\{[A, B] : A \in \mathcal{U}, B \in \mathcal{U}_{m-1}\}.$$

We will see below that $\mathcal{U}_{n-1} = \{\lambda\mathtt{E}_{1n} : \lambda \in \mathscr{R}\}$ and $\mathcal{U}_m = \{0\}$ for $m \geq n$.

This gives the lower central series of $\mathcal{U}$,

$$\{0\} = \mathcal{U}_n \leq \mathcal{U}_{n-1} \leq \cdots \leq \mathcal{U}_2 \leq \mathcal{U}_1 = \mathcal{U}.$$

Proposition 3.5.3. *For $m = 1, 2 \dots, n-1$, $\mathcal{U}_m$ is a free $\mathscr{R}$-module with basis $\{\mathtt{E}_{ij} : j - i \geq m\}$, and also a Lie ideal of T_n.*

Proof. The basis claim is clear from the indexing of the matrices. We show that $\mathcal{U}_m$ is a Lie ideal. Let $A \in T_n$ and $B \in \mathcal{U}_m$. We know that $b_{ij} = 0$ unless $j \geq i + m$ and $i \leq j - m$. Hence, for $C = AB - BA$ we have

$$c_{ij} = \sum_{k=i}^{j-m} a_{ik} b_{kj} - \sum_{k=i+m}^{j} b_{ik} a_{kj} \,.$$

For any c_{ij} such that $j < i + m$, both these sums are zero. Hence $[A, B] \in \mathcal{U}_m$, which is thus an ideal in T_n. □

Proposition 3.5.4. *$\mathcal{U}_{n-1}$ is $Z(\mathcal{U})$, the center of $\mathcal{U}$.*

Proof. By Proposition 3.5.3, we have $\mathcal{U}_{n-1} = \{\lambda \mathtt{E}_{1n} : \lambda \in \mathscr{R}\}$ and $AB = 0 = BA$ for $A \in \mathcal{U}_{n-1}$ and $B \in \mathcal{U}_m$ for any m. This shows that $\mathcal{U}_{n-1}$ is contained in the center $Z(\mathcal{U})$.

As in the proof of Proposition 3.4.1, let $j > i$ so that $\mathtt{E}_{ij} \in \mathcal{U}$, and suppose there exists some $S \in \mathcal{U}$ such that $\mathtt{E}_{ij} S = S \mathtt{E}_{ij}$. Again, the result on commuting matrices in Lemma 2.3.3, along with the extra restriction on the indices, gives that $S = \lambda I + \mu \mathtt{E}_{1n}$ for some $\lambda, \mu \in \mathscr{R}$. However, since $S \in \mathcal{U}$, then we must have $\lambda = 0$ and so $\mathcal{U}_{n-1} = Z(\mathcal{U})$. □

Proposition 3.5.5. *For any $\varphi \in \mathrm{Aut}_L(T_n)$, we have $\varphi(\mathcal{U}_m) = \mathcal{U}_m$.*

Proof. Since φ is a linear bijection and must preserve brackets, we have:

$$\begin{aligned}\varphi(\mathcal{U}) = \varphi([T_n, T_n]) &= \mathrm{span}\{\varphi[A, B] : A, B \in T_n\} \\ &= \mathrm{span}\{[\varphi(A), \varphi(B)] : A, B \in T_n\} \\ &= [\varphi(T_n), \varphi(T_n)] = [T_n, T_n] = \mathcal{U}.\end{aligned}$$

By induction, for $m \geq 2$ we then have

$$\varphi(\mathcal{U}_m) = \varphi([\mathcal{U}, \mathcal{U}_{m-1}]) = [\varphi(\mathcal{U}), \varphi(\mathcal{U}_{m-1})] = [\mathcal{U}, \mathcal{U}_{m-1}] = \mathcal{U}_m \,,$$

as required. □

We next define a sequence of ideals using matrices in T_n where only first-row entries are nonzero.

Definition 3.5.6. Let $\mathcal{A}$ be the free $\mathscr{R}$-module with basis $\{\mathtt{E}_{1j} : 1 \leq j \leq n\}$, and let $\mathcal{A}_m = \mathcal{A} \cap \mathfrak{U}_m$ be the free $\mathscr{R}$-module with basis $\{\mathtt{E}_{1j} : m < j \leq n\}$.

We note that $\mathcal{A}_m$ is an associative ideal of T_n. If we take $\mathtt{E}_{ij} \in T_n$ and $\mathtt{E}_{1k} \in \mathcal{A}_m$, then $\mathtt{E}_{ij}\mathtt{E}_{1k} = \mathtt{E}_{1k}$ if $i = 1 = j$ and zero otherwise. On the other hand $\mathtt{E}_{1k}\mathtt{E}_{ij} = \mathtt{E}_{1j}$ if $i = k$ and zero otherwise. In this case, since $j \geq i$, we have $j \geq k$ and so $\mathtt{E}_{1j} \in \mathcal{A}_m$.

We recall that every associative ideal will also be a Lie ideal. For $\varphi \in \mathrm{Aut}_L(T_n)$, the Lie bracket preservation property means that $\varphi(\mathcal{A}_m)$ is also a Lie ideal of T_n.

3.6 Main Results (Triangular)

The main result for this chapter is

Theorem 3.6.1. *Let $\mathscr{R}$ be a commutative ring. Then,*

(1) $\mathrm{Aut}_L(T_n) = G_1$ *for* $n = 1$*;*

(2) $\mathrm{Aut}_L(T_n) = G_0 \times G_1$ *for* $n = 2$*;*

(3) $\mathrm{Aut}_L(T_n) = (G_0 \times G_1) \rtimes G_2$ *for* $n > 2$*,*

where $\rtimes$ *denotes the (internal) semi-direct product of groups.*

Comments: Let $\varphi_S \in G_0$ and $\psi_\alpha \in G_1$, and take $id = \omega_1 \in G_2$. In Proposition 3.4.1, we saw that, for $A \in T_n$,

$$\begin{aligned}\psi_\alpha \circ \varphi_S \circ id(A) &= SAS^{-1} + \mathrm{tr}(D_\alpha SAS^{-1})I \\ &= SAS^{-1} + \mathrm{tr}(D_\alpha A)I.\end{aligned}$$

Alternatively, for $\omega = \omega_0 \in G_2$,

$$\begin{aligned}\psi_\alpha \circ \varphi_S \circ \omega(A) &= -SJA^\top JS^{-1} - \mathrm{tr}(D_\alpha SJA^\top JS^{-1})I \\ &= -SJA^\top JS^{-1} - \mathrm{tr}(JD_\alpha JA^\top)I \\ &= -SJA^\top JS^{-1} - \mathrm{tr}(D_\alpha^\top A)I \\ &= -SJA^\top JS^{-1} - \mathrm{tr}(D_\beta A)I,\end{aligned}$$

where β is the reverse ordering of the n-tuple α.

Combining the above two forms, for any idempotent $\varepsilon \in \mathscr{R}$, by linearity we have the general result

$$\begin{aligned}\psi_\alpha \circ \varphi_S \circ \omega_\varepsilon(A) &= \psi_\alpha \circ \varphi_S(\varepsilon A + (1-\varepsilon)\omega A)\\ &= \varepsilon(\psi_\alpha \circ \varphi_S)(A) + (1-\varepsilon)(\psi_\alpha \circ \varphi_S \circ \omega)(A)\\ &= \varepsilon(SAS^{-1} + \operatorname{tr}(D_\alpha A)I) - (1-\varepsilon)(SJA^\top JS^{-1} + \operatorname{tr}(D_\beta A)I).\end{aligned}$$

If $\mathscr{R}$ is a connected ring, with the only idempotents being 1 and 0, then this reduces to the previous two cases.

Proof. For $n = 1$, by Lemma 3.4.3 we have $G_0 = \{id\}$ and $G_2 \subseteq G_1$. Since the ring is commutative, any bracket is zero, so an automorphism is just a bijective linear map. This is determined by where I is sent, and so is isomorphic to the group of units of the ring. Hence, $\operatorname{Aut}_L(T_n) \cong \operatorname{Aut}_L(\mathscr{R}) \cong G_1$.

From this point forward, we assume $n > 1$. Let $G = G_0 \times G_1$ if $n = 2$ and $G = (G_0 \times G_1) \rtimes G_2$ if $n > 2$. In either case, $G \leq \operatorname{Aut}_L(T_n)$ so it suffices to show that every $\varphi \in \operatorname{Aut}_L(T_n)$ belongs to G.

In order to prove this, we will start with an arbitrary $\varphi \in \operatorname{Aut}_L(T_n)$ and compose it with suitable elements of G until only an element of G remains. Thus, over several steps we shall repeatedly replace φ with $\psi \circ \varphi$ for different suitably chosen $\psi \in G$.

The first goal is to get to the stage where $\varphi(\mathtt{E}_{1k}) = \mathtt{E}_{1k}$ for $2 \leq k \leq n$. We saw that $\varphi(\mathfrak{U}_{n-1}) = \mathfrak{U}_{n-1}$, so $\varphi(\mathtt{E}_{1n}) = \alpha\mathtt{E}_{1n}$ for some $\alpha \in \mathscr{R}^\times$.

Preserving $\mathtt{E}_{1n}$

Let $S = I + (\alpha - 1)\mathtt{E}_{nn}$. Then,

$$\begin{aligned}\varphi_S(\mathtt{E}_{1n}) &= (I + (\alpha-1)\mathtt{E}_{nn})\mathtt{E}_{1n}(I + (\alpha^{-1}-1)\mathtt{E}_{nn})\\ &= \mathtt{E}_{1n}(I + (\alpha^{-1}-1)\mathtt{E}_{nn})\\ &= \mathtt{E}_{1n} + (\alpha^{-1}-1)\mathtt{E}_{1n}\\ &= \alpha^{-1}\mathtt{E}_{1n}\,.\end{aligned}$$

Hence, by replacing φ with $\varphi_S \circ \varphi$, we may assume that $\varphi(\mathtt{E}_{1n}) = \mathtt{E}_{1n}$.

If $n = 2$, the first goal is achieved. Until further notice, we will assume that $n > 2$.

Preserving $\mathtt{E}_{1,n-1}$

For the next element, since $\mathtt{E}_{1,n-1} \in \mathcal{A}_{n-2} \subseteq \mathcal{U}_{n-2}$ and $\varphi(\mathcal{A}_{n-2}) \subseteq \varphi(\mathcal{U}_{n-2}) = \mathcal{U}_{n-2}$, we have

$$\varphi(\mathtt{E}_{1,n-1}) = \alpha\mathtt{E}_{1,n-1} + \beta\mathtt{E}_{1n} + \gamma\mathtt{E}_{2n} = \begin{bmatrix} \cdots & 0 & \alpha & \beta \\ & & 0 & \gamma \\ & & & 0 \\ & & & \vdots \end{bmatrix},$$

for some $\alpha, \beta, \gamma \in \mathscr{R}$.

As $\varphi(\mathcal{A}_{n-2})$ is an ideal of T_n, we have

$$\alpha\mathtt{E}_{1,n-1} + \beta\mathtt{E}_{1n} = [\mathtt{E}_{11}, \varphi(\mathtt{E}_{1,n-1})] \in \varphi(\mathcal{A}_{n-2}).$$

Hence, by re-arranging

$$\gamma\mathtt{E}_{2n} = \varphi(\mathtt{E}_{1,n-1}) - [\mathtt{E}_{11}, \varphi(\mathtt{E}_{1,n-1})] \in \varphi(\mathcal{A}_{n-2}),$$

and since $\varphi(\mathtt{E}_{1n}) = \mathtt{E}_{1n}$, we also have

$$\alpha\mathtt{E}_{1,n-1} = [\mathtt{E}_{11}, \varphi(\mathtt{E}_{1,n-1})] - \varphi(\beta\mathtt{E}_{1n}) \in \varphi(\mathcal{A}_{n-2}).$$

This implies that

$$\varphi(\mathcal{A}_{n-2}) = \mathscr{R}\mathtt{E}_{1n} \oplus \mathscr{R}\alpha\mathtt{E}_{1,n-1} \oplus \mathscr{R}\gamma\mathtt{E}_{2n}\,,$$

as an $\mathscr{R}$-module. However, since $\mathcal{A}_{n-2}$ is only rank 2, this means that

$$\mathscr{R}\alpha\mathtt{E}_{1,n-1} \oplus \mathscr{R}\gamma\mathtt{E}_{2n} \cong \mathscr{R}.$$

To simplify notation in this section, let $E := \mathtt{E}_{1,n-1}$ and $F := \mathtt{E}_{2n}$. For a general commutative ring, we examine when

$$\mathscr{R} \cong \mathscr{R}\alpha E \oplus \mathscr{R}\gamma F.$$

Let $\theta : \mathscr{R} \to \mathscr{R}\alpha E \oplus \mathscr{R}\gamma F$ denote this module isomorphism. Then, for some $u, v \in \mathscr{R}$, we have

$$\theta(1) = u\alpha E + v\gamma F.$$

From the bijection, there exists $\varepsilon, \eta \in \mathscr{R}$ such that

$$\begin{aligned} \theta(\varepsilon) &= u\alpha E, \\ \theta(\eta) &= v\gamma F. \end{aligned}$$

By the linearity and bijectivity of the isomorphism, we have $1 = \varepsilon + \eta$. We also note that

$$\varepsilon v\gamma F = \varepsilon\theta(\eta) = \theta(\varepsilon\eta) = \eta\theta(\varepsilon) = \eta y\alpha E.$$

Thus, $\theta(\varepsilon\eta) \in \mathscr{R}\alpha E \cap \mathscr{R}\gamma F = \{0\}$ and, by bijectivity, this implies that $\varepsilon\eta = 0$. Using this fact gives

$$\theta(\varepsilon) = \varepsilon\theta(1) = \varepsilon\theta(\varepsilon) + \varepsilon\theta(\eta) = \theta(\varepsilon^2) + \theta(\varepsilon\eta) = \theta(\varepsilon^2),$$

and so $\varepsilon = \varepsilon^2$ is an idempotent. The same argument applies to $\eta = \eta^2$ and, by construction, we see that $\eta = 1 - \varepsilon$.

From the bijection, there exists some $\xi \in \mathscr{R}$ such that $\varphi(\xi) = \alpha E$. This then means that $\varphi(u\xi) = u\alpha E = \varphi(\varepsilon)$. Thus, $u\xi = \varepsilon$. However,

$$\alpha E = \varphi(\xi) = \varphi(\xi(\varepsilon + \eta)) = \xi\varphi(\varepsilon) + \xi\varphi(\eta) = \xi u\alpha E + \xi v\gamma F$$

Hence, $\alpha = \xi u\alpha = \varepsilon\alpha$, which then implies that $(1 - \varepsilon)\alpha = 0$. By the same argument applied to γF, we have $\varepsilon\gamma = (1 - \eta)\gamma = 0$.

We now return to the regular notation for $E = \mathtt{E}_{1,n-1}$ and $F = \mathtt{E}_{2n}$. Having found this idempotent $\varepsilon \in \mathscr{R}$, we now replace φ with $\omega_\varepsilon \circ \varphi$ to get

$$\begin{aligned}\varphi(\mathtt{E}_{1n}) = \omega_\varepsilon(\mathtt{E}_{1n}) &= \varepsilon \mathtt{E}_{1n} - (1-\varepsilon)\mathtt{E}_{1n} \\ &= (2\varepsilon - 1)\mathtt{E}_{1n}\,,\end{aligned}$$

and, with $\varepsilon\gamma = 0$, $(1-\varepsilon)\alpha = 0$ and $\varepsilon\alpha = \alpha$, we have

$$\begin{aligned}\varphi(\mathtt{E}_{1,n-1}) &= \omega_\varepsilon(\alpha \mathtt{E}_{1,n-1} + \beta \mathtt{E}_{1n} + \gamma \mathtt{E}_{2n}) \\ &= \varepsilon(\alpha \mathtt{E}_{1,n-1} + \beta \mathtt{E}_{1n} + \gamma \mathtt{E}_{2n}) + (1-\varepsilon)\omega(\alpha \mathtt{E}_{1,n-1} + \beta \mathtt{E}_{1n} + \gamma \mathtt{E}_{2n}) \\ &= \varepsilon\alpha \mathtt{E}_{1,n-1} + \varepsilon\beta \mathtt{E}_{1n} - (1-\varepsilon)\beta \mathtt{E}_{1n} - (1-\varepsilon)\gamma \mathtt{E}_{1,n-1} \\ &= (\alpha - \gamma)\mathtt{E}_{1,n-1} + (2\varepsilon - 1)\beta \mathtt{E}_{1n}\,.\end{aligned}$$

Let $S = I - 2(1-\varepsilon)\mathtt{E}_{nn}$ so that $S^2 = I$. Replacing φ with $\varphi_S \circ \varphi$ then gives

$$\begin{aligned}\varphi(\mathtt{E}_{1n}) &= S(2\varepsilon - 1)\mathtt{E}_{1n} S \\ &= (I - 2(1-\varepsilon)\mathtt{E}_{nn})(2\varepsilon - 1)\mathtt{E}_{1n}(I - 2(1-\varepsilon)\mathtt{E}_{nn}) \\ &= (2\varepsilon - 1)\mathtt{E}_{1n}(I - 2(1-\varepsilon)\mathtt{E}_{nn}) \\ &= (2\varepsilon - 1)\mathtt{E}_{1n} + 2(1-\varepsilon)\mathtt{E}_{1n} \\ &= \mathtt{E}_{1n}\,,\end{aligned}$$

$$\begin{aligned}\varphi(\mathtt{E}_{1,n-1}) &= \varphi_S((\alpha - \gamma)\mathtt{E}_{1,n-1} + (2\varepsilon - 1)\beta \mathtt{E}_{1n}) \\ &= (I - 2(1-\varepsilon)\mathtt{E}_{nn})((\alpha - \gamma)\mathtt{E}_{1,n-1} + (2\varepsilon - 1)\beta \mathtt{E}_{1n})(I - 2(1-\varepsilon)\mathtt{E}_{nn}) \\ &= ((\alpha - \gamma)\mathtt{E}_{1,n-1} + (2\varepsilon - 1)\beta \mathtt{E}_{1n})(I - 2(1-\varepsilon)\mathtt{E}_{nn}) \\ &= (\alpha - \gamma)\mathtt{E}_{1,n-1} + (2\varepsilon - 1)\beta \mathtt{E}_{1n} + 2(1-\varepsilon)\beta \mathtt{E}_{1n} \\ &= (\alpha - \gamma)\mathtt{E}_{1,n-1} + \beta \mathtt{E}_{1n}\,.\end{aligned}$$

Relabelling $\hat{\alpha} := (\alpha - \gamma)$, we may then write $\varphi(\mathtt{E}_{1,n-1}) = \hat{\alpha}\mathtt{E}_{1,n-1} + \beta \mathtt{E}_{1n}$. Lastly, since $\varphi(\mathfrak{U}_{n-2}) = \mathfrak{U}_{n-2}$, for some $\alpha', \beta', \gamma' \in \mathscr{R}$, we have

$$\varphi(\mathtt{E}_{2n}) = \alpha' \mathtt{E}_{1,n-1} + \beta' \mathtt{E}_{1n} + \gamma' \mathtt{E}_{2n}\,.$$

Since $\{\mathtt{E}_{1n}, \mathtt{E}_{1,n-1}, \mathtt{E}_{2n}\}$ is a basis for the free $\mathscr{R}$-module $\mathfrak{U}_{n-2}$ and $\varphi(\mathfrak{U}_{n-2}) = \mathfrak{U}_{n-2}$, it follows that $\{\varphi(\mathtt{E}_{1n}), \varphi(\mathtt{E}_{1,n-1}), \varphi(\mathtt{E}_{2n})\}$ is also a basis for $\mathfrak{U}_{n-2}$.

Using this independence,

$$\det \begin{bmatrix} 1 & 0 & 0 \\ \beta & \hat{\alpha} & 0 \\ \beta' & \alpha' & \gamma' \end{bmatrix} = \hat{\alpha}\gamma' \in \mathscr{R}^{\times}.$$

Hence, $\hat{\alpha} \in \mathscr{R}^{\times}$. This allows us to define $S = I + (\hat{\alpha}^{-1} - 1)(\mathtt{E}_{11} + \mathtt{E}_{nn})$ and $S^{-1} = I + (\hat{\alpha} - 1)(\mathtt{E}_{11} + \mathtt{E}_{nn})$.

Replacing φ by $\varphi_S \circ \varphi$, we then have

$$\begin{aligned} \varphi(\mathtt{E}_{1n}) &= \varphi_S(\mathtt{E}_{1n}) = S\mathtt{E}_{1n}S^{-1} \\ &= (I + (\hat{\alpha}^{-1} - 1)(\mathtt{E}_{11} + \mathtt{E}_{nn}))\mathtt{E}_{1n}S^{-1} \\ &= \hat{\alpha}^{-1}\mathtt{E}_{1n}(I + (\hat{\alpha} - 1)(\mathtt{E}_{11} + \mathtt{E}_{nn})) \\ &= \hat{\alpha}^{-1}\hat{\alpha}\mathtt{E}_{1n} \\ &= \mathtt{E}_{1n}\,, \end{aligned}$$

$$\begin{aligned} \varphi(\mathtt{E}_{1,n-1}) &= \varphi_S(\hat{\alpha}\mathtt{E}_{1,n-1} + \beta\mathtt{E}_{1n}) = S(\hat{\alpha}\mathtt{E}_{1,n-1} + \beta\mathtt{E}_{1n})S^{-1} \\ &= (I + (\hat{\alpha}^{-1} - 1)(\mathtt{E}_{11} + \mathtt{E}_{nn}))(\hat{\alpha}\mathtt{E}_{1,n-1} + \beta\mathtt{E}_{1n})S^{-1} \\ &= (\hat{\alpha}^{-1}\hat{\alpha}\mathtt{E}_{1,n-1} + \hat{\alpha}^{-1}\beta\mathtt{E}_{1n})(I + (\hat{\alpha} - 1)(\mathtt{E}_{11} + \mathtt{E}_{nn})) \\ &= \mathtt{E}_{1,n-1} + \hat{\alpha}^{-1}\hat{\alpha}\beta\mathtt{E}_{1n} \\ &= \mathtt{E}_{1,n-1} + \beta\mathtt{E}_{1n}\,. \end{aligned}$$

Setting up another replacement, let $B = I + \beta\mathtt{E}_{n-1,n}$ so that $B^{-1} = I - \beta\mathtt{E}_{n-1,n}$. Replacing φ by $\varphi_B \circ \varphi$, we then have

$$\varphi(\mathtt{E}_{1n}) = \varphi_B(\mathtt{E}_{1n}) = B\mathtt{E}_{1n}B^{-1} = \mathtt{E}_{1n}\,,$$

$$\begin{aligned} \varphi(\mathtt{E}_{1,n-1}) &= \varphi_B(\mathtt{E}_{1,n-1} + \beta\mathtt{E}_{1n}) = B(\mathtt{E}_{1,n-1} + \beta\mathtt{E}_{1n})B^{-1} \\ &= (I + \beta\mathtt{E}_{n-1,n})(\mathtt{E}_{1,n-1} + \beta\mathtt{E}_{1n})B^{-1} \\ &= (\mathtt{E}_{1,n-1} + \beta\mathtt{E}_{1n})(I - \beta\mathtt{E}_{n-1,n}) \\ &= \mathtt{E}_{1,n-1} + \beta\mathtt{E}_{1n} - \beta\mathtt{E}_{1n} \\ &= \mathtt{E}_{1,n-1}\,. \end{aligned}$$

Thus, we have gotten to the step where $\varphi(\mathtt{E}_{1n}) = \mathtt{E}_{1n}$ and $\varphi(\mathtt{E}_{1,n-1}) = \mathtt{E}_{1,n-1}$. If $n = 3$, this achieves the first goal. In the case where $n > 3$, we continue with an induction argument.

Preserving $\mathtt{E}_{1k}$ for $k > 1$

Assume that $\varphi(\mathtt{E}_{1k}) = \mathtt{E}_{1k}$ for $m < k \le n$ where $2 \le m \le n-2$. This implies that $\varphi(A) = A$ for all $A \in \mathcal{A}_m$.

Since $\mathcal{A}_m = \mathcal{A}_{m-1} \cap \mathfrak{U}_m$ and φ acts trivially on $\mathcal{A}_m$, we have

$$\mathcal{A}_m = \varphi(\mathcal{A}_m) = \varphi(\mathcal{A}_{m-1} \cap \mathfrak{U}_m) = \varphi(\mathcal{A}_{m-1}) \cap \varphi(\mathfrak{U}_m) = \varphi(\mathcal{A}_{m-1}) \cap \mathfrak{U}_m \,.$$

Note that the following argument requires some unavoidable index gymnastics. Since $\mathtt{E}_{1m} \in \mathfrak{U}_{m-1}$, by Proposition 3.5.3 for some $\alpha_{ij} \in \mathscr{R}$ we have

$$\varphi(\mathtt{E}_{1m}) = \sum_{j-i \ge m-1} \alpha_{ij} \mathtt{E}_{ij} \,.$$

For $k = 2, 3, \ldots, n-m$, we have $\mathtt{E}_{k,k+1} \in \mathfrak{U}$ and

$$\begin{aligned} [\mathtt{E}_{k,k+1}, \varphi(\mathtt{E}_{1m})] &= \sum_{j-i \ge m-1} \alpha_{ij} \mathtt{E}_{k,k+1} \mathtt{E}_{ij} - \sum_{j-i \ge m-1} \alpha_{ij} \mathtt{E}_{ij} \mathtt{E}_{k,k+1} \\ &= \sum_{j \ge m+k} \alpha_{k+1,j} \mathtt{E}_{kj} - \sum_{i \le k-m+1} \alpha_{ik} \mathtt{E}_{i,k+1} \,. \end{aligned}$$

We know $\mathtt{E}_{1m} \in \mathcal{A}_{m-1}$ and so $\varphi(\mathtt{E}_{1m}) \in \varphi(\mathcal{A}_{m-1})$. Since $\varphi(\mathcal{A}_{m-1})$ is an ideal, then $[\mathtt{E}_{k,k+1}, \varphi(\mathtt{E}_{1m})]$ is also in $\varphi(\mathcal{A}_{m-1})$. In addition, $\mathcal{A}_{m-1} \subseteq \mathfrak{U}_{m-1}$ and $\varphi(\mathfrak{U}_{m-1}) = \mathfrak{U}_{m-1}$. This means that $[\mathtt{E}_{k,k+1}, \varphi(\mathtt{E}_{1m})]$ is in $[\mathfrak{U}, \mathfrak{U}_{m-1}] = \mathfrak{U}_m$.

Hence, $[\mathtt{E}_{k,k+1}, \varphi(\mathtt{E}_{1m})] \in \varphi(\mathcal{A}_{m-1}) \cap \mathfrak{U}_m = \mathcal{A}_m$. In the sum, any coefficients of $\mathtt{E}_{ij}$ with $i > 1$ must be zero. Elements in the two sums can only cancel if $i = k$ and $j = k+1$. However, this implies that $m = 1$, contradicting our assumptions.

Examining the left sum in the bracket expansion above, since $k > 1$, we must have $\alpha_{k+1,j} = 0$ for $j \ge m+k$, because the coefficient of $\mathtt{E}_{kj}$ must be zero.

With this restriction on the coefficients, we may rewrite

$$\varphi(\mathtt{E}_{1m}) = \sum_{j=m}^{n} \alpha_{1j}\mathtt{E}_{1j} + \sum_{j=m+1}^{n} \alpha_{2j}\mathtt{E}_{2j} \, .$$

For k from $m+1$ to $n-1$ we may again take

$$[\varphi(\mathtt{E}_{1m}), \mathtt{E}_{k,k+1}] = \alpha_{1k}\mathtt{E}_{1,k+1} + \alpha_{2k}\mathtt{E}_{2,k+1} \, ,$$

which, by the same argument as above, is in $\varphi(\mathcal{A}_{m-1}) \cap \mathfrak{U}_m = \mathcal{A}_m$. Hence, $\alpha_{2k} = 0$ and so

$$\varphi(\mathtt{E}_{1m}) = \sum_{j=m}^{n} \alpha_{1j}\mathtt{E}_{1j} \quad + \quad \alpha_{2n}\mathtt{E}_{2n} \, .$$

Since $n-2 \geq m$ by assumption, $\mathtt{E}_{2n} \in \mathfrak{U}_m$. Hence,

$$[\varphi(\mathtt{E}_{1m}), \mathtt{E}_{nn}] = \alpha_{1n}\mathtt{E}_{1n} + \alpha_{2n}\mathtt{E}_{2n} \quad \in \varphi(\mathcal{A}_{m-1}) \cap \mathfrak{U}_m = \mathcal{A}_m \, ,$$

so $\alpha_{2n} = 0$ and

$$\varphi(\mathtt{E}_{1m}) = \sum_{j=m}^{n} \alpha_{1j}\mathtt{E}_{1j} \quad \in \mathcal{A}_{m-1} \, .$$

Since φ acts as the identity on $\mathcal{A}_m$, we have $\alpha_{1m} \in \mathscr{R}^{\times}$. If we then define

$$S = I + (\alpha_{1m} - 1)\mathtt{E}_{mm}$$

we have

$$\begin{aligned}
\varphi_S(\mathtt{E}_{1m}) &= (I + (\alpha_{1m} - 1)\mathtt{E}_{mm})\mathtt{E}_{1m}(I + (\alpha_{1m}^{-1} - 1)\mathtt{E}_{mm}) \\
&= \mathtt{E}_{1m}(I + (\alpha_{1m}^{-1} - 1)\mathtt{E}_{mm}) \\
&= \mathtt{E}_{1m} + (\alpha_{1m}^{-1} - 1)\mathtt{E}_{1m} \\
&= \alpha_{1m}^{-1}\mathtt{E}_{1m} \, ,
\end{aligned}$$

and for $k > m$,

$$\begin{aligned}\varphi_S(\mathtt{E}_{1k}) &= (I + (\alpha_{1m} - 1)\mathtt{E}_{mm})\mathtt{E}_{1k}(I + (\alpha_{1m}^{-1} - 1)\mathtt{E}_{mm}) \\ &= \mathtt{E}_{1k}(I + (\alpha_{1m}^{-1} - 1)\mathtt{E}_{mm}) \\ &= \mathtt{E}_{1k}\,.\end{aligned}$$

Hence, if we replace φ with $\varphi_S \circ \varphi$ we may assume that $\alpha_{1m} = 1$.

Taking $B = I + \sum_{k=m+1}^{n} \alpha_{1k}\mathtt{E}_{mk}$ with $B^{-1} = I - \sum_{k=m+1}^{n} \alpha_{1k}\mathtt{E}_{mk}$ then, for $k > m$,

$$\begin{aligned}\varphi_B(\mathtt{E}_{1k}) &= (I + \sum_{k=m+1}^{n} \alpha_{1k}\mathtt{E}_{mk})\mathtt{E}_{1k}(I - \sum_{k=m+1}^{n} \alpha_{1k}\mathtt{E}_{mk}) \\ &= \mathtt{E}_{1k}(I - \sum_{k=m+1}^{n} \alpha_{1k}\mathtt{E}_{mk}) \\ &= \mathtt{E}_{1k}\,,\end{aligned}$$

$$\begin{aligned}\varphi_B(\mathtt{E}_{1m}) &= (I + \sum_{k=m+1}^{n} \alpha_{1k}\mathtt{E}_{mk})\mathtt{E}_{1m}(I - \sum_{k=m+1}^{n} \alpha_{1k}\mathtt{E}_{mk}) \\ &= \mathtt{E}_{1m}(I - \sum_{k=m+1}^{n} \alpha_{1k}\mathtt{E}_{mk}) \\ &= \mathtt{E}_{1m} - \sum_{k=m+1}^{n} \alpha_{1k}\mathtt{E}_{1k}\,.\end{aligned}$$

Hence, by replacing φ with $\varphi_B \circ \varphi$, we have $\varphi(\mathtt{E}_{1k}) = \mathtt{E}_{1k}$ for all $k \geq m$.

This completes the induction step and the first goal is achieved, so that $\varphi(\mathtt{E}_{1k}) = \mathtt{E}_{1k}$ for $2 \leq k \leq n$.

Effect on $\mathtt{E}_{11}$

Next we examine the effect of φ on $\mathtt{E}_{11}$. For $\alpha_{ij} \in \mathscr{R}$, we write

$$\varphi(\mathtt{E}_{11}) = \sum_{i \leq j} \alpha_{ij}\mathtt{E}_{ij}\,.$$

The Lie product of any two elements of T_n is in $\mathcal{U}$. So,

$$[\varphi(\mathtt{E}_{11}), \mathtt{E}_{nn}] \;=\; \sum_{i=1}^{n} \alpha_{in}\mathtt{E}_{in} - \alpha_{nn}\mathtt{E}_{nn} \;=\; \sum_{i=1}^{n-1} \alpha_{in}\mathtt{E}_{in}\,,$$

belongs to $\varphi(\mathcal{A}) \cap \mathcal{U} = \varphi(\mathcal{A} \cap \mathcal{U}) = \varphi(\mathcal{A}_1) = \mathcal{A}_1$, and we conclude that $\alpha_{in} = 0$ for $2 \leq i \leq n-1$. Next, for $n > 2$, we find that

$$\begin{aligned}[\varphi(\mathtt{E}_{11}), \mathtt{E}_{n-1,n-1}] &= \sum_{i=1}^{n-1} \alpha_{i,n-1}\mathtt{E}_{i,n-1} - \sum_{j=n-1}^{n} \alpha_{n-1,j}\mathtt{E}_{n-1,j} \\ &= \sum_{i=1}^{n-2} \alpha_{i,n-1}\mathtt{E}_{i,n-1} \quad \in \mathcal{A}_1\,,\end{aligned}$$

so by the same argument, $\alpha_{i,n-1} = 0$ for $2 \leq i \leq n-2$.

Repeating this for $[\varphi(\mathtt{E}_{11}), \mathtt{E}_{kk}]$ for $k = n, n-1, \ldots, 3$, we may conclude that $\alpha_{ij} = 0$ for $1 < i < j \leq n$. Hence, the only nonzero entries of $\varphi(\mathtt{E}_{11})$ are in the first row or on the diagonal.

This is vacuously true if $n = 2$. From here onward, we assume $n \geq 2$.

As $[\mathtt{E}_{11}, \mathtt{E}_{1k}] = \mathtt{E}_{1k}$ for $k > 1$, and the automorphism preserves brackets, we have

$$\begin{aligned}\mathtt{E}_{1k} &= \varphi(\mathtt{E}_{1k}) \\ &= \varphi[\mathtt{E}_{11}, \mathtt{E}_{1k}] \\ &= [\varphi(\mathtt{E}_{11}), \varphi(\mathtt{E}_{1k})] \\ &= [\varphi(\mathtt{E}_{11}), \mathtt{E}_{1k}].\end{aligned}$$

On the other hand, since $\alpha_{kj} = 0$ for $1 < k < j \leq n$,

$$\begin{aligned}[\varphi(\mathtt{E}_{11}), \mathtt{E}_{1k}] &= \sum_{i \leq 1} \alpha_{i1}\mathtt{E}_{ik} - \sum_{k \leq j} \alpha_{kj}\mathtt{E}_{1j} \\ &= \alpha_{11}\mathtt{E}_{1k} - \sum_{j=k}^{n} \alpha_{kj}\mathtt{E}_{1j} \\ &= \alpha_{11}\mathtt{E}_{1k} - \alpha_{kk}\mathtt{E}_{1k} \\ &= (\alpha_{11} - \alpha_{kk})\mathtt{E}_{1k}\,.\end{aligned}$$

Thus, $\alpha_{kk} = \alpha_{11} - 1$ for $k > 1$. It then follows that

$$\varphi(\mathtt{E}_{11}) = (\alpha_{11} - 1)I + \mathtt{E}_{11} + \sum_{j=2}^{n} \alpha_{1j}\mathtt{E}_{1j} .$$

If we define $S = I + \sum_{j=2}^{n} \alpha_{1j}\mathtt{E}_{1j}$, we have $S^{-1} = I - \sum_{j=2}^{n} \alpha_{1j}\mathtt{E}_{1j}$ and for $k > 1$,

$$\begin{aligned}\varphi_S(\mathtt{E}_{1k}) &= (I + \sum_{j=2}^{n} \alpha_{1j}\mathtt{E}_{1j})\mathtt{E}_{1k}(I - \sum_{j=2}^{n} \alpha_{1j}\mathtt{E}_{1j}) \\ &= \mathtt{E}_{1k}(I - \sum_{j=2}^{n} \alpha_{1j}\mathtt{E}_{1j}) \ = \ \mathtt{E}_{1k} ,\end{aligned}$$

and

$$\begin{aligned}\varphi_S(\mathtt{E}_{11}) &= (I + \sum_{j=2}^{n} \alpha_{1j}\mathtt{E}_{1j})\mathtt{E}_{11}(I - \sum_{j=2}^{n} \alpha_{1j}\mathtt{E}_{1j}) \\ &= \mathtt{E}_{11}(I - \sum_{j=2}^{n} \alpha_{1j}\mathtt{E}_{1j}) \\ &= \mathtt{E}_{11} - \sum_{j=2}^{n} \alpha_{1j}\mathtt{E}_{1j} .\end{aligned}$$

By replacing φ with $\varphi_S \circ \varphi$ we may assume that $\varphi(\mathtt{E}_{1k}) = \mathtt{E}_{1k}$ for $1 < k \leq n$, and $\varphi(\mathtt{E}_{11}) = (\alpha_{11} - 1)I + \mathtt{E}_{11}$.

To simplify notation, let $\alpha_1 := \alpha_{11} - 1$ so that we may write $\varphi(\mathtt{E}_{11}) = \alpha_1 I + \mathtt{E}_{11}$.

Effect off the diagonal and first row

We now examine the effect of φ on elements $\mathtt{E}_{rs}$ which are off the diagonal and first row. Let $1 < r < s \leq n$. Then $\mathtt{E}_{rs} \in \mathfrak{U}_{s-r}$ and so $\varphi(\mathtt{E}_{rs}) \in \mathfrak{U}_{s-r}$. So, for some $\alpha_{ij} \in \mathscr{R}$,

$$\varphi(\mathtt{E}_{rs}) = \sum_{j-i \geq s-r} \alpha_{ij}\mathtt{E}_{ij} .$$

For $k \neq r$, we have $[\mathtt{E}_{1k}, \mathtt{E}_{rs}] = 0$ so applying the bracket also gives $[\mathtt{E}_{1k}, \varphi(\mathtt{E}_{rs})] = 0$. For this fixed $k \neq r$,

$$\begin{aligned} 0 &= \sum_{j-i \geq s-r} \alpha_{ij} \mathtt{E}_{1k} \mathtt{E}_{ij} - \sum_{j-i \geq s-r} \alpha_{ij} \mathtt{E}_{ij} \mathtt{E}_{1k} \\ &= \sum_{j-k \geq s-r} \alpha_{kj} \mathtt{E}_{1j} - \sum_{1-i \geq s-r} \alpha_{i1} \mathtt{E}_{ik} \\ &= \sum_{j-k \geq s-r} \alpha_{kj} \mathtt{E}_{1j} \,. \end{aligned}$$

Consequently, $\alpha_{ij} = 0$ for all $i \neq r$, and so

$$\varphi(\mathtt{E}_{rs}) = \sum_{j \geq s} \alpha_{rj} \mathtt{E}_{rj} \,.$$

Since $[\mathtt{E}_{1r}, \mathtt{E}_{rs}] = \mathtt{E}_{1s}$, this also means $[\mathtt{E}_{1r}, \varphi(\mathtt{E}_{rs})] = \mathtt{E}_{1s}$. Hence,

$$\begin{aligned} \mathtt{E}_{1s} &= \sum_{j \geq s} \alpha_{rj} \mathtt{E}_{1r} \mathtt{E}_{rj} - \sum_{j \geq s} \alpha_{rj} \mathtt{E}_{rj} \mathtt{E}_{1r} \\ &= \sum_{j \geq s} \alpha_{rj} \mathtt{E}_{1j} - \sum_{1 \geq s} \alpha_{r1} \mathtt{E}_{rr} \\ &= \sum_{j \geq s} \alpha_{rj} \mathtt{E}_{1j} \,, \end{aligned}$$

and so $\alpha_{rs} = 1$ and $\alpha_{rj} = 0$ for $j > s$. Thus, $\varphi(\mathtt{E}_{rs}) = \mathtt{E}_{rs}$, making φ the identity on $\mathfrak{U}$.

Effect on the rest of the diagonal

To summarize, we have shown that $\varphi(\mathtt{E}_{ij}) = \mathtt{E}_{ij}$ for $j > i$ and $\varphi(\mathtt{E}_{11}) = \alpha_1 I + \mathtt{E}_{11}$. We examine the other diagonal basis elements.

Now, let $2 \leq k \leq n$. Then $[\mathtt{E}_{1s}, \mathtt{E}_{kk}] = 0$ for $s \neq k$ and, by applying φ, we also have $[\mathtt{E}_{1s}, \varphi(\mathtt{E}_{kk})] = 0$ for $s \neq k$.

By Lemma 2.3.3, if $\varphi(\mathtt{E}_{kk}) = \sum_{i\le j} \alpha_{ij}\mathtt{E}_{ij}$ commutes with $\mathtt{E}_{1s}$ then $\alpha_{sj} = 0$ for $j \neq s$ and $\alpha_{11} = \alpha_{ss}$. We may also see this explicitly,

$$\begin{aligned} 0 &= \sum_{i\le j} \alpha_{ij}\mathtt{E}_{1s}\mathtt{E}_{ij} - \sum_{i\le j} \alpha_{ij}\mathtt{E}_{ij}\mathtt{E}_{1s} \\ &= \sum_{s\le j} \alpha_{sj}\mathtt{E}_{1j} - \sum_{i\le 1} \alpha_{i1}\mathtt{E}_{is} \\ &= \sum_{j=s}^{n} \alpha_{sj}\mathtt{E}_{1j} - \alpha_{11}\mathtt{E}_{1s}\,. \end{aligned}$$

Since the above assumes that $s \neq k$,

$$\varphi(\mathtt{E}_{kk}) = \sum_{i\neq k} \alpha_{11}\mathtt{E}_{ii} + \sum_{j=k}^{n} \alpha_{kj}\mathtt{E}_{kj}\,.$$

Define $\alpha_k := \alpha_{11}$ and $\beta_j := \alpha_{kj}$ for $j > k$. Finally, let $\beta_k := \alpha_{kk} - \alpha_{11}$. Then we may write

$$\varphi(\mathtt{E}_{kk}) = \alpha_k I + \sum_{j=k}^{n} \beta_j \mathtt{E}_{kj}.$$

Since $[\mathtt{E}_{1k}, \mathtt{E}_{kk}] = \mathtt{E}_{1k}$, applying φ gives $[\mathtt{E}_{1k}, \varphi(\mathtt{E}_{kk})] = \mathtt{E}_{1k}$ and

$$\begin{aligned} \mathtt{E}_{1k} &= \sum_{j=k}^{n} \beta_j \mathtt{E}_{1k}\mathtt{E}_{kj} - \sum_{j=k}^{n} \beta_j \mathtt{E}_{kj}\mathtt{E}_{1k} \\ &= \sum_{j=k}^{n} \beta_j \mathtt{E}_{1j}\,. \end{aligned}$$

Hence, $\beta_k = 1$ and $\beta_j = 0$ for $j > k$. Thus, combining this with the previous result for $\mathtt{E}_{11}$, for all i we may write

$$\varphi(\mathtt{E}_{ii}) = \alpha_k I + \mathtt{E}_{ii}\,.$$

Along with the fact that $\varphi(\mathtt{E}_{ij}) = \mathtt{E}_{ij}$ for $j > i$, this allows us to write φ in the form

$$\varphi(\mathtt{E}_{ij}) = \mathtt{E}_{ij} + \delta_{ij}\alpha_i I\,, \qquad i \le j\,,$$

or $\varphi(A) = \psi_\alpha(A) = A + \mathrm{tr}(D_\alpha A)I$, for all $A \in T_n$. Since $\varphi(I) = (1 + \sum \alpha_i)I$ and φ is a linear automorphism, $\mathscr{R}$ must contain an inverse for $1 + \sum \alpha_i$. Hence, $1 + \sum \alpha_i \in \mathscr{R}^\times$, showing that $\varphi \in G_1$.

This completes the decomposition of an arbitrary automorphism from $\mathrm{Aut}_L(T_n)$, proving Theorem 3.6.1. $\square$

Chapter 4

Block-Triangular Algebras

The methods in this chapter were inspired by the paper by Marcoux and Sourour [7] on Lie isomorphisms of nest algebras. In particular, we adapt their general approach to finite dimensional block-triangular matrix algebras over a unique factorization domain (UFD), not of characteristic 2 or 3.

We follow the same broad steps in this paper, but many of the proofs are changed significantly, particularly where the authors relied on Hilbert space results. Only the most elementary properties of a UFD are invoked, in particular the cancellation property and the unique factorization, giving greatest common divisors.

As mentioned in the Background chapter we adopt some terminology from linear algebra (over a field), rather than use the equivalent terms specific to module theory. We feel this is justified so as to emphasize the places that the theory is substantially similar. The tools of linear algebra are heavily used throughout many branches of mathematics, so this abuse of terminology will hopefully be forgiven in the interests of making this work more accessible.

We will use notation established in Section 2.3. For an algebra $\mathcal{T} := T(n_1, \dots, n_k)(\mathscr{R}) \subseteq M_n(\mathscr{R})$, the letter k denotes the number of blocks on the diagonal of the matrices and each $m_t = \sum_{i=1}^{t} n_i$, for $1 \leq t \leq k$, denotes the dimension of the nest subspace $N_t \in \mathcal{N}$. There are some minor differences in notation from Chapter 3 to more closely match the paper by Marcoux and Sourour [7], in case the reader wishes to compare the approach used here.

4.1 Discussion of Main Theorem (Block-Triangular)

Since $\mathscr{R}$ is a UFD, it is connected and the only idempotents are 1 and 0. Our main result echoes the restricted form in the comments following Theorem 3.6.1. The condition that the UFD not have characteristic 2 or 3 follows from several steps required in the proofs. This condition is also present in the general theorem of Martindale for Lie isomorphisms of simple rings (Theorem 2.4.6).

Theorem 4.1.1 (Main)**.** *Let $\mathscr{R}$ be a Unique Factorization Domain, not of characteristic 2 or 3. Let $T(n_1,\ldots,n_k)(\mathscr{R})$ and $T(p_1,\ldots,p_\ell)(\mathscr{R})$ be upper block-triangular sub-algebras of $M_n(\mathscr{R})$ for $n \geq 2$. An $\mathscr{R}$-linear map $\varphi : T(n_1,\ldots,n_k)(\mathscr{R}) \to T(p_1,\ldots,p_\ell)(\mathscr{R})$ is a Lie isomorphism if and only if $\ell = k$ and either*

(a) $n_t = p_t$ for $1 \leq t \leq k$ so that $T(n_1,\ldots,n_k)(\mathscr{R}) = T(p_1,\ldots,p_\ell)(\mathscr{R})$, and there exists an invertible $Y \in T(n_1,\ldots,n_k)(\mathscr{R})$ and a generalized trace τ on $T(n_1,\ldots,n_k)(\mathscr{R})$ satisfying $1+\tau(I) \in \mathscr{R}^\times$, such that

$$\varphi(T) = Y^{-1}TY + \tau(T)I;$$

or

(b) $n_t = p_{k+1-t}$ for $1 \leq t \leq k$ so that $T(n_k,\ldots,n_1)(\mathscr{R}) = T(p_1,\ldots,p_\ell)(\mathscr{R})$, and there exists an invertible $Y \in T(p_1,\ldots,p_\ell)(\mathscr{R})$ and a generalized trace τ as above, such that

$$\varphi(T) = -Y^{-1}JT^\top JY + \tau(T)I,$$

where $J = [\delta_{i,n+1-i}]$ is the anti-diagonal permutation matrix.

If we look at the nest structure with the two algebras, $\mathcal{T}(\mathcal{N}) = T(n_1,\ldots,n_k)(\mathscr{R})$ and $\mathcal{T}(\mathcal{M}) = T(p_1,\ldots,p_\ell)(\mathscr{R})$ then, in the first case $\mathcal{N} = \mathcal{M}$ and in the second case $\mathcal{N} = \mathcal{M}^\angle$.

Note that $T(1,1)(\mathscr{R})$ is just the 2×2 triangular algebra $T_2(\mathscr{R})$ and Theorem 3.6.1 covers this case. Also, $T(2)(\mathscr{R})$ is the full matrix algebra $M_2(\mathscr{R})$, which is addressed below. Unless expressly stated, for the rest of the paper we assume that $n > 2$.

To prove the result we will first show the effect of the Lie isomorphism φ on idempotents in the algebra. From this we will be able to establish the order and dimension-preserving properties of φ. Finally, we use projections to decompose φ into compositions of inner automorphisms, negative anti-isomorphisms and what we previously termed trace isomorphisms.

4.2 Idempotents

Unless explicitly stated we assume that $\mathscr{R}$ is a UFD, not of characteristic 2 or 3. We note again that, although we will use the nest structure in many of these proofs, the main results apply to the matrix algebras themselves, without reference to being endomorphisms of the free module $\mathscr{R}^n$.

The skeleton of the following lemma is drawn from Lemma 3.1 in Marcoux and Sourour's paper [7]. If a property of an element $A \in \mathcal{T}(\mathcal{N})$ can be phrased as a bracket (Lie product) relation, then that property will be preserved under a Lie isomorphism.

In particular, we will be using nest projections and more general idempotent elements $A = A^2$ to decompose our Lie isomorphism. Hence, seeing how idempotents interact with the Lie bracket is of fundamental importance to the proofs that follow.

Lemma 4.2.1 (Idempotent-Bracket Lemma)**.** *Let $\mathcal{T}$ be an upper block-triangular algebra and $\mathcal{N}$ be the corresponding nest. Then,*

(a) A is the sum of a "scalar" and an idempotent if and only if

$$[A, [A, [A, T]]] = [A, T] \quad \textit{for every} \;\; T \in \mathcal{T}. \tag{1}$$

(b) A is the sum of a "scalar" and an idempotent whose range belongs to $\mathcal{N}$ if and only if

$$[A, [A, T]] = [A, T] \quad \textit{for every} \;\; T \in \mathcal{T}. \tag{2}$$

The proof requires several steps.

Proof. **Lemma 4.2.1, Part (a)** (⇒)

We first note

$$\begin{aligned}[A,[A,T]] &= [A,(AT - TA)] \\ &= A(AT - TA) - (AT - TA)A \\ &= A^2T - 2ATA + TA^2,\end{aligned}$$

and so

$$\begin{aligned}[A,[A,[A,T]]] &= [A, A^2T - 2ATA + TA^2] \\ &= A^3T - 3A^2TA + 3ATA^2 - TA^3.\end{aligned}$$

If $A = E + \alpha I$ is the sum of a "scalar" αI and an idempotent $E^2 = E$, then the scalar portion vanishes under the bracket and $[A,T] = [E,T]$, since αI is contained in the center of $\mathfrak{T}$. Then,

$$\begin{aligned}[A,[A,[A,T]]] &= E^3T - 3E^2TE + 3ETE^2 - TE^3 \\ &= ET - TE = [E,T] = [A,T].\end{aligned}$$

Proof of Lemma 4.2.1, Part (a) (⇐)

Conversely, let $[A,[A,[A,T]]] = [A,T]$ for every $T \in \mathfrak{T}$, so that

$$(A^3 - A)T - 3A^2TA + 3ATA^2 - T(A^3 - A) = 0. \tag{3}$$

Case 1

Suppose that the bottom right-most block of A is zero except for a single value along the diagonal, i.e., $A_{kk} = \lambda I_{kk}$ for some $\lambda \in \mathscr{R}$. We may translate A by a scalar matrix without affecting the relation in the theorem, so we may assume that $A_{kk} = [0]$.

For any fixed $1 \leq t \leq n$, take $T = \mathtt{E}_{tn} \in \mathfrak{T}$, a basis matrix with a single 1 in the final column. Writing $A = [a_{ij}]$, we know the bottom row $a_{nj} = 0$ for all $1 \leq j \leq n$ and so $TA = \mathtt{E}_{tn}A = 0$. Thus, equation (3) gives $(A^3 - A)T = 0$.

Let $\hat{A} = A^3 - A$. Then, for some $\hat{\alpha}_{ij} \in R$, we may write $\hat{A} = [\hat{\alpha}_{ij}]$, and

$$0 = \hat{A}T = \left(\sum_{i,j} \hat{\alpha}_{ij}\mathtt{E}_{ij}\right)\mathtt{E}_{tn} = \sum_{i} \hat{\alpha}_{it}\mathtt{E}_{in}.$$

By linear independence of the basis matrices, $\hat{\alpha}_{it} = 0$ for each fixed t and all i. Hence, $A^3 - A = \hat{A} = 0$ and so $A^3 = A$.

Next, we have two "almost" idempotents which we denote

$$\begin{aligned} E &= (A^2 + A), \text{ and} \\ F &= (A^2 - A), \end{aligned}$$

where $EF = FE = 0$, $E^2 = 2E$, and $F^2 = 2F$. This allows us to write

$$\begin{aligned} 2A &= E - F, \quad \text{and} \\ 2A^2 &= E + F. \end{aligned}$$

With $A^3 = A$, equation (3) gives us $A^2TA = ATA^2$ for all $T \in \mathcal{T}$. So, multiplying through by 4 gives

$$\begin{aligned} (2A^2)T(2A) = (2A)T(2A^2) \quad &\Rightarrow \quad (E+F)T(E-F) = (E-F)T(E+F) \\ &\Rightarrow \quad 2FTE = 2ETF \\ &\Rightarrow \quad 0 = (ETF)E = FTE^2 = 2FTE = 2ETF. \end{aligned}$$

Thus, $FTE = 0 = ETF$. Suppose that both E and F are nonzero. We seek a contradiction.

We first note that the diagonal blocks E_{ii} and F_{ii} cannot all be zero, otherwise E or F would be nilpotent, not "almost" idempotent with $E^2 = 2E$ and $F^2 = 2F$. Let

$$\begin{aligned} t &= \min\{i : E_{ii} \neq 0\}, \text{ and} \\ s &= \min\{i : F_{ii} \neq 0\}. \end{aligned}$$

Without loss of generality, we may assume $t \leq s$. (If not, we may use FTE below, instead of ETF.)

Suppose $t = s$. There then exists e_i such that $Ee_i = x \neq 0$ for some $m_{t-1} < i \leq m_t$. Likewise, there exists e_j such that $F^\top e_j = y \neq 0$ for some $m_{t-1} < j \leq m_t$. We see that the matrix $T = e_i e_j^\top = \mathtt{E}_{ij}$ has a single 1 entry in the tt block on the diagonal, hence $e_i e_j^\top \in \mathfrak{T}$.

Otherwise, if $t < s$, then there exists e_i such that $Ee_i = x \neq 0$ for some $m_{t-1} < i \leq m_t$. Likewise, there exists e_j such that $F^\top e_j = y \neq 0$ for some $m_{s-1} < j \leq m_s$. Since $t < s$ we must have $i < j$, and we again have $e_i e_j^\top \in \mathfrak{T}$.

As constructed, let $T = e_i e_j^\top = \mathtt{E}_{ij}$. Then,

$$ETF \;=\; E(e_i e_j^\top)F \;=\; (Ee_i)(F^\top e_j)^\top \;=\; xy^\top \;\neq\; 0,$$

which is a contradiction. Therefore, one of E or F must be zero. If $(A^2 - A) = F = 0$ then $A^2 = A$ is idempotent. If $(A^2 + A) = E = 0$, then $(-A)^2 = A^2 = -A$ and so $I - (-A) = I + A$ is idempotent. This concludes the case where the bottom corner is zero (or a scalar).

Case 2

Suppose $A_{kk} \neq \lambda I_{kk}$. We note that this means $n_k = \dim(N_{m_{k-1}}^\perp) \geq 2$.

With this condition, there exists an element $y \in N_{m_{k-1}}^\perp$ such that $A^\top y$ and y are linearly independent. By Lemma 2.1.1, we may then find some $z \in \mathscr{R}^n$ such that $(z|y) = z^\top y = 0$ but $(Az|y) = (z|A^\top y) = z^\top(A^\top y) \neq 0$.

For any $x \in \mathscr{R}^n$, if we let $T = xy^\top$, we note that $T \in \mathfrak{T}$ because $y \in N_{k-1}^\perp$. Taking equation (3) and applying it to the chosen $T = xy^\top$ and z, gives

$$\begin{aligned} 0 &= (A^3 - A)(xy^\top)z - 3A^2(xy^\top)Az + 3A(xy^\top)A^2 z - (xy^\top)(A^3 - A)z \\ &= (z|y)(A^3 - A)x - 3(Az|y)A^2 x + 3(A^2 z|y)Ax - ((A^3 - A)z|y)x, \\ &= -3(Az|y)A^2 x + 3(A^2 z|y)Ax - ((A^3 - A)z|y)x. \end{aligned}$$

So, for all $x \in \mathscr{R}^n$ we have

$$aA^2 = bA + cI$$

for some $a, b, c \in \mathscr{R}$ with $a \neq 0$. Putting this back into equation (3), gives $(b^2 - a^2 + 4ac)AT = (b^2 - a^2 + 4ac)TA$. Since the assumptions in Case 2 exclude the possibility that A is a scalar, we have $4ac = a^2 - b^2$.

We multiply through by $a \neq 0$ and substitute for aA^2 in two steps:

$$\begin{gathered} a(A^3 - A)T + 3aATA^2 = aT(A^3 - A) + 3aA^2TA \\ (A(bA + cI) - aA)T + 3AT(bA + cI) = T(A(bA + cI) - aA) + 3(bA + cI)TA \\ bA^2T + (4c - a)AT = bTA^2 + (4c - a)TA \\ b(aA^2)T + a(4c - a)AT = bT(aA^2) + a(4c - a)TA \\ b(bA + cI)T + a(4c - a)AT = bT(bA + cI) + a(4c - a)TA \\ (b^2 - a^2 + 4ac)AT = (b^2 - a^2 + 4ac)TA \end{gathered}$$

Let $\mathbb{K}$ be the fraction field of $\mathscr{R}$ and consider A as a matrix in $M_n(\mathbb{K})$.

Then $c = \frac{a^2 - b^2}{4a} \in \mathbb{K}$ and

$$A^2 = \tfrac{b}{a}A + \tfrac{a^2 - b^2}{4a^2}I.$$

We then observe that $A + \frac{a-b}{2a}I$ is idempotent in $M_n(\mathbb{K})$,

$$\begin{aligned} (A + \tfrac{a-b}{2a}I)^2 &= A^2 + \tfrac{a-b}{a}A + \tfrac{(a-b)^2}{4a^2}I \\ &= (\tfrac{b}{a}A + \tfrac{a^2-b^2}{4a^2}I) + (1 - \tfrac{b}{a})A + \tfrac{(a-b)^2}{4a^2}I \\ &= A + \tfrac{a-b}{2a}I. \end{aligned}$$

We show that this implies that $\frac{a-b}{2a}$ is in $\mathscr{R}$.

Proposition 4.2.2. *Let $\mathscr{R}$ be a unique factorization domain and $\mathbb{K}$ the fraction field of $\mathscr{R}$. Let $A \in M_n(\mathbb{K})$ and $\alpha, \beta \in \mathscr{R}$ such that $\frac{\beta}{\alpha} \in \mathbb{K}$. If $A \in M_n(\mathscr{R})$ and $A + \frac{\beta}{\alpha}I \in M_n(\mathbb{K})$ is a nontrivial idempotent, then $\frac{\beta}{\alpha} \in \mathscr{R}$ and $A + \frac{\beta}{\alpha}I \in M_n(\mathscr{R})$.*

Proof. If $\beta = 0$ we are done. We assume otherwise.

Given that $A + \frac{\beta}{\alpha} I$ is a nontrivial idempotent in $M_n(\mathbb{K})$, it has eigenvalues 1 and 0. Hence, A has eigenvalues $-\frac{\beta}{\alpha}$ and $-\frac{\beta}{\alpha} + 1$. In a UFD there is a well-defined notion of gcd, so we may assume that α and β are coprime in $\mathscr{R}$ (up to multiplication by a unit).

Since the entries of A come from $\mathscr{R}$, the determinant of A is also in $\mathscr{R}$. The determinant is also the product of powers of the eigenvalues, so

$$(-1)^{s+t} \left(\tfrac{\beta}{\alpha}\right)^s \left(\tfrac{\beta}{\alpha} - 1\right)^t = \det(A) \ \in \mathscr{R}$$

for some $s + t = n$ with $t \geq 1$. This implies that α^n divides $\beta^s(\beta - \alpha)^t$. Since α and β are coprime, $\beta - \alpha$ and α are also coprime. Hence, α must be a unit in $\mathscr{R}$ and so $\frac{\beta}{\alpha} = \beta\alpha^{-1} \in \mathscr{R}$, as desired. □

For the proof of the larger lemma, since $A + \frac{a-b}{2a} I$ is idempotent in $M_n(\mathbb{K})$ and $A \in M_n(\mathscr{R})$, by Proposition 4.2.2 we have have $\frac{a-b}{2a} \in \mathscr{R}$ and $A + \frac{a-b}{2a} I \in M_n(\mathscr{R})$, as desired. It is worth noting that in the proof of the Proposition we assumed α and β were coprime, so this does not imply that $2a$ is a unit in $\mathscr{R}$, only that it divides $a - b$.

This concludes Case 2 of Lemma 4.2.1, Part (a).

Proof of Lemma 4.2.1, Part (b) $(\Rightarrow)$

As above, if $A = E + \alpha I$ is the sum of a "scalar" αI and an idempotent $E^2 = E$, then

$$\begin{aligned} [A, [A, T]] &= E^2 T - 2ETE + TE^2 \\ &= ET - TE + 2TE - 2ETE \\ &= [E, T] + 2(I - E)TE. \end{aligned}$$

If the range of E belongs to $\mathcal{N}$, then $\operatorname{ran}(E) = N_i$ for some $N_i \in \mathcal{N}$. Since each subspace in the nest is invariant under T, we have $\operatorname{ran}(TE) \subseteq N_i = \operatorname{ran}(E)$ This then means that $(I - E)TE = 0$, and hence

$$[A, [A, T]] = [E, T] = [A, T].$$

Proof of Lemma 4.2.1, Part (b) ($\Leftarrow$)

Conversely, if A satisfies equation (2), it also satisfies equation (1) and, hence, $A = E + \lambda I$ for some $\lambda \in \mathscr{R}$ and $E^2 = E$. Then, for all $T \in \mathcal{T}$,

$$\begin{aligned} & & [E,[E,T]] &= [E,T] \\ &\Rightarrow & E^2T - 2ETE + TE^2 &= ET - TE \\ &\Rightarrow & ETE &= TE \qquad (*) \end{aligned}$$

We use the following proposition to show that this implies $\operatorname{ran}(E) \in \mathcal{N}$.

Proposition 4.2.3. *Let $\mathcal{T}(\mathcal{N})$ be an upper block-triangular algebra over a unique factorization domain $\mathscr{R}$, with corresponding nest $\mathcal{N}$. If the range of an idempotent $E \in \mathcal{E}(\mathcal{N})$ is invariant under all $T \in \mathcal{T}(\mathcal{N})$, then* $\operatorname{ran}(E) \in \mathcal{N}$.

Proof. There is a smallest subspace in the nest containing the range of E. Let $N_t = \min\{N_j \in \mathcal{N} : \operatorname{ran}(E) \subseteq N_j\}$. We recall that $N_t = \operatorname{span}\{e_i\}_{i=1}^{m_t}$.

For some $x = \sum x_i e_i \in \operatorname{ran}(E)$ we have $x_k \neq 0$ for $m_{t-1} < k \leq m_t$ since N_t is the smallest of the subspaces in the nest containing all such x.

Taking $T = \mathtt{E}_{kk} \in \mathcal{T}$ gives

$$TEx = Tx = x_k e_k \quad \text{and} \quad ETEx = Ex_k e_k \,.$$

The invariance of $\operatorname{ran}(E)$ is equivalent to $ETE = TE$ for all $T \in \mathcal{T}(\mathcal{N})$, and so we have $x_k e_k \in \operatorname{ran}(E)$. Using linearity and the cancellation property of $\mathscr{R}$ as an integral domain, the relation $x_k(Ee_k - e_k) = 0$ means $e_k \in \operatorname{ran}(E)$.

For every $s < k$ we may take $T = \mathtt{E}_{sk} \in \mathcal{T}$ where

$$TEe_k = Te_k = e_s \quad \text{and} \quad ETEe_k = Ee_s \,,$$

and so $e_s \in \operatorname{ran}(E)$.

In the other direction, for any $k < s \leq m_t$ we may use the basis matrix $\mathtt{E}_{sk}$ which has its 1 entry in the tt-block, so that $T = \mathtt{E}_{sk}$ is still in $\mathcal{T}$. Then, we again have $TEe_k = Te_k = e_s$ and $ETEe_k = Ee_s$, and so $e_s \in \operatorname{ran}(E)$.

Combined, we have $\text{span}\{e_i\}_{i=1}^{m_t} = N_t \subseteq \text{ran}(E)$. Therefore, $\text{ran}(E) = N_t \in \mathcal{N}$. □

We could have used the same basic argument to show that if $W \subseteq \mathscr{R}^n$ is a subspace invariant under all $T \in \mathcal{T}$, then $W = \lambda N_t$ for some $\lambda \in \mathscr{R}$ and $N_t \in \mathcal{N}$. However, we will not need this more general statement in this paper. In our case, since the subspace of interest is the range of an idempotent, we were able to use cancellation to show that $\lambda = 1$.

This result completes the proof of Lemma 4.2.1. □

– – – – – – –

Corollary 4.2.4. *Let $\mathcal{N}$ and $\mathcal{M}$ be nests and suppose that $\varphi : \mathcal{T}(\mathcal{N}) \to \mathcal{T}(\mathcal{M})$ is a Lie isomorphism. Then*

(a) $\varphi(I) = \lambda I$ where $\lambda \in \mathscr{R}^\times$ is an invertible scalar.

(b) If E is an idempotent in $\mathcal{T}(\mathcal{N})$, it follows that $\varphi(E) = F + \alpha_E I$ where $\alpha_E \in \mathscr{R}$ and F is an idempotent in $\mathcal{T}(\mathcal{M})$.

(c) If E is an idempotent in $\mathcal{T}(\mathcal{N})$ and $\text{ran}(E) \in \mathcal{N}$, then $\varphi(E) = F + \alpha_E I$, where $\alpha_E \in \mathscr{R}$, F is an idempotent in $\mathcal{T}(\mathcal{M})$ and $\text{ran}(F) \in \mathcal{M}$.

Furthermore, if $0 \neq E \neq I$, then both the scalar α_E and the idempotent F occurring above are uniquely determined.

Proof. Part (a) follows from Corollary 2.3.4, which shows that the center of $\mathcal{T}$ is the scalars. The scalar λ must be invertible, since if $\varphi^{-1}(I) = \eta I$, then $I = \varphi^{-1}(\varphi(I)) = \varphi^{-1}(\lambda I) = \lambda\eta I$.

Parts (b) and (c) follow directly from Lemma 4.2.1, since the Lie isomorphism φ necessarily preserves the bracket.

For the final assertion, if E is not a scalar, it is not in the center, and so $\varphi(E)$ is also not a scalar. Suppose $\varphi(E) = F_1 + \alpha_1 I = F_2 + \alpha_2 I$. Clearly, $\alpha_1 = \alpha_2 \iff F_1 = F_2$.

Supposing that $\alpha_1 \neq \alpha_2$, then

$$F_1 = F_2 + (\alpha_2 - \alpha_1)I = F_2^2 + 2(\alpha_2 - \alpha_1)F_2 + (\alpha_2 - \alpha_1)^2 I, \quad \text{and so}$$
$$(\alpha_2 - \alpha_1)I = 2(\alpha_2 - \alpha_1)F_2 + (\alpha_2 - \alpha_1)^2 I,$$

which implies that F_2 is a scalar, contradicting $\varphi(E)$ not being a scalar. $\square$

Definition 4.2.5 (Auxiliary Functions)**.** Let $\mathcal{N}$ and $\mathcal{M}$ be nests and suppose $\varphi : \mathcal{T}(\mathcal{N}) \to \mathcal{T}(\mathcal{M})$ is a Lie isomorphism. We define a function

$$\begin{aligned} \hat{\varphi} &: \mathcal{E}^0(\mathcal{N}) \to \mathcal{E}^0(\mathcal{M}), \quad \text{by} \\ \hat{\varphi}(E) &= \varphi(E) - \alpha_E I. \end{aligned}$$

In the notation of Corollary 4.2.4, if $\varphi(E) = F + \alpha_E I$, then $\hat{\varphi}(E) = F$. In addition, we define

$$\begin{aligned} \tilde{\varphi} &: \mathcal{N}^0 \to \mathcal{M}^0, \quad \text{by} \\ \tilde{\varphi}(N_i) &= \operatorname{ran}(\varphi(P_i) - \alpha_{P_i} I). \end{aligned}$$

In other words, $\tilde{\varphi}(N_i) = \operatorname{ran}(\hat{\varphi}(P_i))$.

Next, we characterize idempotents that have the same range in terms of commutators.

Lemma 4.2.6. *Let E and F be nonzero idempotent linear maps on a finite rank free module over an integral domain, not of characteristic 2. The following are equivalent:*

(a) $EF - FE = F - E + \lambda I$ *for a scalar* λ;

(b) $EF - FE = F - E$;

(c) $EF = F$ *and* $FE = E$.

Proof. We see directly that (c) $\Rightarrow$ (b) $\Rightarrow$ (a). To prove (a) $\Rightarrow$ (c), we assume that $EF - FE = F - E + \lambda I$ and consider

$$\begin{aligned} 0 &= (I - E)(EF - FE + F)E = (I - E)(F - E + \lambda I + F)E \\ &= 2(I - E)FE, \quad \text{and} \\ 0 &= E(EF - FE)E = E(F - E + \lambda I)E = EFE - (1 - \lambda)E. \end{aligned}$$

Combined, this gives

$$FE = EFE = (1 - \lambda)E.$$

Thus, if $x \in \operatorname{ran}(E)$, then $Fx = (1-\lambda)x$, which shows that $(1-\lambda)$ is an eigenvalue of F. Since F is idempotent and the ring is an integral domain (where the only idempotents are 0 and 1), we have $\lambda = 0$ or 1. Similarly,

$$\begin{aligned} 0 &= (I-F)(EF-FE-E)F = (I-F)(F-E+\lambda I-E)F \\ &= -2(I-F)EF, \quad \text{and} \\ 0 &= F(EF-FE)F = F(F-E+\lambda I)F = (1+\lambda)F - FEF, \end{aligned}$$

which yields $EF = (1+\lambda)F$, and so $(1+\lambda)$ is an eigenvalue of E. Together with the above, this means $\lambda = 0$ and $EF = F$, $FE = E$, as desired. □

Proposition 4.2.7. *Let $\varphi : \mathcal{T}(\mathcal{N}) \to \mathcal{T}(\mathcal{M})$ be a Lie isomorphism of upper block-triangular algebras, for associated nests $\mathcal{N}$ and $\mathcal{M}$. If $E, F \in \mathcal{E}^0(\mathcal{N})$ such that* $\operatorname{ran}(E) = \operatorname{ran}(F)$*, then $\alpha_E = \alpha_F$ and* $\operatorname{ran}\hat{\varphi}(E) = \operatorname{ran}\hat{\varphi}(F)$.

Proof. Since $\operatorname{ran}(E) = \operatorname{ran}(F)$, we have $EF = F$ and $FE = E$. Hence, $[E,F] = F - E$. It follows from the Lie isomorphism that $[\varphi(E), \varphi(F)] = \varphi([E,F]) = \varphi(F) - \varphi(F)$ which can be rewritten

$$[\hat{\varphi}(E), \hat{\varphi}(F)] = \hat{\varphi}(E) - \hat{\varphi}(F) + (\alpha_F - \alpha_E)I.$$

By Lemma 4.2.6 we get $\alpha_E = \alpha_F$, that $\hat{\varphi}(E)\hat{\varphi}(F) = \hat{\varphi}(F)$ and $\hat{\varphi}(F)\hat{\varphi}(E) = \hat{\varphi}(E)$. Hence, $\operatorname{ran}\hat{\varphi}(E) = \operatorname{ran}\hat{\varphi}(F)$. □

4.3 Order Structure

The next step is to show that $\hat{\varphi}$ preserves the *order* structure of the idempotents in $\mathcal{E}(\mathcal{N})$.

Definition 4.3.1. Let E_1 and E_2 be two idempotents. We say that $E_1 \le E_2$ if $E_1E_2 = E_1 = E_2E_1$. This is equivalent to E_1 and E_2 commuting and $\operatorname{ran}(E_1) \subseteq \operatorname{ran}(E_2)$. We say that $E_1 < E_2$ if $E_1 \le E_2$ but $E_1 \ne E_2$.

We recall that it is possible to have $E_1E_2 = E_2E_1$, but not have the idempotents satisfy the containment condition for the ranges.

Lemma 4.3.2. *Let $\varphi : \mathcal{T}(\mathcal{N}) \to \mathcal{T}(\mathcal{M})$ be a Lie isomorphism of upper block-triangular algebras, for associated nests $\mathcal{N}$ and $\mathcal{M}$. Suppose that $E_1, E_2 \in \mathcal{E}(\mathcal{N})$ such that $0 < E_1 < E_2 < I$. Set $F_i = \hat{\varphi}(E_i)$ in $\mathcal{E}(\mathcal{M})$ for $i = 1, 2$. Then, either $0 < F_1 < F_2 < I$ or $0 < F_2 < F_1 < I$.*

This shows that φ preserves comparability within the partially ordered set $\mathcal{E}(\mathcal{N})$.

Proof. We observe that $0 = \varphi([E_1, E_2]) = [\varphi(E_1), \varphi(E_2)] = [\hat{\varphi}(E_1), \hat{\varphi}(E_2)] = [F_1, F_2]$, where we may move from φ to $\hat{\varphi}$, since the bracket operation ignores scalars. Hence, F_1 and F_2 commute.

By assumption, we have $E_1, E_2 \notin \mathscr{R}I$. We see that

$$(E_2 - E_1)^2 = E_2^2 - 2E_2E_1 + E_1^2 = E_2 - 2E_1 + E_1 = E_2 - E_1.$$

In addition, $E_2(E_2 - E_1) = (E_2 - E_1) = (E_2 - E_1)E_2$ so $(E_2 - E_1) < E_2$. Hence, $(E_2 - E_1) \notin \mathscr{R}I$. If $E \in \mathcal{E}^0(\mathcal{N})$ then $\varphi(E)$ is also not a scalar, since φ is bijective and maps the center $Z(\mathcal{T}(\mathcal{N})) = \mathscr{R}I$ to the center $Z(\mathcal{T}(\mathcal{M})) = \mathscr{R}I$. Hence $F_1, F_2, (F_2 - F_1) \notin \mathscr{R}I$ and $F_1 \neq F_2$

From Corollary 4.2.4 we know that $\varphi(E_2 - E_1)$ is a scalar plus an idempotent. Hence, the eigenvalues of $(F_2 - F_1)$ are λ and $\lambda + 1$ for some $\lambda \in \mathscr{R}$. We may view F_1 and F_2 as matrices in $M_n(\mathbb{K})$, where $\mathbb{K}$ is the fraction field of $\mathscr{R}$. We may choose a basis for $\mathbb{K}^n$ that diagonalizes F_1 and F_2 simultaneously. If F_1 and F_2 are not comparable, then the eigenvalues of $(F_2 - F_1)$ include 1 and -1, which is a contradiction. □

Lemma 4.3.3. *Let $\varphi : \mathcal{T}(\mathcal{N}) \to \mathcal{T}(\mathcal{M})$ be a Lie isomorphism of upper block-triangular algebras, for associated nests $\mathcal{N}$ and $\mathcal{M}$. Suppose that $E_1, E_2, E_3 \in \mathcal{E}(\mathcal{N})$ such that $0 < E_1 < E_2 < E_3 < I$. Set $F_i = \hat{\varphi}(E_i)$ in $\mathcal{E}(\mathcal{M})$ for $i = 1, 2, 3$.*

(a) If $F_1 < F_2$, then $0 < F_1 < F_2 < F_3 < I$,

(b) If $F_1 > F_2$, then $I > F_1 > F_2 > F_3 > 0$.

Consequently, the map $\hat{\varphi}$ (and also $\tilde{\varphi}$) is either order preserving or order reversing.

Proof. The idempotents F_1, F_2 and F_3 are distinct, nonzero, and mutually comparable by Lemma 4.3.2.

(a) We note that $E_1 - (E_2 - E_3)$ is idempotent and so $F_1 - F_2 + F_3$ is a scalar plus an idempotent, with eigenvalues λ and $\lambda + 1$ for some $\lambda \in \mathscr{R}$. As with Lemma 4.3.2, we may view F_1, F_2 and F_3 as matrices in $M_n(\mathbb{K})$, and choose a basis for $\mathbb{K}^n$ that diagonalizes them simultaneously. However, if $F_1 < F_3 < F_2$ or if $F_3 < F_1$ then the eigenvalues of $F_1 - F_2 + F_3$ include 0, 1 and -1, which is a contradiction.

Part (b) can be proven in an analogous manner. $\square$

We shall refer to the Lie isomorphism φ itself as being *order preserving* or *order reversing* according to whether $\hat{\varphi}$ is order preserving or order reversing, respectively. We now extend the definition of $\hat{\varphi}$ and $\tilde{\varphi}$ to all of $\mathcal{E}(\mathcal{N})$ and $\mathcal{N}$, respectively.

Definition 4.3.4. If $\varphi : \mathcal{T}(\mathcal{N}) \to \mathcal{T}(\mathcal{M})$ is an order preserving Lie isomorphism, then we define

$$\hat{\varphi}(0) = 0, \quad \hat{\varphi}(I) = I, \quad \tilde{\varphi}(\{0\}) = \{0\}, \quad \tilde{\varphi}(\mathscr{R}^n) = \mathscr{R}^n.$$

If φ is order reversing, we define

$$\hat{\varphi}(0) = I, \quad \hat{\varphi}(I) = 0, \quad \tilde{\varphi}(\{0\}) = \mathscr{R}^n, \quad \tilde{\varphi}(\mathscr{R}^n) = \{0\}.$$

4.4 Dimension Preservation

Theorem 4.4.1. *Let $\varphi : \mathcal{T}(\mathcal{N}) \to \mathcal{T}(\mathcal{M})$ be a Lie isomorphism of upper block-triangular algebras, for associated nests $\mathcal{N}$ and $\mathcal{M}$. Then either*

(a) $\tilde{\varphi}$ is a bijective dimension-preserving order-preserving map between $\mathcal{N}$ and $\mathcal{M}$, or

(b) the map $\tilde{\varphi}^{\angle}$ defined via $\tilde{\varphi}^{\angle}(N) = (\tilde{\varphi}(N))^{\angle}$ is a bijective dimension-preserving order-preserving map between $\mathcal{N}$ and $\mathcal{M}^{\angle}$.

Proof. The order properties of $\tilde{\varphi}$ were established in Lemma 4.3.3. For each $N_i \in \mathcal{N}$, $\tilde{\varphi}(N_i)$ is in $\mathcal{M}$ (and $\tilde{\varphi}$ is injective) so the number of nest subsets in $\mathcal{N}$ is less than or equal to the number of subsets in $\mathcal{M}$. This is also true for $\widetilde{\varphi^{-1}}$ in the opposite direction, so $\tilde{\varphi}$ is a bijection with $\tilde{\varphi}^{-1} = \widetilde{\varphi^{-1}}$. We wish to show that $\tilde{\varphi}$ is dimension preserving.

Let $\mathcal{T}(\mathcal{N}) = T(n_1, \ldots, n_k)(\mathscr{R}) \subseteq M_n(\mathscr{R})$. As in the notation section, we define $m_0 = 0$ and $m_t = \sum_{i=1}^{t} n_i$ for $1 \leq t \leq k$. We see that $m_k = n$ and $n_t = m_t - m_{t-1}$. Define $N_0 = \{0\}$ and otherwise $N_t = \operatorname{span}\{e_i\}_{i=1}^{m_t}$ so the nest of subspaces $\mathcal{N} = \{N_t\}_{t=0}^{k}$. Recall that $N_t \ominus N_{t-1} = \operatorname{span}\{e_i\}_{i=m_{t-1}+1}^{m_t}$. For $t \in \{1, \ldots, k\}$ we have $\dim(N_t \ominus N_{t-1}) = m_t - m_{t-1} = n_t$.

Define the intermediate projections in $\mathcal{T}(\mathcal{N})$ to be $Q_j = \sum_{i=1}^{m_{t-1}+j} \mathtt{E}_{ii}$ for $j \in \{0, \ldots, n_t\}$, so that

$$P_{t-1} = Q_0 < Q_1 < Q_2 < \cdots < Q_{n_t} = P_t.$$

Let $F_j = \hat{\varphi}(Q_j)$ for $j \in \{0, \ldots, n_t\}$, where $\operatorname{ran}(F_0) = \tilde{\varphi}(N_{t-1})$ and $\operatorname{ran}(F_{n_t}) = \tilde{\varphi}(N_t)$ are subspaces in the nest $\mathcal{M}$.

(a) Assuming that $\hat{\varphi}$ is order preserving, we wish to determine the dimension of $\tilde{\varphi}(N_t) \ominus \tilde{\varphi}(N_{t-1})$. In $\mathcal{T}(\mathcal{M})$ we have

$$F_0 < F_1 < F_2 < \cdots < F_{n_t},$$

where $\operatorname{ran}(F_{i-1}) \subsetneq \operatorname{ran}(F_i)$ for all $i \in \{1, \ldots, n_t\}$. For each such i, we may find $x_i \in \operatorname{ran}(F_i)$ such that $x_i \notin \operatorname{ran}(F_{i-1})$. We show that $\{x_i\}_{i=1}^{n_t}$ is an independent set.

Suppose that $\sum_{i=1}^{n_t} \alpha_i x_i = 0$ for some $\alpha_i \in \mathscr{R}$. Since $x_i \in \operatorname{ran}(F_j)$ for $j \geq i$, then $(I - F_j)x_i = 0$ if $j \geq i$, but $(I - F_j)x_i \neq 0$ if $j < i$.

Starting at n_t, we have $0 = (I - F_{n_t-1}) \sum_{i=1}^{n_t} x_i = \alpha_{n_t}(I - F_{n_t-1})x_{n_t}$. Since $(I - F_{n_t-1})x_{n_t} \neq 0$, we have $\alpha_{n_t} = 0$. Working backwards from n_t to 1, we then see that $\alpha_i = 0$ for all $i \in \{1, \ldots, n_t\}$, and so $\{x_i\}_{i=1}^{n_t}$ is an independent set.

We observe that, since $\mathscr{R}$ is an integral domain with cancellation, if $x_i \notin \operatorname{ran}(F_j)$, then $\operatorname{span}\{x_i\} \cap \operatorname{ran}(F_j) = \{0\}$. Explicitly, if $\lambda x_i \in \operatorname{ran}(F_j)$ then $\lambda(F_j x_i - x_i) = 0$, which implies $\lambda = 0$.

Here, all $x_i \in \operatorname{ran}(F_{n_t}) = \tilde{\varphi}(N_t)$ but all $x_i \notin \operatorname{ran}(F_0) = \tilde{\varphi}(N_{t-1})$. Hence $\operatorname{span}\{x_i\}_{i=1}^{n_t} \subseteq \tilde{\varphi}(N_t) \ominus \tilde{\varphi}(N_{t-1})$ and so we have $\dim(\tilde{\varphi}(N_t) \ominus \tilde{\varphi}(N_{t-1})) \geq n_t$. Applying the same argument to φ^{-1} proves the reverse inequality and hence $\tilde{\varphi}$ is dimension preserving, as claimed.

(b) If $\hat{\varphi}$ is order reversing, we have

$$F_0 > F_1 > F_2 > \cdots > F_{n_t}\,.$$

Using the reversed order and the same range inclusion argument from part (a), we get $\dim(\tilde{\varphi}(N_{t-1}) \ominus \tilde{\varphi}(N_t)) \geq n_t$. Again, the same observation applied to φ^{-1} proves the reverse inequality and hence $\tilde{\varphi}$ is also dimension preserving in the order reversing case.

It then follows that the map $\tilde{\varphi}^{\angle}$ defined in (b) is a bijective dimension-preserving order-preserving map between $\mathcal{N}$ and $\mathcal{M}^{\angle}$. □

4.5 Full Matrix (Trivial Nest) Case

In this case, let $\mathcal{T}(\mathcal{N}) = M_n(\mathscr{R})$ be the full matrix algebra with $n \geq 2$, where the associated nest $\mathcal{N} = \{\{0\}, \mathscr{R}^n\}$ is trivial. As before, let $\varphi : \mathcal{T}(\mathcal{N}) \to \mathcal{T}(\mathcal{M})$ be a Lie isomorphism.

Given any subspace $M_i \in \mathcal{M}$, let $F \in \mathcal{T}(\mathcal{M})$ be the projection onto M_i. Using Corollary 4.2.4 (c), we know that $\varphi^{-1}(F) = E + \alpha_F I$ where E is an idempotent whose range is in $\mathcal{N}$. However, $\mathcal{N}$ is trivial and so E is either I or 0. Hence, $\mathcal{M}$ contains only two subspaces, making $\mathcal{M}$ also trivial.

Since φ is also a (dimension/rank preserving) linear isomorphism of modules, $\mathcal{T}(\mathcal{N}) = M_n(\mathscr{R}) = \mathcal{T}(\mathcal{M})$ and thus, φ is a Lie automorphism. Hence, it is sufficient to consider Lie automorphisms of $M_n(\mathscr{R})$.

Let $\mathbb{K}$ be the fraction field of $\mathscr{R}$. Since φ is a linear bijection, we may embed $M_n(\mathscr{R})$ in $M_n(\mathbb{K})$ and extend the map by linearity to $\check{\varphi} : M_n(\mathbb{K}) \to M_n(\mathbb{K})$. The Lie product is preserved under φ and is determined by its action on the basis matrices $\{\mathtt{E}_{ij}\}$ via the structure equations. Since the set $\{\mathtt{E}_{ij}\}$ gives the same basis (the standard basis as a module/vector space) for both $M_n(\mathscr{R})$ and $M_n(\mathbb{K})$, and the Lie product is bilinear, $\check{\varphi}$ is also a Lie automorphism.

$$\begin{array}{ccc} M_n(\mathscr{R}) & \xrightarrow{\ \varphi\ } & M_n(\mathscr{R}) \\ \downarrow{\scriptstyle incl} & & \downarrow{\scriptstyle incl} \\ M_n(\mathbb{K}) & \xrightarrow{\ \check{\varphi}\ } & M_n(\mathbb{K}) \end{array}$$

When $\mathbb{K}$ is a field, it is a straightforward exercise to show that $M_n(\mathbb{K})$ is a simple ring. By Theorem 2.4.6 (Martindale), $\check{\varphi}$ is of the form $\sigma + \tau$ where τ is an additive map into the center which maps commutators to zero, and σ is either an automorphism or the negative of an anti-automorphism of $M_n(\mathbb{K})$.

An additive map on $M_n(\mathbb{K})$ into the center which maps commutators to zero is a generalized trace (the center of $M_n(\mathbb{K})$ is isomorphic to $\mathbb{K}$) and so given an arbitrary $A \in M_n(\mathbb{K})$ we have $\tau(A) = \frac{\beta}{\alpha} I$ for some $\alpha, \beta \in \mathscr{R}$ such that $\frac{\beta}{\alpha} \in \mathbb{K}$.

We consider the automorphism case first. By Theorem 2.4.2 (Skolem-Noether) associative automorphisms of a full matrix algebra over a field are inner, so $\sigma(A) = S^{-1}AS$ for some $S \in GL_n(\mathbb{K})$.

Given $A \in M_n(\mathscr{R}) \subseteq M_n(\mathbb{K})$, since $\check{\varphi}|_{M_n(\mathscr{R})} = \varphi$, we have $\varphi(A) = \check{\varphi}(A) = S^{-1}AS + \frac{\beta}{\alpha} I$ for some $\alpha, \beta \in \mathscr{R}$ such that $\frac{\beta}{\alpha} \in \mathbb{K}$. We may assume that α and β are coprime.

We note that φ is a linear bijection, so $S^{-1}AS + \frac{\beta}{\alpha} I \in M_n(\mathscr{R})$ even though $S \in GL_n(\mathbb{K})$ and $\frac{\beta}{\alpha} \in \mathbb{K}$. We wish to show that $\frac{\beta}{\alpha} \in \mathscr{R}$ and $S^{-1}AS \in M_n(\mathscr{R})$.

The inner automorphism is characterized by its action on the basis matrices $\{\mathtt{E}_{ij}\}$. As we noted in the discussion of traces, if $i \neq j$ we can write $\mathtt{E}_{ij} = [\mathtt{E}_{ii}, \mathtt{E}_{ij}]$ to get $\tau(\mathtt{E}_{ij}) = 0$, and so $S^{-1}\mathtt{E}_{ij}S \in M_n(\mathscr{R})$. Given $A = \mathtt{E}_{ii}$ for any $i \in \{1, \ldots, n\}$, assume that

$$\varphi(\mathtt{E}_{ii}) = \check{\varphi}(\mathtt{E}_{ii}) = S^{-1}\mathtt{E}_{ii}S + \frac{\beta}{\alpha} I \quad \in M_n(\mathscr{R}).$$

We note that $S^{-1}\mathtt{E}_{ii}S$ is an idempotent in $M_n(\mathbb{K})$ and so a slight variation on Proposition 4.2.2 gives us $\frac{\beta}{\alpha} \in \mathscr{R}$ and $S^{-1}\mathtt{E}_{ii}S \in M_n(\mathscr{R})$.

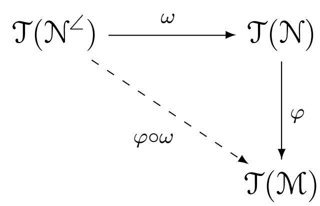
As with the field case, every associative automorphism of $M_n(\mathscr{R})$ is inner when $\mathscr{R}$ is a UFD (Theorem 2.4.3). Hence $S \in GL_n(\mathscr{R})$ and this gives case (a) of Theorem 4.1.1 for the trivial nest.

If σ is the negative of an anti-automorphism, we may compose it with ω. As an automorphism, $\sigma \circ \omega = -\sigma(JA^\top J) = S^{-1}AS$ for some $S \in GL_n(\mathscr{R})$. But then $\sigma(A) = -S^{-1}JA^\top JS$, which matches case (b) of Theorem 4.1.1.

We note that in the full matrix case, $JS \in GL_n(\mathscr{R})$ and so we could have viewed the negative anti-automorphism as just the map $A \mapsto -A^\top$.

4.6 Decomposition using Projections

From this point forward, we restrict the arguments to the case where φ is order preserving. If φ is order reversing, we may compose with ω to form an order-preserving Lie isomorphism $\varphi \circ \omega : \mathcal{T}(\mathcal{N}^{\angle}) \to \mathcal{T}(\mathcal{M})$ where $\varphi \circ \omega(A) = -\varphi(JA^{\top}J)$.



If $\varphi : \mathcal{T}(\mathcal{N}) \to \mathcal{T}(\mathcal{M})$ is order preserving and dimension preserving, the nest structure of $\mathcal{M}$ is the same as the nest structure of $\mathcal{N}$. Hence, as sets of matrices, we have $\mathcal{T}(\mathcal{N}) = \mathcal{T}(\mathcal{M})$. Thus, from here forward we will consider our map to be a Lie automorphism $\varphi : \mathcal{T}(\mathcal{N}) \to \mathcal{T}(\mathcal{N})$.

For such a Lie automorphism $\varphi : \mathcal{T}(\mathcal{N}) \to \mathcal{T}(\mathcal{N})$ and nest projections $\mathcal{P}(\mathcal{N}) = \{P_t\}_{t=0}^{k}$, we know from the previous results that $\hat{\varphi}(P(\mathcal{N})) = \{\hat{\varphi}(P_t)\}_{t=0}^{k}$ will be a collection of idempotents with ranges corresponding to the nest subspaces in $\mathcal{N}$. However, these idempotents may not be the "standard" nest projections.

Lemma 4.6.1. *Let $\varphi : \mathcal{T}(\mathcal{N}) \to \mathcal{T}(\mathcal{N})$ be an order-preserving Lie automorphism. Then there exists an invertible matrix $R \in \mathcal{T}(\mathcal{N})$ such that $P_t = R^{-1}\hat{\varphi}(P_t)R$ for all $P_t \in \mathcal{P}(\mathcal{N})$.*

Proof. For each $P_t \in \mathcal{P}(\mathcal{N})$ denote the idempotents $F_t = \hat{\varphi}(P_t)$. Given that $\mathcal{T}(\mathcal{N}) = T(n_1, \ldots, n_k)(\mathscr{R})$, recall the index $m_t = \sum_{i=1}^{k} n_i$ for $1 \leq t \leq k$.

Let $1 \leq t < k$ be fixed. Since the range of the idempotent F_t is in the nest, $\operatorname{ran}(F_t) = N_t$, then it preserves the first m_t basis elements (from the automorphism being order and dimension preserving).

If $F_t e_j = e_j$ for $j \in \{1, \ldots, m_t\}$, then the first m_t columns of F_t correspond to the standard basis elements. In other words the upper left corner $m_t \times m_t$ block is the identity matrix and the block below it is zero. As F_t is idempotent, the bottom rows from index $m_t + 1$ to n will be zero, since nothing in the span of $\{e_j\}_{j=m_t+1}^{n}$ is in the image of F_t.

Divided at m_t for the columns and rows, F_t can be written as a block matrix,

$$F_t = \left[\begin{array}{c:c} I & X \\ \hdashline 0 & 0 \end{array}\right].$$

Hence, we may construct

$$S_t = \left[\begin{array}{c:c} I & -X \\ \hdashline 0 & I \end{array}\right], \quad \text{and} \qquad S_t^{-1} = \left[\begin{array}{c:c} I & X \\ \hdashline 0 & I \end{array}\right], \quad \text{so} \qquad S_t^{-1} F_t S_t = \left[\begin{array}{c:c} I & 0 \\ \hdashline 0 & 0 \end{array}\right] = P_t.$$

Given $t < k$, we know that $F_{t+1} > F_t$ and since $N_t \subseteq N_{t+1}$ we may use the same block form to write

$$F_{t+1} = \left[\begin{array}{c:c} I & B_{12} \\ \hdashline 0 & B_{22} \end{array}\right], \quad \text{so that} \quad S_t^{-1} F_{t+1} S_t = \left[\begin{array}{c:c} I & \hat{B}_{12} \\ \hdashline 0 & B_{22} \end{array}\right].$$

where $\hat{B}_{12} = B_{12} + XB_{22} - B_{11}X$.

Then, since $F_t < F_{t+1}$ we have

$$P_t(S_t^{-1} F_{t+1} S_t) = (S_t^{-1} F_t S_t)(S_t^{-1} F_{t+1} S_t) = S_t^{-1}(F_t F_{t+1}) S_t = S_t^{-1} F_t S_t = P_t,$$

or explicitly,

$$\left[\begin{array}{c:c} I & \hat{B}_{12} \\ \hdashline 0 & 0 \end{array}\right] = \left[\begin{array}{c:c} I & 0 \\ \hdashline 0 & 0 \end{array}\right].$$

Hence, $\hat{B}_{12} = 0$, giving

$$S_t^{-1} F_{t+1} S_t = \left[\begin{array}{c:c} I & 0 \\ \hdashline 0 & B_{22} \end{array}\right],$$

where B_{22} is an $(n - m_t) \times (n - m_t)$ idempotent matrix in $T(n_{t+1}, \ldots, n_k)(\mathscr{R})$. We can think of this smaller upper block-triangular algebra $T(n_{t+1}, \ldots, n_k)(\mathscr{R})$ as acting on the nest $\hat{\mathcal{N}} = \{N_i \ominus N_t : t \leq i \leq k\}$. Since $\operatorname{ran}(F_{t+1}) = N_{t+1}$, we have $\operatorname{ran}(B_{22}) = (N_{t+1} \ominus N_t) \in \hat{\mathcal{N}}$.

As in the above process, we can then find a matrix $\hat{S}_{t+1}$ such that $\hat{S}_{t+1}^{-1} B_{22} \hat{S}_{t+1}$ acts as the standard projection on $(N_{t+1} \ominus N_t)$. We define

$$S_{t+1} = \left[\begin{array}{c:c} I & 0 \\ \hdashline 0 & \hat{S}_{t+1} \end{array}\right],$$

so that $(S_{t+1}^{-1} S_t^{-1}) F_{t+1} (S_{t+1} S_t) = P_{t+1}$.

It is useful to note that $(S_{t+1}^{-1} S_t^{-1}) F_t (S_t S_{t+1}) = S_{t+1}^{-1} P_t S_{t+1} = P_t$, since the lower right block of S_{t+1} does not affect P_t.

Using induction we can define $R = \prod_{t=1}^{k} S_t$, which is in $\mathcal{T}$ since it is the product of matrices $S_t \in \mathcal{T}$. We have now found an invertible matrix $R \in \mathcal{T}$ such that $R^{-1} F_t R = P_t$ for all $F_t = \hat{\varphi}(P_t)$, as desired. □

Remark 4.6.2. By composing φ with the map ω if necessary, we have an order and dimension preserving isomorphism from $\mathcal{T}(\mathcal{N})$ to $\mathcal{T}(\mathcal{M})$. Hence $\mathcal{N} = \mathcal{M}$ and so, as mentioned above, we consider only order-preserving Lie automorphisms $\varphi : \mathcal{T}(\mathcal{N}) \to \mathcal{T}(\mathcal{N})$.

We use the notation $\mathrm{Ad}_R(A) = R^{-1} A R$ and note that Ad_R is an (associative) inner automorphism of $\mathcal{T}(\mathcal{N})$, and thus an order-preserving Lie automorphism. By composition,

$$\mathcal{T}(\mathcal{N}) \xrightarrow{\ \varphi\ } \mathcal{T}(\mathcal{N}) \xrightarrow{\ \mathrm{Ad}_R\ } \mathcal{T}(\mathcal{N}),$$

$$\mathcal{P}(\mathcal{N}) \xrightarrow{\ \hat{\varphi}\ } \hat{\varphi}(\mathcal{P}(\mathcal{N})) \xrightarrow{\ \mathrm{Ad}_R\ } \mathcal{P}(\mathcal{N}).$$

We may thus define a new map $\rho := \mathrm{Ad}_R \circ \varphi$ to get an order-preserving Lie automorphism of $\mathcal{T}(\mathcal{N})$ such that

$$\rho(P_t) = P_t + \lambda_{P_t} I,$$

for all $P_t \in \mathcal{P}(\mathcal{N})$.

Hence, we have reduced the problem of classifying Lie isomorphisms between two upper block-triangular algebras to that of characterizing Lie automorphisms of a fixed upper block-triangular algebra which preserve the nest projections modulo the scalars.

Remark 4.6.3. For the following we will use the **Peirce decomposition** of a matrix, using a projection $P \in \mathcal{P}(\mathcal{N})$. We may write any $A \in \mathcal{T}(\mathcal{N})$ as

$$A = PAP + PA(I-P) + (I-P)AP + (I-P)A(I-P)\,.$$

This is sometimes written in block-matrix form as

$$A = \left[\begin{array}{c:c} PAP & PA(I-P) \\ \hdashline (I-P)AP & (I-P)A(I-P) \end{array}\right]$$

It is important to recognize that this is an abuse of notation. For example, PAP is a full matrix and not a submatrix restricted to the upper left block. However, since it has zeros in all the other blocks, it is often convenient to consider it as equivalent to its restriction to this smaller block.

In our case, since the algebra is upper block-triangular, if $A \in \mathcal{T}(\mathcal{N})$ we have $(I - P)AP = 0$ for any $P \in \mathcal{P}(\mathcal{N})$.

Proposition 4.6.4. *Let $\mathcal{T}$ be an upper block-triangular algebra with the associated nest $\mathcal{N}$, and let $P \in \mathcal{P}(\mathcal{N})$. Suppose that $\rho : \mathcal{T} \to \mathcal{T}$ is a Lie automorphism that for each $Q \in \mathcal{P}(\mathcal{N})$,*

$$\rho(Q) = Q + \lambda_Q I,$$

for some $\lambda_Q \in \mathscr{R}$. Then,

(a) $\rho(P\mathcal{T}(I-P)) = P\mathcal{T}(I-P)$;

(b) $\rho(P\mathcal{T}P) + \mathscr{R}I = P\mathcal{T}P + \mathscr{R}I$.

Proof. (a) If $Z \in P\mathcal{T}(I-P)$, then $Z = [P, Z]$ and hence

$$\rho(Z) = \rho[P,Z] = [\rho(P), \rho(Z)] = [P + \lambda_P I, \rho(Z)] = [P, \rho(Z)],$$

implying that $\rho(Z) \in P\mathcal{T}(I-P)$.

This proves the inclusion $\rho(P\mathcal{T}(I-P)) \subseteq P\mathcal{T}(I-P)$. The reverse inclusion follows from using the same argument for ρ^{-1}.

(b) Given $T \in P\mathfrak{T}P$, we have $[T, P] = 0$ so $[\rho(T), P] = 0$ by a similar argument to above. Hence, P commutes with $\rho(T)$ and therefore $P\rho(T)(I - P) = 0$.

Define the projections $Q_i = P + \sum_{j=p+1}^{p+i} \mathtt{E}_{jj} = \sum_{j=1}^{p+i} \mathtt{E}_{jj}$ for $i \in \{1, 2, \ldots, n-p\}$. We note that $P < Q_1 < Q_2 < \cdots < Q_{n-p} = I$, where the dimension of the range of each Q_i is one greater than the projection before it. By the order preservation properties of $\hat{\rho}$ there exists idempotents in $\mathcal{E}(\mathcal{N})$ which we denote $F_i = \hat{\rho}^{-1}(Q_i)$ such that $P < F_1 < F_2 < \cdots < F_{n-p} = I$.

Since $F_i \geq P$, we know that $F_i P = P = P F_i$ and so $[T, F_i] = 0$ as well. Thus, for $i \in \{1, 2, \ldots, n-p\}$,

$$0 = \rho([T, F_i]) = [\rho(T), \rho(F_i)] = [\rho(T), \hat{\rho}(F_i)] = [\rho(T), Q_i].$$

For example, with $A = \rho(T)$ and $i = 1$ we have

$$Q_1 = \left[\begin{array}{c|ccc} I & & 0 & \\ \hline & 1 & 0 & \\ 0 & 0 & 0 & \\ & & & \ddots \end{array}\right], \text{ and } A = \left[\begin{array}{c|ccc} A_{11} & & 0 & \\ \hline & a_{p+1,p+1} & a_{p+1,p+2} & \\ 0 & a_{p+2,p+1} & a_{p+2,p+2} & \\ & & & \ddots \end{array}\right], \text{ so}$$

$$Q_1 A = \left[\begin{array}{c|ccc} A_{11} & & 0 & \\ \hline & a_{p+1,p+1} & a_{p+1,p+2} & \cdots \\ 0 & 0 & 0 & \\ & & & \ddots \end{array}\right] = \left[\begin{array}{c|ccc} A_{11} & & 0 & \\ \hline & a_{p+1,p+1} & 0 & \\ 0 & a_{p+2,p+1} & 0 & \\ & \vdots & & \ddots \end{array}\right] = AQ_1 .$$

By ranging over all the Q_i this implies that

$$(I - P)\rho(T)(I - P) = \sum_{j=p+1}^{p+i} a_{jj}\mathtt{E}_{jj} ,$$

where the only nonzero entries are on the diagonal of the matrix. We seek to show that all these a_{jj} are the same.

We define $\mathcal{L}_P := P\mathcal{T}$, which is a Lie ideal in $\mathcal{T}$. Hence, the image $\rho(\mathcal{L}_P)$ is also a Lie ideal in $\mathcal{T}$.

For any element $T \in \mathcal{L}_P$, the "top right corner" $PT(I-P)$ is mapped by ρ back into $P\mathcal{T}(I-P)$ by part (a) above. The "top left corner" PTP is mapped into $P\mathcal{T}P$ plus nonzero entries along the diagonal below the p row.

Let $A = \rho(T) \in \rho(\mathcal{L}_P)$. Taking $\mathtt{E}_{in}$ with $p < i < n$, we see that $[A, \mathtt{E}_{in}] \in \rho(\mathcal{L}_P)$, but the in^{th}-entry of $[A, \mathtt{E}_{in}]$ will be $a_{ii} - a_{nn}$ which equals zero, since anything off the diagonal in $(I-P)\rho(\mathcal{L}_P)(I-P)$ is zero. This means that $a_{ii} = a_{nn}$ for all $p < i < n$ and so

$$A = PAP + \lambda(I-P)$$

for some $\lambda \in \mathscr{R}$. By adjusting the top left corner by a scalar if needed, we get $\rho(T) = S + \lambda I$ for $S \in P\mathcal{T}P$, so

$$\rho(P\mathcal{T}P) + \mathscr{R}I \subseteq P\mathcal{T}P + \mathscr{R}I.$$

By using the same argument for ρ^{-1}, we get the reverse inclusion and so equality follows. □

Proposition 4.6.5. *Let $\mathcal{T}(\mathcal{N})$ be an upper block-triangular algebra with the associated nest $\mathcal{N}$, and take $P \in \mathcal{P}(\mathcal{N})$ where $P \neq 0, I$. Suppose that $\rho : \mathcal{T}(\mathcal{N}) \to \mathcal{T}(\mathcal{N})$ is an order-preserving Lie automorphism of $\mathcal{T}$ and that for each $Q \in \mathcal{P}(\mathcal{N})$, $\rho(Q) = Q + \lambda_Q I$. Then for every $T \in P\mathcal{T}P$, we can write*

$$\rho(T) = \psi_1(T) + \gamma_1(T)I,$$

where ψ_1 is an associative automorphism of $P\mathcal{T}P$, and γ_1 is a linear functional on $P\mathcal{T}P$ that annihilates all commutators $[A, B]$ for $A, B \in P\mathcal{T}P$.

Proof. By Proposition 4.6.4, we see that if $T \in P\mathcal{T}P$, then $\rho(T) \in P\mathcal{T}P + \mathscr{R}I$. Define $\gamma_1(T)$ to be the scalar that appears in $(I-P)\rho(T)(I-P)$ and set $\psi_1(T) = \rho(T) - \gamma_1(T)I$. Since ρ is linear, and γ_1 is linear by construction, ψ_1 is linear as well. Also by the equality in Proposition 4.6.4 (b) we see that ψ_1 maps $P\mathcal{T}P$ onto itself.

We seek to show that ψ_1 is an associative algebra homomorphism. First let $T \in P\mathfrak{T}P$ and $X \in P\mathfrak{T}(I-P)$. From Proposition 4.6.4 (a) we have that $\rho(X) \in P\mathfrak{T}(I-P)$. Hence,

$$\begin{aligned}\rho(TX) &= \rho([T,X]) = [\rho(T), \rho(X)] \\ &= [\psi_1(T) + \gamma_1(T)I, \rho(X)] = [\psi_1(T), \rho(X)] = \psi_1(T)\rho(X).\end{aligned}$$

Thus if $T_1, T_2 \in P\mathfrak{T}P$, we see

$$\psi_1(T_1T_2)\rho(X) = \rho(T_1T_2X) = \psi_1(T_1)\rho(T_2X) = \psi_1(T_1)\psi_1(T_2)\rho(X).$$

Since $X \in P\mathfrak{T}(I-P)$ is arbitrary and, by Proposition 4.6.4 (a), ρ maps $P\mathfrak{T}(I-P)$ onto itself, we get

$$\psi_1(T_1T_2) = \psi_1(T_1)\psi_1(T_2). \qquad \square$$

Remark 4.6.6. By making the appropriate changes to Propositions 4.6.4 and 4.6.5 we can also show that for $T \in (I-P)\mathfrak{T}(I-P)$,

$$\rho(T) = \psi_2(T) + \gamma_2(T)I$$

where ψ_2 is an associative automorphism of $(I-P)\mathfrak{T}(I-P)$ and γ_2 is a linear functional annihilating all commutators in $(I-P)\mathfrak{T}(I-P)$.

– – – – – – –

By combining the results of 4.6.4, 4.6.5, and 4.6.6 we see that with respect to the decomposition $\mathscr{R}^n = \mathscr{R}^p \oplus \mathscr{R}^{n-p} = P\mathscr{R}^n \oplus (I-P)\mathscr{R}^n$, we have

$$\rho \left[\begin{array}{c|c} T_{11} & T_{12} \\ \hline 0 & T_{22}\end{array}\right] = \left[\begin{array}{c|c} \psi_1(T_{11}) & \xi(T_{12}) \\ \hline 0 & \psi_2(T_{22})\end{array}\right] + \tau\left(\left[\begin{array}{c|c} T_{11} & T_{12} \\ \hline 0 & T_{22}\end{array}\right]\right) I,$$

where $\tau : \mathfrak{T} \to \mathscr{R}$ is a linear functional that annihilates $P\mathfrak{T}(I-P)$, and ξ is linear. Acting on T, we have $\tau(T) = \gamma_1(T_{11}) + \gamma_2(T_{22})$, so τ annihilates all commutators since if T is a commutator, then T_{11} and T_{22} are commutators in the corresponding restricted algebras.

Since ψ_1 and ψ_2 are associative automorphisms of block-triangular algebras, they are inner by Theorem 2.4.4. Hence, there exists invertible $V \in \mathcal{T}(P\mathcal{N}) \subseteq M_p(\mathscr{R})$ and $W \in \mathcal{T}((I-P)\mathcal{N}) \subseteq M_{n-p}(\mathscr{R})$ such that

$$\begin{aligned} \psi_1(T_{11}) &= V^{-1}T_{11}V, &&\text{for} && T_{11} \in P\mathcal{T}P, \\ \psi_2(T_{22}) &= W^{-1}T_{22}W, &&\text{for} && T_{22} \in (I-P)\mathcal{T}(I-P). \end{aligned}$$

Definition 4.6.7. Let $U = V \oplus W$ and define $\rho_0 = (\mathrm{Ad}_U)^{-1} \circ \rho = \mathrm{Ad}_{U^{-1}} \circ \rho$. We then have $\rho_0(T) = \pi(T) + \tau(T)I$ where

$$\pi \left[\begin{array}{c|c} T_{11} & T_{12} \\ \hline 0 & T_{22} \end{array}\right] = \left[\begin{array}{c|c} T_{11} & \mu(T_{12}) \\ \hline 0 & T_{22} \end{array}\right]$$

for the linear map $\mu(T_{12}) = V\xi(T_{12})W^{-1}$.

Since $\pi = \rho_0 - I\tau$ is the sum of two Lie homomorphisms, π is also a Lie homomorphism. From Proposition 4.6.4 (a), we saw that $\rho(P\mathcal{T}(I-P)) = P\mathcal{T}(I-P)$ and so ξ is bijective, meaning that μ is also bijective. Hence, π is bijective.

Lemma 4.6.8. *Let $\mathcal{N}$ be a nest and $P \in \mathcal{P}(\mathcal{N})$, $P \neq 0, I$. With respect to the decomposition $\mathscr{R}^n = P\mathscr{R}^n \oplus (I-P)\mathscr{R}^n$, let*

$$\pi \left[\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right] = \left[\begin{array}{c|c} A & \mu(B) \\ \hline 0 & C \end{array}\right]$$

be a Lie isomorphism of $\mathcal{T}(\mathcal{N})$. Then

(a) μ is a left $\mathcal{T}(P\mathcal{N})$-module and a right $\mathcal{T}((I-P)\mathcal{N})$-module map;

(b) π is an associative algebra automorphism of $\mathcal{T}(\mathcal{N})$.

Proof. (a) Let $A \in \mathcal{T}(P\mathcal{N})$, $B \in M_{p\times(n-p)}(\mathscr{R})$. Then, since π is a Lie homomorphism,

$$\begin{aligned} \left[\begin{array}{c|c} 0 & \mu(AB) \\ \hline 0 & 0 \end{array}\right] = \pi \left[\begin{array}{c|c} 0 & AB \\ \hline 0 & 0 \end{array}\right] &= \pi \left[\left[\begin{array}{c|c} A & 0 \\ \hline 0 & 0 \end{array}\right], \left[\begin{array}{c|c} 0 & B \\ \hline 0 & 0 \end{array}\right] \right] \\ &= \left[\left[\begin{array}{c|c} A & 0 \\ \hline 0 & 0 \end{array}\right], \left[\begin{array}{c|c} 0 & \mu(B) \\ \hline 0 & 0 \end{array}\right] \right] = \left[\begin{array}{c|c} 0 & A\mu(B) \\ \hline 0 & 0 \end{array}\right]. \end{aligned}$$

Thus, $\mu(AB) = A\mu(B)$. Similarly $\mu(BD) = \mu(B)D$ for $D \in \mathfrak{T}((I-P)\mathcal{N})$.

(b) Using the bimodule property just established, then for $T, S \in \mathfrak{T}$ we have

$$\begin{aligned}
\pi(T)\pi(S) &= \left[\begin{array}{c:c} T_{11} & \mu(T_{12}) \\ \hdashline 0 & T_{22} \end{array}\right] \left[\begin{array}{c:c} S_{11} & \mu(S_{12}) \\ \hdashline 0 & S_{22} \end{array}\right] \\
&= \left[\begin{array}{c:c} T_{11}S_{11} & T_{11}\mu(S_{12}) + \mu(T_{12})S_{22} \\ \hdashline 0 & T_{22}S_{22} \end{array}\right] \\
&= \left[\begin{array}{c:c} T_{11}S_{11} & \mu(T_{11}S_{12} + T_{12}S_{22}) \\ \hdashline 0 & T_{22}S_{22} \end{array}\right] = \pi(TS),
\end{aligned}$$

and so π is associative algebra automorphism. □

Observation 4.6.9. We again invoke Theorem 2.4.4, to get

$$\pi(T) = \mathrm{Ad}_Z(T) = Z^{-1}TZ,$$

for some invertible $Z \in \mathfrak{T}(\mathcal{N})$.

By stepping back through each of these steps, we may now prove our main result.

Proof of Theorem 4.1.1. From Observation 4.6.9, we have

$$\pi(T) = Z^{-1}TZ,$$

while from Definition 4.6.7, we get

$$\rho(T) = U^{-1}\pi(T)U + \tau(T)I = U^{-1}Z^{-1}TZU + \tau(T)I.$$

In Remark 4.6.2, we constructed $\rho = \mathrm{Ad}_R \circ \varphi$ and so

$$\varphi(T) = (RU^{-1}Z^{-1})T(ZUR^{-1}) + \tau(T)I.$$

By defining $Y = ZUR^{-1}$ we then achieve the main result

$$\varphi(T) = Y^{-1}TY + \tau(T)I,$$

as desired. From Corollary 4.2.4 (a), we have $\varphi(I) = \lambda I$ for some $\lambda \in \mathscr{R}^{\times}$ so, in this notation, $\lambda = 1 + \tau(I)$ is invertible in $\mathscr{R}$.

In the case where φ was order reversing, the composition $\varphi \circ \omega$ was order preserving so

$$\begin{aligned}
\varphi \circ \omega(T) &= Y^{-1}TY + \tau(T)I, \text{ or} \\
\varphi(T) &= -Y^{-1}JT^{\top}JY + \tau(-JT^{\top}J)I, \\
\varphi(T) &= -Y^{-1}JT^{\top}JY + \hat{\tau}(T)I,
\end{aligned}$$

which matches the second case, for the generalized trace $\hat{\tau}(T) := \tau(-JT^{\top}J)$.

This completes the main theorem (4.1.1) for block upper-triangular algebras. □

Bibliography

[1] Y. A. Cao. Automorphisms of Certain Lie Algebras of Upper Triangular Matrices over a Commutative Ring. *Journal of Algebra*, 189(2):506–513, 1997.

[2] W. S. Cheung. *Mappings on Triangular Algebras*. PhD thesis, University of Victoria, 2000.

[3] T. W. Hungerford. *Algebra*, volume 73 of *Graduate Texts in Mathematics*. Springer, New York, 2003.

[4] I. M. Isaacs. Automorphisms of matrix algebras over commutative rings. *Linear Algebra and its Applications*, 31:215–231, 1980.

[5] T. P. Kezlan. A note on algebra automorphisms of triangular matrices over commutative rings. *Linear Algebra and its Applications*, 135:181–184, 1990.

[6] L. W. Marcoux and A. R. Sourour. Commutativity preserving linear maps and Lie automorphisms of triangular matrix algebras. *Linear Algebra and its Applications*, 288:89–104, 1999.

[7] L. W. Marcoux and A. R. Sourour. Lie Isomorphisms of Nest Algebras. *Journal of Functional Analysis*, 164(1):163–180, 1999.

[8] W. S. Martindale. Lie Isomorphisms of Prime Rings. *Trans. Amer. Math. Soc.*, 142:437 – 455, 1969.

[9] D. Ž. Đoković. Automorphisms of the Lie Algebra of Upper Triangular Matrices over a Connected Commutative Ring. *Journal of Algebra*, 170(1):101–110, 1994.