MATRIX GROUPS AND LIE ALGEBRAS:

AN ALGEBRAIC TREATMENT

by

TUCKER JEROME ERVIN

MARTIN EVANS, COMMITTEE CHAIR
MARTYN DIXON
JON CORSON
MATTHEW FEMINELLA

A THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Arts
in the Department of Mathematics
in the Graduate School of
The University of Alabama

TUSCALOOSA, ALABAMA

2019

ABSTRACT

There exist two correspondences between groups and Lie algebras. One occurs between matrix Lie groups and Lie algebras. The other concerns itself with complete groups of unitriangular matrices and Lie algebras. Titled the Lie and Mal'cev correspondences respectively, the purpose of this paper is to explore the two. We begin with an introduction to the basic properties of Lie algebras and other preliminary material followed by a construction of free Lie rings and algebras as well as by other interesting material discovered along the way. We then dive into the Lie correspondence, with which we contrast the Mal'cev correspondence in the section after.

DEDICATION

I would like to dedicate this thesis to my future wife. May this be the first of many papers she proofreads.

LIST OF ABBREVIATIONS AND SYMBOLS

$a^b = b^{-1}ab$	The conjugate of a by b for any two group elements a and b
$F_n(A, B)$	A homogeneous polynomial of weight n in variables A and B
[,]	The Lie bracket
A^-	The Lie ring (algebra) formed from the ring (algebra) A
$M_n(\mathbb{R})$	The matrix ring of square matrices over a ring R
$T_n(\mathbb{R})$	The group of unitriangular square matrices over a ring R
$U_n(\mathbb{R})$	The group of strictly upper triangular square matrices over a ring R

ACKNOWLEDGMENTS

I would like to first thank my advisor, Dr. Martin Evans, without whom none of this would have been possible.

For agreeing to serve on the committee, I thank Dr. Martyn Dixon, Dr. Jon Corson, and Dr. Matthew Feminella.

Finally, I would like to thank my mother for her help editing this paper.

CONTENTS

CHAPTER 1

INTRODUCTION AND OTHER PRELIMINARY MATERIAL

We begin with a discussion of Lie rings, Lie algebras, and a few of their properties. Then we will demonstrate how to form a Lie ring (algebra) given an arbitrary ring (algebra). Following soon after is an overview of nilpotence and commutators before we finish the section by examining the matrix groups that we are concerned with.

Definition 1. *A non-associative ring is an abelian group A with two binary products $+$ and $*$ that satisfies the distributive laws for all a, b, and c in A:*

1. $(a + b) * c = (a * c) + (b * c)$.

2. $a * (b + c) = (a * b) + (a * c)$.

In other words, a non-associative ring is a ring, with or without a multiplicative identity, that drops the requirement of associativity of $*$.

Note that non-associative does not mean strictly not associative. It is not *necessarily* associative.

We can then define non-associative algebras.

Definition 2. *A non-associative algebra A is a vector space over a field F that is a non-associative ring with respect to a product $*$ such that for all $a, b \in A$ and $\alpha \in F$*

$$(\alpha a) * b = \alpha(a * b) = a * (\alpha b).$$

Thus $*$ is a bilinear binary product. The non-associative algebra A can be referred to as a non-associative F-algebra.

Lie rings and algebras are special types of non-associative rings and algebras.

Definition 3. *A Lie ring is a non-associative ring L with a binary product $+$ and a non-associative binary product $*$ satisfying two properties:*

1. $x * x = 0$ *for all* $x \in L$. *(Alternation)*

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

Instead of the placeholder $*$, Lie rings and algebras often use $[\ ,\ ]$ to denote their non-associative binary product. The properties above then become:

1. $[x, x] = 0$ for all $x \in L$.

2. $[[x, y], z] + [[y, z], x] + [[z, x], y] = 0$ for all $x, y, z \in L$.

We will often refer to $[\ ,\ ]$ as the Lie bracket and $[x, y]$ as a Lie product.

Note that the first property of the Lie bracket implies a third. Let L be a Lie ring. If $[z, z] = 0$ for any $z \in L$, then for all $x, y \in L$

$$0 = [x + y, x + y] = [x + y, x] + [x + y, y]$$

$$0 = [x, x] + [y, x] + [x, y] + [y, y]$$

$$0 = [y, x] + [x, y].$$

Thus, for all $x, y \in L$,

$$[x, y] = -[y, x].$$

We say that the Lie bracket is anticommutative.

If we assume a given non-associative ring A possesses an anticommutative product $*$, then

$$x * x - x * x = x * x + x * x = 2x * x,$$

and we deduce that $x * x = 0$ for all $x \in A$ in the case that $(A, +)$ has no elements of order 2.

If we restrict our Lie rings L to ones without elements of order two, we could very well replace the condition $[x, x] = 0$ with $[x, y] = -[y, x]$ for all $x, y \in L$. Also, note that

$$0 = [[x, y], z] + [[y, z], x] + [[z, x], y]$$

$$0 = -[x, [y, z]] - [y, [z, x]] - [z, [x, y]]$$

$$0 = -([x, [y, z]] + [y, [z, x]] + [z, [x, y]]),$$

and if $0 = [x, [y, z]] + [y, [z, x]] + [z, [x, y]]$, then

$$0 = -[[x, y], z] - [[y, z], x] - [[z, x], y]$$

$$0 = -([[x, y], z] + [[y, z], x] + [[z, x], y]).$$

Thus the Jacobi identity could be restated as $[x, [y, z]]+[y, [z, x]]+[z, [x, y]] = 0$. However, following the notation in Khukhro [2], we will define our simple Lie products as $[x_1, x_2, x_3] = [[x_1, x_2], x_3]$ and $[x_1, \ldots, x_n] = [[x_1, \ldots, x_{n-1}], x_n]$ for all $n > 3$; we say that our products are *left-normed*. In addition, if we have a simple Lie product $[x_1, x_2, \ldots, x_n]$, we say it is of *weight* n. We define the weights of an arbitrary Lie product as the sum of the weights of its subproducts. Thus, our version of the Jacobi identity is left-normed and is less cumbersome when written as $[x, y, z] + [y, z, x] + [z, x, y] = 0$.

Definition 4. *A Lie algebra, specifically an F-algebra, is a Lie ring that is also a non-associative algebra over the field F. In other words, it satisfies (1) and (2) as in the definition*

of Lie rings.

For one of the most important examples of Lie algebras, take $\mathbb{R}^3$ as a vector space with the well known cross product $\times$. If i, j, and k are the basis vectors of $\mathbb{R}^3$ and $v = \langle v_1, v_2, v_3 \rangle$ and $w = \langle w_1, w_2, w_3 \rangle$ are arbitrary vectors, the cross product is defined as

$$v \times w = (v_2w_3 - w_2v_3)i - (v_1w_3 - w_1v_3)j + (v_1w_2 - w_1v_2)k.$$

Switching the w_i with the v_i produces $-(v_2w_3 - w_2v_3)i + (v_1w_3 - w_1v_3)j - (v_1w_2 - w_1v_2)k$. Thus $v \times w = -w \times v$. Now, let $z = \langle z_1, z_2, z_3 \rangle$ be another arbitrary vector. Then

$$(v \times w) \times z + (w \times z) \times v + (z \times v) \times w$$

$$= \langle v_2w_3 - w_2v_3, -v_1w_3 + w_1v_3, v_1w_2 - w_1v_2 \rangle \times z +$$

$$\langle w_2z_3 - z_2w_3, -w_1z_3 + z_1w_3, w_1z_2 - z_1w_2 \rangle \times v + \langle z_2v_3 - v_2z_3, -z_1v_3 + v_1z_3, z_1v_2 - v_1z_2 \rangle \times w.$$

Looking solely at the first coordinate,

$$(-v_1w_3 + w_1v_3)z_3 - z_2(v_1w_2 - w_1v_2) + (-w_1z_3 + z_1w_3)v_3 - v_2(w_1z_2 - z_1w_2) +$$

$$(-z_1v_3 + v_1z_3)w_3 - w_2(z_1v_2 - v_1z_2) = 0.$$

Similarly, the other coordinates also simplify to 0. Thus $(v \times w) \times z + (w \times z) \times v + (z \times v) \times w$ becomes the 0 vector for all vectors in $\mathbb{R}^3$. The product $\times$ is therefore anticommutative and satisfies the Jacobi identity, transforming $\mathbb{R}^3$ into a Lie algebra. Given the ubiquitous use of the cross product in physics, $\mathbb{R}^3$ becomes the go-to example of a Lie algebra.

To explore Lie algebras in general, we have one major theorem concerning the link between the more common associative rings (algebras) and Lie rings (algebras).

Theorem 1. *Every associative ring $(R, +, *)$ can be viewed as a Lie ring R^- in the following*

way. We define a Lie product $[\ ,\ ]$ *on the set* R *by* $[a,b] = a*b - b*a$ *for all* $a, b \in R$*. Then* $(R, +, [\ ,\ ])$ *is a Lie ring, denoted by* R^-*.*

Proof. It is necessary to show that $[\ ,\ ]$ satisfies the three properties of a Lie bracket: the distributive laws, alternation, and the Jacobi identity. The abelian group $(R, +)$ remains unchanged.

Let $a, b, c \in R$. Then

$$[(a+b), c] = (a+b)*c - c*(a+b) = a*c + b*c - c*a - c*b = [a,c] + [b,c]$$

and

$$[a, (b+c)] = a*(b+c) - (b+c)*a = a*b + a*c - b*a - c*a = [a,b] + [a,c].$$

Since $[\ ,\ ]$ satisfies the two distributive laws, $(R, +, [\ ,\ ])$ is a non-associative ring. Additionally,

$$[a,a] = a*a - a*a = 0,$$

satisfying the first property of a Lie bracket.

The Jacobi identity is also satisfied as $*$ is an associative product and

$$[[a,b],c] + [[b,c],a] + [[c,a],b]$$

$$= [a,b]*c - c*[a,b] + [b,c]*a - a*[b,c] + [c,a]*b - b*[c,a]$$

$$= ([a,b]*c - a*[b,c]) + ([c,a]*b - c*[a,b]) + ([b,c]*a - b*[c,a])$$

$$= a*(c*b) - (b*a)*c + c*(b*a) - (a*c)*b + b*(a*c) - (c*b)*a = 0.$$

Therefore $(R, +, [\ ,\ ])$, which we denote by R^-, is a Lie Ring.

□

Obviously, this implies the following corollary.

Corollary 1. *If A is an associative algebra, then A^- as defined in the proof of Theorem 1 is a Lie algebra.*

The main concern of this paper will be a discussion of the different kinds of correspondences that arise between special kinds of groups and Lie algebras. Before delving further into this discussion, we first remind the reader of definitions of nilpotence for groups and non-associative rings, as both will be needed later on.

To begin, we define the commutators of a group.

Definition 5. *Let G be a group with $a, b \in G$. Then the commutator of a and b is $[a, b] = a^{-1}b^{-1}ab$. As before, we let $[x_1, x_2, x_3, \ldots, x_n] = [[[[x_1, x_2], x_3], \ldots], x_n]$.*

Additionally, we define $[M, N] = \langle [m, n] \ : m \in M,\ n \in N \rangle$ for any two subsets M and N of G.

We have a few useful formulas that deal with commutators [2]. Here and throughout we will write $a^b = b^{-1}ab$ to denote conjugation of a by b for any group elements a and b.

Lemma 1. ***Commutator Identities***

Let a, b, and c be any elements of any group G and ϕ a homomorphism from G to any group. Then the following identities hold:

1. $a^b = a[a, b]$,
2. $[ab, c] = [a, c]^b[b, c] = [a, c][a, c, b][b, c]$,
3. $[a, bc] = [a, c][a, b]^c = [a, c][a, b][a, b, c]$,
4. $[a, b]^{-1} = [b, a]$,
5. $\phi([a, b]) = [\phi(a), \phi(b)]$,
6. $[a, b^{-1}, c]^b[b, c^{-1}, a]^c[c, a^{-1}, b]^a = 1$ *(The Hall-Witt Identity).*

Proof. The proofs of these identities are relatively elementary and consist of direct calculations. They are given below.

1.
$$a^b = b^{-1}ab = aa^{-1}b^{-1}ab = a[a,b].$$

2.
$$[ab,c] = b^{-1}a^{-1}c^{-1}abc = b^{-1}a^{-1}c^{-1}acbb^{-1}c^{-1}bc = [a,c]^b[b,c]$$
$$= [a,c][a,c]^{-1}[a,c]^b[b,c] = [a,c][a,c,b][b,c].$$

3.
$$[a,bc] = a^{-1}c^{-1}b^{-1}abc = a^{-1}c^{-1}acc^{-1}a^{-1}b^{-1}abc = [a,c][a,b]^c$$
$$= [a,c][a,b][a,b]^{-1}[a,b]^c = [a,c][a,b][a,b,c].$$

4.
$$[a,b]^{-1} = (a^{-1}b^{-1}ab)^{-1} = b^{-1}a^{-1}ba = [b,a].$$

5.
$$\phi([a,b]) = \phi(a^{-1}b^{-1}ab) = \phi(a)^{-1}\phi(b)^{-1}\phi(a)\phi(b) = [\phi(a),\phi(b)].$$

6.
$$[a,b^{-1},c]^b[b,c^{-1},a]^c[c,a^{-1},b]^a$$
$$= b^{-1}[a,b^{-1}]^{-1}c^{-1}[a,b^{-1}]cbc^{-1}[b,c^{-1}]^{-1}a^{-1}[b,c^{-1}]aca^{-1}[c,a^{-1}]^{-1}b^{-1}[c,a^{-1}]ba$$
$$= b^{-1}[b^{-1},a]c^{-1}[a,b^{-1}]cbc^{-1}[c^{-1},b]a^{-1}[b,c^{-1}]aca^{-1}[a^{-1},c]b^{-1}[c,a^{-1}]ba$$
$$= b^{-1}ba^{-1}b^{-1}ac^{-1}a^{-1}bab^{-1}cbc^{-1}cb^{-1}c^{-1}ba^{-1}b^{-1}cbc^{-1}aca^{-1}ac^{-1}a^{-1}cb^{-1}c^{-1}aca^{-1}ba$$
$$= 1.$$

□

We next define the lower central series of a group.

Definition 6. *The lower central series of a group G is the series*

$$\gamma_1(G) \geq \gamma_2(G) \geq \cdots \geq \gamma_k(G) \geq \ldots$$

defined recursively by $\gamma_1(G) = G$ and $\gamma_n(G) = [\gamma_{n-1}(G), \gamma_1(G)]$ for all $n \geq 2$.

We may now define nilpotent groups.

Definition 7. *A group G is said to be nilpotent if $\gamma_n(G) = 1$ for some $n \in \mathbb{N}$. The group is said to be nilpotent of class c if $\gamma_{c+1}(G) = 1$ and $\gamma_c(G) \neq 1$.*

In comparison, the definition for a nilpotent non-associative ring is more intuitive.

Definition 8. *An element a in a non-associative ring R (with non-associative product $*$) is said to be nilpotent if $a^n = 0$ for some $n \in \mathbb{N}$, where $a^1 = a$ and $a^n = a^{n-1} * a$ for all $n \geq 2$.*

Similarly, R is said to be nilpotent if $R^n = 0$ for some $n \in \mathbb{N}$, where $R^1 = R$ and $R^n = R^{n-1}R = \{\sum_{i=1}^{k} a_i * b_i \ : a_i \in R^{n-1},\ b_i \in R,\ k \in \mathbb{N}\}$ for all $n \geq 2$. We say that R is nilpotent of class c if $R^{c+1} = 0$ and $R^c \neq 0$.

There is an important fact concerning the Lie ring (algebra) formed from an associative ring (algebra) A.

Lemma 2. *If an associative ring (algebra) A is nilpotent, then A^- is also nilpotent.*

Proof. To demonstrate this, we will proceed by induction to show that the simple Lie products of weight n in A^- are a linear combination of products of weight n in A. For $n = 2$, $[x_1, x_2] = x_1 * x_2 - x_2 * x_1$, which is a linear combination of products of weight 2 in A, completing the base case.

Suppose $[x_1, \ldots, x_n]$ is a linear combination of products of weight n in A. Then

$$[x_1, \ldots, x_n, x_{n+1}] = [x_1, \ldots, x_n] * x_{n+1} - x_{n+1} * [x_1, \ldots, x_n].$$

Following from our assumption, $[x_1, \ldots, x_n, x_{n+1}]$ is a linear combination of products of weight $n+1$ in A. Thus $[x_1, \ldots, x_n]$ is a linear combination of products of weight n in A for all $n \geq 2$.

Therefore, if A is nilpotent of class c, $[x_1, x_2, \ldots, x_{c+1}] = 0$ for all $x_i \in A$, forcing A^- to be nilpotent of class at most c.

□

Now we define the two main classes of groups we are concerned with in this work: matrix Lie groups and matrix Mal'cev groups.

Definition 9. *A matrix Lie group G is a subgroup of the group of n-by-n invertible real matrices with the additional restriction that G is closed under nonsingular limits.*

To demonstrate what is meant by a nonsingular limit of matrices, we must first define the absolute value of a real matrix. Our treatment of matrix Lie groups is based on the material of Stillwell [7] and Hall [1].

Definition 10. *If $A = (a_{ij})$ is any real-valued matrix, the absolute value of that matrix is defined as*

$$|A| = \sqrt{\sum_{i,j} |a_{ij}|^2}.$$

Note, if A is a square n-by-n matrix, the absolute value of A is the euclidean norm of A when $M_n(\mathbb{R})$ is viewed as the space $\mathbb{R}^{n^2}$. As such, it possesses the usual properties of a norm.

1. $|A| \geq 0$.
2. $|A| = 0$ if and only if $A = 0$.
3. $|\alpha A| = |\alpha||A|$ for all real numbers α.
4. $|A + B| \leq |A| + |B|$ for all square matrices A and B.

In addition to the properties afforded to a norm, the absolute value of an n-by-n matrix is submultiplicative, i.e. $|AB| \leq |A||B|$ for all real n-by-n matrices A and B, following directly from the triangle and Cauchy-Schwarz inequalities of the real numbers.

By analogy with the definition of the limit of a sequence of real numbers, we say that a sequence $\{A_k\}_{k=1}^{\infty}$ of real n-by-n matrices has the real n-by-n matrix A as a limit if for every $\epsilon > 0$, there is some positive integer K such that $|A - A_k| < \epsilon$ for every $k > K$. If this limit exists and is nonsingular (i.e. invertible), then the sequence $\{A_k\}_{k=1}^{\infty}$ is said to have a nonsingular limit A.

Given our new definitions, we can show that if $\{A_k\}_{k=1}^{\infty}$ is a sequence of square matrices such that the series $\sum_{k=1}^{\infty} |A_k|$ converges, then $\sum_{k=1}^{\infty} A_k$ also converges. As proof, let $T_j = \sum_{i=1}^{j} |A_i|$ and $S_j = \sum_{i=1}^{j} A_i$. Then the $\sum_{k=1}^{\infty} |A_k|$ and $\sum_{k=1}^{\infty} A_k$ are the limits of T_j and S_j respectively as j increases. Fix $\epsilon > 0$. Since $\{T_j\}_{j=1}^{\infty}$ converges in the real numbers, it forms a Cauchy sequence and $|T_n - T_m| < \epsilon$ for all $n > m$ greater than some positive integer K. It follows that

$$|S_n - S_m| = |A_n + A_{n-1} + \cdots + A_{m+1}|$$

$$\leq |A_n| + |A_{n-1}| + \cdots + |A_{m+1}|$$

$$= T_n - T_m < \epsilon,$$

providing the desired result, which will be most useful in the section concerning the Lie correspondence.

Now we can understand the definition of a matrix Lie group. If a group of n-by-n invertible real matrices G is a matrix Lie group, then every nonsingular limit A of a sequence $\{A_k\}_{k=1}^{\infty}$, with each $A_k \in G$, is contained in G.

For the Mal'cev correspondence, we focus on matrix Mal'cev groups. Ian Stewart's treatment is our main source for this correspondence, and he begins with unitriangular matrix groups over the rationals that possess nth roots for every $n \in \mathbb{Z}$ [6]. He does not define any specific groups as matrix Mal'cev groups, but we will refer to groups of this type

as matrix Mal'cev groups. First, we must define what it means for a group to be complete and what an R-group is.

Definition 11. *A group G is complete if for every non-zero integer n and every $g \in G$ there exists $h \in G$ with $h^n = g$.*

Such groups are also called *radicable* or *divisible* groups.

Definition 12. *A matrix Mal'cev group is a complete group of unitriangular matrices over the rationals.*

It is important to observe that in nilpotent torsion-free groups, extraction of roots, when possible, is unique.

Lemma 3. *If G is a nilpotent torsion-free group and $a^n = b^n$ for some $a, b \in G$ and $n \in \mathbb{Z}$ then $a = b$.*

Proof. We argue by induction on the class c of nilpotence of G. If $c = 1$, then G is abelian and $(ab^{-1})^n = 1$. Since G is torsion-free, ab^{-1} does not have finite order. Thus $ab^{-1} = 1$ implies that $a = b$, proving the base case.

For the inductive step, assume the hypothesis holds for groups of class $c - 1$. If G is a nilpotent torsion-free group of class c, then $G/Z(G)$ is a nilpotent group of class $c - 1$. Additionally, this quotient group is torsion-free [4, Theorem 5.2.19]. Thus if $(aZ(G))^n = (bZ(G))^n$, $aZ(G) = bZ(G)$ by our inductive hypothesis. This implies that $a = bz$ for some $z \in Z(G)$. Then $a^n = b^n z^n = a^n z^n$, further implying that $z^n = 1$ and $z = 1$ as before. Therefore, $a = b$ if $a^n = b^n$ for all $n \in \mathbb{Z}$ and all groups G of class $c > 0$. □

We say that a group G is an R-group if extraction of roots, when possible, is unique. Thus, every nilpotent torsion-free group is an R-group. For complete nilpotent torsion-free groups G, this result has an important consequence. We may define g^q for every $g \in G$ and $q = a/b \in \mathbb{Q}$ (where $a, b \in \mathbb{Z}$) as the unique $h \in G$ such that $h^b = g^a$. This allows us to take rational roots in G. In the terminology of Khukhro [2, 1.40], complete nilpotent torsion-free groups are $\mathbb{Q}$-powered.

CHAPTER 2

FREE LIE RINGS AND OTHER IMPORTANT THEOREMS

We devote this section to some interesting results that are related to the Lie and Mal'cev correspondences. We begin with the construction of free Lie rings and then end with a short discussion of the Poincaré-Birkhoff-Witt theorem, Ado's theorem, and their importance for the theory of Lie algebras.

We recreate the construction of the free Lie rings and free Lie $\mathbb{Q}$-algebras out of an associative free $\mathbb{Q}$-algebra A following the treatment of Khukhro [2, Chapter 5.3]. Let $\{x_n\}_{n\in B}$ be a set of generators of A, where B is a well-ordered index set. As A is the algebra formed from $\mathbb{Q}$-linear combinations of monomials, we can alternately view A as a non-commutative polynomial algebra over $\mathbb{Q}$.

For ease of use, we are using natural numbers as indices, but the construction works for well-ordered sets of arbitrary cardinalities. Form the monomials of A. The monomials of degree 1 are all the x_n. The monomials of degree 2 are $x_n x_m$, and so on and so on. The set of all monomials then forms a basis for A as a vector space, and they are often referred to as words, in which case the x_n forming the word are referred to as letters. We can also refer to $\{x_i\}_{i\in B}$ as an alphabet.

Now, let L be the Lie subring generated by the set $\{x_1, x_2, \dots\}$ of A^-. We then show that the set $\mathbb{Q}L = \{ql : q \in \mathbb{Q},\ l \in L\}$ becomes a Lie $\mathbb{Q}$-algebra.

Lemma 4. *$\mathbb{Q}L = \{ql : q \in \mathbb{Q},\ l \in L\}$ is a Lie $\mathbb{Q}$-algebra.*

Proof. Let $ql, pk \in \mathbb{Q}L$, where $p, q \in \mathbb{Q}$ and $k, l \in L$. Then $[ql, pk] = qlpk - pkql$. As A is a

$\mathbb{Q}$-algebra,

$$qlpk - pkql = qplk - pqkl = pq(lk - kl) = pq[l, k].$$

Now, let t be any rational number such that tq and tp are integers. Then $ql + pk = t^{-1}(tql + tpk)$. Since L is the Lie ring generated by the generators of A, both tql and tqk are contained in L. Thus $ql + pk \in \mathbb{Q}L$.

Since this holds for any $p, q \in \mathbb{Q}$ and $k, l \in L$, $\mathbb{Q}L$ is a non-associative subring of A^-. Now, if $[\ ,\ ]$ is bilinear with respect to $\mathbb{Q}$, $\mathbb{Q}L$ will be a Lie $\mathbb{Q}$-algebra.

Let $t \in \mathbb{Q}$. Then

$$[tql, pk] = tqlpk - pktql = t(qlpk - pkql) = t[ql, pk],$$

and

$$[ql, tpk] = qltpk - tpkql = t(qlpk - pkql) = t[ql, pk].$$

Therefore, $\mathbb{Q}L$ is a subalgebra of A^-, making it a Lie $\mathbb{Q}$-algebra.

□

Note that $L \subset \mathbb{Q}L \neq A^-$ as the monomial $x_n x_m$ is not contained in $\mathbb{Q}L$. If it were, then $x_n x_m$ would be a linear combination of Lie products in $\mathbb{Q}L$. However, each Lie product in $\mathbb{Q}L$ is a rational multiple of a linear combination of monomials in A. Thus $x_n x_m \in \mathbb{Q}L$ contradicts the linear independence of the basis elements of A.

Our goal is to show that the Lie ring and Lie $\mathbb{Q}$-algebra we have just constructed are free. To this end we need to define basic Lie products and an order on the monomials of A.

Definition 13. *We define an order $<$ on the set of monomials of A in the following way:*

1. *The monomials of degree 1 are ordered according to the order of B on their indices, i.e. $x_1 < x_2 < x_3 < \ldots$,*

2. *If we have two words $x_{i_1} \ldots x_{i_n}$ and $x_{j_1} \ldots x_{j_m}$ with n and m greater than 1, then*

$x_{i_1} \dots x_{i_n} < x_{j_1} \dots x_{j_m}$ *implies either that there is some* $k \in \mathbb{N}$ *such that* $x_{i_k} < x_{j_k}$ *and* $x_{i_s} = x_{j_s}$ *for all* s *less than* k *or that* $x_{j_1} \dots x_{j_m}$ *is a proper initial segment of* $x_{i_1} \dots x_{i_n}$.

Note that this is not the usual lexicographic ordering in which $u < uv$ whenever v has length at least 1 for any two words u and v.

Using this order, we may define what it means for a word to be regular.

Definition 14. *A word* u *is regular if* u *is greater than each of its cyclic permutations: if* $u = vw$ *is any non-trivial decomposition, then* $u > wv$.

Let u denote any word in A. In general, there exists a myriad of ways in which we could bracket u to obtain a Lie product. For example, if $u = x_1x_2x_1x_2x_1x_1$, we could form the Lie products

$$[[x_1, x_2], x_1, [x_2, x_1], x_1],$$

$$[[x_1, x_2], x_1, x_2, [x_1, x_1]],$$

or

$$[x_1, x_2, x_1, [x_2, x_1], x_1].$$

For any given word u, we could form possibly hundred of Lie products. Let us write $[u]$ to denote an arbitrary but fixed way of bracketing u as to form a Lie product. However, finding the underlying word u from a Lie product $[u]$ is as simple as removing the brackets. For example, $[x_1, x_2, x_1, [x_2, x_1], x_1]$ has underlying word $x_1x_2x_1x_2x_1x_1$.

Note further that our left-normed style of bracketing means that

$$[[x_1, x_2], x_1, [x_2, x_1], x_1]$$

represents

$$[[[[x_1, x_2], x_1], [x_2, x_1]], x_1].$$

When written formally in this way, we see that each occurrence of the symbols $x_1, x_2, \dots$

appears either to the left of a right bracket, $x_i]$, or the right of a left bracket, $[x_i$. This observation is vitally important for some of our upcoming lemmas.

Definition 15. *The basic Lie products of weight 1 are the symbols $[x_i]$, where $i \in B$. We define the basic Lie products of arbitrary weight inductively. We say that $[[b_1], [b_2]]$ is a basic Lie product of weight the sum of the weights of $[b_1]$ and $[b_2]$ if*

1. *both $[b_1]$ and $[b_2]$ are basic Lie products,*

2. *$b_1 > b_2$ when considered with the order defined previously,*

3. *if the weight of $[b_1]$ is greater than 1 and $[b_1] = [[b_{1,1}], [b_{1,2}]]$, then $b_{1,2} \leq b_2$.*

As Khukhro is our main reference, we use the basic products of A. I. Shirshov [2, Page 66], also referred to as Lyndon words. In other treatments, the basic commutators of P. Hall and M. Hall are used.

Some quick examples of basic Lie products are x_k, $[x_j, x_k]$, $[x_j, [x_j, x_k]]$, and $[[x_i, x_k], [x_j, x_k]]$ for any $k < j < i \in B$. The Lie product $[x_1, x_2]$ would be the simplest non-example. Note that we adopt the convention that basic Lie products of weight 1 do not need to keep the bracket notation. Thus, we may write $[x_j, x_k]$ instead of $[[x_j], [x_k]]$.

What this given order and these basic Lie products will allow us to do is show that any Lie ring is generated by the images of basic Lie products from L. We will then be able to show that the basic Lie products in L are linearly independent. The proof that L is a free Lie ring will follow naturally from that, and $\mathbb{Q}L$ will be a free Lie $\mathbb{Q}$-algebra as a corollary. To go about doing so, we must first prove a few lemmas which all follow the method of Khukhro [2].

First, we provide an example to illustrate the concepts we will use. Suppose that A is the associative free $\mathbb{Q}$-algebra generated by three elements, x_1, x_2, and x_3 — which can also be viewed as the non-commutative polynomial ring in three indeterminates over $\mathbb{Q}$. There are precisely three basic Lie products of weight 1: $[x_1]$, $[x_2]$, and $[x_3]$. Since $x_1 < x_2 < x_3$,

the basic Lie products of weight two are $[x_3, x_2]$, $[x_3, x_1]$, and $[x_2, x_1]$. As

$$x_1 < x_2x_1 < x_2 < x_3x_1 < x_3x_2 < x_3,$$

we find the basic Lie products of weight three to be

$$[x_3, [x_3, x_2]], [x_3, [x_3, x_1]], [x_3, [x_2, x_1]]$$

$$[[x_3, x_2], x_2]$$

$$[[x_3, x_1], x_2], [[x_3, x_1], x_1]$$

$$[x_2, [x_2, x_1]]$$

$$[[x_2, x_1], x_1].$$

The underlying words for these basic Lie products are as follows: $x_3x_3x_2$, $x_3x_3x_1$, $x_3x_2x_1$, $x_3x_2x_2$, $x_3x_1x_2$, $x_3x_1x_1$, $x_2x_2x_1$, $x_2x_1x_1$. Each of these words is a regular word, and our next two lemmas prove that this holds in general.

Lemma 5. *If $[v]$ is a basic Lie product and x_m is the least letter in v, then x_m is not the first letter in v.*

Proof. Let $[v]$ be a basic Lie product of weight 2. Each $[v]$ is of the form $[x_i, x_j]$ for some $i, j \in B$, where $x_i > x_j$ by our definition of a basic Lie product. Thus, the least letter is not the first one in the word v, completing the case for a weight of 2. Assume now that basic Lie products of weight less than k do not have their least letter as their first letter. Let $[v]$ be a basic Lie product of weight k.

If v's first letter is its least, x_m say, then $[v] = [x_m, [v_1]]$ or $[v] = [[v_2], [v_3]]$ by the definition of the basic Lie products, where $[v_1], [v_2]$, and $[v_3]$ are basic Lie products of weight less than k with v_2's least letter being x_m as well. By the inductive hypothesis, $[v_2]$ cannot have x_m as its first letter. Thus $[v] = [x_m, [v_1]]$. However, if $[b]$ is a basic Lie product such that

$[b] = [[b_1], [b_2]]$, where $[b_1]$ and $[b_2]$ are also basic Lie products, then b_1 must be greater than b_2. Therefore, x_m, the least letter in the word v, cannot be the first letter of v.

By induction, if $[v]$ is a basic Lie product and x_m is the least letter in v, then x_m is not the first letter in v.

□

The above lemma is a crucial step in proving the next result.

Lemma 6. *If $[u]$ is a basic Lie product, the underlying associative word u is regular.*

Proof. By definition, the basic Lie products of weight 1 are regular, completing the base case. We then argue by induction on the weight.

Suppose the conclusion holds for basic Lie products of weight less than n. Let $[u]$ be a basic Lie product of weight n. Let x_m be the least letter in the word u. By the previous lemma, x_m is not the first letter of u. Then the first occurrence of x_m in u is immediately after some x_k. Recall our observation that each symbol x_j in $[u]$ appears to the right of a left bracket, $[x_j$, or to the left of a right bracket, $x_j]$. By lemma 5, no occurrence of x_m is of the form $[x_m$. Consider a subproduct $[a, x_m]$ of $[u]$. If the weight of $[a]$ is greater than 1, then $[a] = [[b], [c]]$ where $[b]$ and $[c]$ are basic Lie products of weight less than $[a]$. By the definition of basic Lie products, we have that $c \leq x_m$, implying that $c = x_m$ as x_m is the least letter in u. It follows that each occurrence of $x_k x_m$ in u comes from a subproduct $[x_k, x_m]$. In other words, the bracket arrangements $x_k]x_m]$, $x_k][x_m$, and $[x_k[x_m$ do not appear in $[u]$.

We introduce a new symbol, $[x_k, x_m]$, into the original alphabet $\{x_i\}_{i \in B}$ and adjust the ordering as $x_1 < x_2 < \cdots < x_{k-1} < [x_k, x_m] < x_k < \ldots$. We order the words in this new alphabet by using the same definition as before (but with the x_{i_r} running through the new alphabet). It is not difficult to see that the order on those words in the new alphabet that have no subwords $x_k x_m$ coincides with the order in the old sense after removing the brackets. Hence, $[u]$ is a basic Lie product of weight less than n in the new alphabet, if we regard all occurrences of $[x_k, x_m]$ in $[u]$ as occurrences of our new letter.

By our induction hypothesis, u is a regular word. By removing the brackets from the new letter, we see that u is greater than any cyclic permutation of u that does not break a subword $x_k x_m$. The cyclic permutations of u that do break some $x_j x_m$ begin with x_m and are therefore less than u as x_m is the least letter of u.

Therefore, if $[u]$ is a basic Lie product, the underlying associative word u is regular.

□

We now show that each regular word has a unique basic Lie product associated with it.

Lemma 7. *Let u be a regular word. Then there is a unique bracketing $[u]$ of u such that $[u]$ is a basic Lie product.*

Proof. We again argue by induction on the degree of u. If the degree is 1, then $[u]$ is the unique basic Lie product for u.

Suppose each regular word of degree less than n has a unique basic Lie product. Let u be a regular word of degree n, and let x_m be the least letter in u. As u is regular, x_m cannot be the first letter of u. Otherwise, $u = x_m v$ would imply that $u < v x_m$, a contradiction.

The first occurrence of x_m in u is then immediately after some $x_k > x_m$. As in the proof of Lemma 6, we replace each subword $x_k x_m$ in u by the symbol $[x_k, x_m]$ and consider the result as a word in the new alphabet containing $[x_k, x_m]$. We order the new alphabet and the words in the new alphabet as in Lemma 6. Then, of course, u is a regular word in this new alphabet. By the induction hypothesis, there is a unique basic product $[u]$ in the new alphabet. When viewed as Lie product $[u]$ in the old alphabet, it is also a basic Lie product. The uniqueness follows from the induction hypothesis and the fact that each subword $x_k x_m$ in u comes from a subproduct $[x_k, x_m]$, as we saw in the proof of Lemma 6.

□

We now prove a couple of properties of regular words and greatest words which will help demonstrate the linear independence of the basic Lie products of L [2].

Lemma 8. *If a is a regular word and $b > a$ for some word b, then $ba > ab$.*

Proof. Either $a = bc$, where c is some other word, or there is some $a_k < b_k$ with $a_j = b_j$, for all $j < k \in B$, and where $a = a_1 a_2 \dots a_k \dots a_n$ and $b = b_1 b_2 \dots b_k \dots b_m$. In the second case, ba will be greater than ab as $b_k > a_k$. In the first case, if $ab \geq ba$, then $bcb \geq bbc$. Thus $cb \geq bc$, contradicting a being a regular word. Therefore, $ba > ab$.

□

Recall that each basic Lie product $[u]$ where u is a word in the symbols $x_1, x_2, \dots$ can be interpreted as a linear combination in A of the monomials of our generating set $\{x_i\}_{i \in B}$ of A. For example, the Lie product $[x_2, [x_2, x_1]]$ is a linear combination of such monomials in the following way:

$$[x_2, [x_2, x_1]] = x_2[x_2, x_1] - [x_2, x_1]x_2$$

$$= x_2 x_2 x_1 - x_2 x_1 x_2 - x_2 x_1 x_2 + x_1 x_2 x_2.$$

We use this thought process and the preceding lemma to talk about the properties of the underlying word of a basic Lie product.

Lemma 9. *If a basic Lie product $[u]$ is expressed as a linear combination of associative monomials in the generators of A, then the underlying word u is the unique greatest word among the monomials.*

Proof. The result follows for basic Lie products of weight 1. Assume that the hypothesis holds for all basic Lie products of weight less than n. Let $[u]$ be a basic Lie product of weight n. By the definition of basic Lie products, $[u] = [[v], [w]]$, where $[v]$ and $[w]$ are basic Lie products of weight less than n such that $v > w$. By the induction hypothesis,

$$[v] = v + \sum_i \alpha_i v_i,$$

and

$$[w] = w + \sum_j \beta_j w_j,$$

where each sum is finite, the i and j are contained in B, $v > v_i$, $w > w_j$, and $\alpha_i, \beta_j \in \mathbb{Z}$. Computing $[[v], [w]]$ then gives

$$[u] = (vw - wv) + \sum_j \beta_j(vw_j - w_j v) + \sum_i \alpha_i(v_i w - wv_i) + \sum_{i,j} \alpha_i\beta_j(v_i w_j - w_j v_i).$$

By Lemma 6, w is regular. By Lemma 8, $vw > wv$. Additionally,

$$v_i w_j < v_i w < vw$$

and

$$w_j v_i < wv_i < wv < vw.$$

Thus $[u]$ is a linear combination of u and some linear combination of smaller monomials, completing the inductive proof. Therefore, if a basic product $[u]$ is expressed as a linear combination of associative monomials in the generators of A, then the underlying word u is the unique greatest word among the monomials.

□

Now, we investigate an arbitrary Lie ring M generated by a set $\{g_i\}_{i \in B}$. Throughout the remainder of this section, we order words in the g_i as above, treating the g_i as formal symbols, so that we can exploit the notions of "basic Lie product" and "regular word".

Lemma 10. *The basic Lie products span M.*

Proof. Again, we prove the hypothesis by induction on both the weight and order of commutators and their underlying associative words. If $[k]$ is any commutator of weight 1 in M, then it is naturally a linear combination of basic Lie products, specifically $[k]$ itself. Assume now that all commutators of weight less than n are sums of basic Lie products. This is our first inductive hypothesis

Let $[k]$ be a commutator of weight n. We know that $[x_1^n] = 0$, making it a sum of basic

Lie products. Assume that $[u]$ is a linear combination of basic Lie products if $u < k$ and $[u]$ has weight n. This is our second inductive hypothesis.

Then $[k] = [[k_1], [k_2]]$, where the weights of $[k_1]$ and $[k_2]$ sum to n. By our first inductive hypothesis, we may assume that $[k_1]$ and $[k_2]$ are linear combinations of basic Lie products. By the bilinear nature of the Lie product, we may assume without loss of generality that $[k_1]$ and $[k_2]$ are themselves basic Lie products. Since $[[k_1], [k_2]] = -[[k_2], [k_1]]$, we may also assume that $k_1 > k_2$, as the only difference is a scalar multiple.

If $[k_1]$ has weight 1, then $[k]$ is a basic Lie product by the definition of a basic Lie product. If it has weight greater than 1, then $[k_1] = [[k_{1,1}], [k_{1,2}]]$, where $[k_{1,1}]$ and $[k_{1,2}]$ are basic Lie products with $k_{1,1} > k_{1,2}$. If $k_{1,2} \leq k_2$, then $[k]$ is a basic Lie product. Otherwise, we have the Lie product $[[[k_{1,1}], [k_{1,2}]], [k_2]]$. By the Jacobi identity,

$$[[[k_{1,1}], [k_{1,2}]], [k_2]] = [[[k_{1,1}], [k_2]], [k_{1,2}]] - [[[k_{1,2}], [k_2]], [k_{1,1}]].$$

As all three of $[k_{1,1}]$, $[k_{1,2}]$, and $[k_2]$ are basic Lie products, the underlying words are regular. As $k_{1,1} > k_{1,2} > k_2$, $k_{1,1}k_{1,2} > k_{1,2}k_{1,1}$, $k_{1,2}k_2 > k_2k_{1,2}$, and $k_{1,1}k_2 > k_2k_{1,1}$. Thus $k = k_{1,1}k_{1,2}k_2 > k_{1,2}k_{1,1}k_2 > k_{1,2}k_2k_{1,1}$ and $k > k_{1,1}k_2k_{1,2}$. By the second inductive hypothesis, these commutators are sums of basic Lie products. Then $[k]$ is a sum of basic Lie products which completes both inductive steps. Any element of M is a linear combination of commutators. Therefore, the basic Lie products span M.

In particular, the basic Lie products in the generators of L span L.

□

We may now prove that L is a free Lie ring.

Theorem 2. *The Lie ring L, constructed from the $\mathbb{Q}$-algebra A on the well ordered set of generators $\{x_n\}_{n\in B}$, where B is an indexing set, is a free Lie ring on the free generators x_i.*

Proof. By the preceding lemma, the basic Lie products in the given generators span L. These basic Lie products are in fact linearly independent. Supposing otherwise, there is a

finite sum $\sum_{i=1}^{n} \alpha_i[u_i]$ with distinct basic Lie products $[u_i]$ and non-zero integers α_i such that $\sum_{i=1}^{n} \alpha_i[u_i] = 0$. Expand each $[u_i]$ as a linear combination of monomials v_j in A. Thus

$$\alpha_i[u_i] = \sum_j \beta_j v_j.$$

By Lemma 9,

$$\alpha_i[u_i] = \beta_i u_i + \sum_k \beta_k w_k,$$

where $u_i > w_k$ for some monomials w_k.

Since this is a finite sum, there is a u_k greater than all the other u_i. There is then a unique basic Lie product $[u_k]$ such that

$$\alpha_k[u_k] + \sum_{i=1}^{k-1} \alpha_i[u_i] + \sum_{i=k+1}^{n} \alpha_i[u_i] = 0.$$

Thus, the sum of monomials

$$\alpha_k u_k + \sum_m \zeta_m z_m = 0,$$

where the z_m are monomials in A, the ζ_m are in $\mathbb{Q}$, and $u_k > z_m$ for all m. This contradicts the linear independence of monomials in A and shows that the basic Lie products in L are linearly independent.

Now let F be a free Lie ring generated by free generators $\{f_n\}_{n \in B}$, where B is the same indexing set for the generators of A (the existence of which we assume). If we order them in the same way, the basic Lie products in the f_i span F. As F is free, we have an onto homomorphism from F to L sending f_i to x_i. Homomorphisms preserve products and thus the image of a basic Lie product in f_i is a basic Lie product in x_i as the two sets of generators have the same order. Since the basic Lie products in L are linearly independent, the only way an element f in F can be sent to the zero element by our homomorphism in L is if

$f = 0$. Thus, we have an isomorphism from F to L. Therefore, L is a free Lie ring on the x_i.

□

Corollary 2. *The Lie $\mathbb{Q}$-algebra $\mathbb{Q}L$ is a free Lie $\mathbb{Q}$-algebra on the generators x_i.*

To conclude this chapter, we record two theorems (without proof) that are important in relation to our treatment of matrix groups in the later chapters: the Poincaré-Birkhoff-Witt Theorem and Ado's Theorem. We must first define a few relevant terms, chief of which is the concept of a universal enveloping algebra.

Definition 16. *We say that an associative algebra A over the field F is a universal enveloping algebra of the Lie F-algebra L if there is an F-linear map ϕ such that*

1. *$\phi(L)$ generates A,*
2. *$\phi([x, y]) = \phi(x)\phi(y) - \phi(y)\phi(x)$ for all $x, y \in L$,*
3. *if B is any associative algebra with a mapping θ from L into B satisfying the previous two properties, there is a unique homomorphism η from A to B such that $\theta(x) = \eta(\phi(x))$ for all $x \in L$.*

We denote a universal enveloping algebra by (A, ϕ).

We will not establish the existence of a universal enveloping algebras for each Lie algebra L as this is well-documented. Naturally, the question arises as to whether there is only one universal enveloping algebra of L: the answer is indeed yes.

Theorem 3. *Let L be a Lie algebra and (A, ϕ) a universal enveloping algebra of L. If (B, θ) is another universal enveloping algebra of L, then $A \cong B$.*

Proof. From the definition of a universal enveloping algebra, there is a homomorphism η from A to B such that $\theta(x) = \eta(\phi(x))$ for all $x \in L$. This homomorphism is surjective as $\phi(L)$ generates B. Since B is also a universal enveloping algebra, we have a surjective map η' from B to A such that $\phi(x) = \eta'(\theta(x))$ for all $x \in L$.

Thus

$$\eta(\eta'(\theta(x))) = \eta(\phi(x)) = \theta(x)$$

and

$$\eta'(\eta(\phi(x))) = \eta'(\theta(x)) = \phi(x).$$

The functions η and η' are then inverses of each other. Therefore, η is an isomorphism and $A \cong B$.

□

This concept is fundamental to the Poincaré-Birkhoff-Witt Theorem, which we now state without proof from Varadarajan [8, Theorem 3.2.2].

Theorem 4. ***Poincaré-Birkhoff-Witt Theorem***

Let L be a Lie algebra over a field F of characteristic 0, (A, ϕ) its universal enveloping algebra, J a linearly ordered set, and $\{x_i : i \in J\}$ a basis for L. Then the elements 1 and $\phi(x_{i_i}) \dots \phi(x_{i_s})$ form a basis for A for all $s \geq 1$ and $i_1 \leq \cdots \leq i_s$. In particular, ϕ is an injection on L.

Important in its own right, the Poincaré-Birkhoff-Witt Theorem demonstrates that any Lie algebra L over a field of characteristic 0 is isomorphic to a Lie subalgebra of the Lie algebra formed on its universal enveloping algebra. Additionally, it explicitly states a basis of the universal enveloping algebra. Both of these results are crucial in the proof of Ado's theorem [8, Theorem 3.17.7]. To state this theorem, we first must define representations of Lie algebras.

Definition 17. *Let L be any Lie algebra over a field F and V a vector space over the same field. A representation of L in V is a map π from L into the endomorphism ring $gl(V)$ such that*

1. *π is F-linear*

2. $\pi([x, y]) = \pi(x)\pi(y) - \pi(y)\pi(x)$ *for all* $x, y \in L$.

We say a representation is faithful if its kernel is trivial.

Theorem 5. ***Ado's Theorem***

Let L *be a finite-dimensional Lie algebra over a field* F *of characteristic 0, and let* N *be the unique maximal nilpotent ideal of* L. *Then there exists a faithful finite-dimensional representation* ρ *of* L *such that* $\rho(N)$ *is nilpotent.*

If the field is $\mathbb{R}$, Ado's theorem proves that every finite-dimensional Lie $\mathbb{R}$-algebra is isomorphic to a matrix Lie algebra over $\mathbb{R}$. If the field is $\mathbb{Q}$, every finite-dimensional Lie $\mathbb{Q}$-algebra L is isomorphic to a matrix Lie algebra over $\mathbb{Q}$.

CHAPTER 3

THE LIE CORRESPONDENCE

Much of the material in this chapter is based on the work of Stillwell [7]. We begin with definitions of paths and tangent vectors. Then we show that the tangent space of a matrix Lie group is a Lie algebra in a natural way. Introducing the logarithm and exponential maps allows us to talk about the relationships between these two structures, culminating with a bijection between certain neighborhoods.

Throughout this chapter, G denotes a matrix Lie group, defined in the first chapter as a subgroup of $GL_l(\mathbb{R})$, the group of real invertible l-by-l matrices, that is additionally closed under nonsingular limits.

Definition 18. *A path from a matrix x_0 to x_1 in the matrix Lie Group G is a continuous map $p(t)$ from a closed interval $[a, b]$ to G such that $p(a) = x_0$ and $p(b) = x_1$.*

Let $p : [a, b] \to G$ be any path in G. Note that $p(t)$ is a matrix in G for each $t \in [a, b]$. If we look at the coordinates of that matrix, $p(t) = (p_{ij}(t))$, then we see that each $p_{ij}(t)$ is a continuous function in $\mathbb{R}$ [3, Theorem 19.6]. The path $p(t)$ is said to be smooth if each $p_{ij}(t)$ is differentiable [7, Chapter 5.1].

The derivative of a path p at t is defined as

$$p'(t) = \lim_{h \to 0} \frac{p(t+h) - p(t)}{h}$$

for every t in the interior of the domain of p such that the limit exists. If $p(t) = (p_{ij}(t))$, then $p'(t) = (p'_{ij}(t))$.

For our purposes, we will almost always be looking at paths with a domain $[a, b]$ that contains 0 in its interior and such that $p(0) = I$. Stealing an idea from calculus, we define a tangent vector at the identity I of a matrix Lie group G in the following way.

Definition 19. *A tangent vector at I, the identity matrix of a matrix Lie group G, is defined as the derivative of p evaluated at 0, where p is a smooth path in G such that 0 is contained in the interior of its domain and $p(0) = I$.*

One can now define the tangent space of a matrix Lie group.

Definition 20. *Let G be a matrix Lie group. The tangent space $T_I(G)$ at the identity I of G is the collection of all tangent vectors at I.*

The tangent space of a matrix Lie group G will be shown to be a Lie algebra, under a natural Lie bracket operation. Moreover, there exist two functions, the logarithmic and the exponential functions, which play a great part in relating properties of G and the Lie algebra $T_I(G)$.

First, we will show that the tangent space of any matrix Lie group is a Lie algebra in a natural way.

Lemma 11. *Let G be a matrix Lie group. Then $T_I(G)$ is a Lie algebra over the real numbers with Lie multiplication given by $[y, x] = yx - xy$ for all $x, y \in T_I(G)$.*

Proof. We know that $T_I(G)$ is non-empty as the path defined by $p(t) = I$ for all $t \in [-1, 1]$ is a smooth continuous map with $p(0) = I$. Thus $p'(0) = 0 \in T_I(G)$.

Suppose $x, y \in T_I(G)$. Then $x = p'(0)$ and $y = q'(0)$ for smooth paths p, q with $p(0) = I = q(0)$. The product of the two paths $pq(t) = p(t)q(t)$ is a smooth path with domain equal to the intersection of the domains of p and q. Its derivative can be calculated from the product rule, which holds for every coordinate function of p and q, $(pq)'(t) = p'(t)q(t) + p(t)q'(t)$. Thus

$$(pq)'(0) = p'(0) + q'(0) = x + y.$$

Therefore, $x + y$ is contained in $T_I(G)$ for any two $x, y \in T_I(G)$.

Now, let $\alpha \in \mathbb{R}$. Then $(\alpha p)'(t) = \alpha p'(t)$. Thus $\alpha p'(0) = \alpha x \in T_I(G)$. The set $T_I(G)$ is then a subspace of the matrix ring containing it, making $T_I(G)$ a vector space over $\mathbb{R}$ and a subspace of $M_l(\mathbb{R})$.

Next, let x, p, y, and q be as above. To see that $T_I(G)$ is a Lie algebra over $\mathbb{R}$, look at the smooth path,

$$r_s(t) = p(s)^{-1} q(t) p(s),$$

where s is a fixed number in $[0, 1]$. As a product of a smooth path and two constant matrices, $r_s(t)$ is a smooth path for every s and

$$r_s(0) = p(s)^{-1} p(s) = I.$$

The matrix $r_s'(0)$ is therefore contained in $T_I(G)$ for all s. Taking the derivative with respect to t produces

$$r_s'(0) = p(s)^{-1} q'(0) p(s) = p(s)^{-1} y p(s).$$

Note that $(p(0)^{-1})' = -p'(0)$ as

$$p(s)p(s)^{-1} = I$$

$$p'(s)p(s)^{-1} + p(s)(p(s)^{-1})' = 0$$

$$p'(s)p(s)^{-1} = -p(s)(p(s)^{-1})'.$$

When $s = 0$,

$$p'(0) = -(p(0)^{-1})'.$$

Let $f(s) = r_s'(0)$. It follows that f is a smooth function of s as it is a product of smooth

functions of s. If we take the derivative of f with respect to s, we find that

$$f'(0) = p(0)^{-1}yp'(0) + (p(0)^{-1})'yp(0)$$

$$= p(0)^{-1}yp'(0) - p'(0)yp(0)$$

$$= yx - xy.$$

Since $T_I(G)$ is a finite-dimensional subspace of $M_l(\mathbb{R})$, it is a normed vector space over the real numbers. Thus $T_I(G)$ is complete, making it a closed subspace of $M_l(\mathbb{R})$ [5, Theorem 3.11]. Since

$$f'(0) = \lim_{h \to 0} \frac{p(h)^{-1}yp(h) - p(0)^{-1}yp(0)}{h},$$

the derivative of f at 0 must be contained in $T_I(G)$, as $f'(0)$ is a limit of elements in $T_I(G)$. Thus, whenever $y, x \in T_I(G)$ then $[y, x] = yx - xy$ is also contained in $T_I(G)$.

Finally, the Lie bracket defined by $[y, x] = yx - xy$ makes $T_I(G)$ into a Lie algebra over the real numbers, as the bracket quite easily satisfies $[x, x] = 0$ and the Jacobi identity.

□

An important result follows directly from our proof that $T_I(G)$ is a Lie $\mathbb{R}$-algebra for any matrix Lie group G. If H is a normal subgroup of G, then we may look at the tangent space $T_I(H)$. What is the relation between the two tangent spaces?

Lemma 12. *If H is a normal subgroup of a matrix Lie group G, then $T_I(H)$ is an ideal of $T_I(G)$.*

Proof. From our previous proof, $T_I(H)$ is a real vector space. As every tangent vector to H at I is a tangent vector to G at I, $T_I(H)$ is a subspace of $T_I(G)$. Since $T_I(H)$ is a matrix Lie algebra in its own right, $T_I(H)$ is a Lie subalgebra of $T_I(G)$. Note that this holds even if H is not a normal subgroup.

Let $x = p'(0) \in T_I(H)$ and $y = q'(0) \in T_I(G)$ for some smooth paths p and q in H and

G respectively, where $p(0) = I = q(0)$. As H is a normal subgroup of G, $q(s)^{-1}p(t)q(s)$ is contained in H for any appropriate s and t. Let $r_s(t) = q(s)^{-1}p(t)q(s)$.

As seen in Lemma 11, $f(s) = r_s'(0) = q(s)^{-1}xq(s)$. Again, as H is a normal subgroup of G, $f(s)$ is a smooth path in $T_I(H)$. Thus $f'(0) = xy - yx = [x, y]$ is contained in $T_I(H)$ as a limit of elements of $T_I(H)$. Therefore, if H is a normal subgroup of the matrix Lie group G, $[x, y] \in T_I(H)$ for any $x \in T_I(H)$ and $y \in T_I(G)$, making $T_I(H)$ an ideal of $T_I(G)$.

□

Now that we know that the tangent space of any matrix Lie group is itself a matrix Lie algebra, we can begin studying the relations between the two. Our previous lemma provides a nice start, but we would like to push it further.

To do so, we will begin by taking a neighborhood around the identity element of the matrix Lie group G and show that there exists a natural bijection between it and a neighborhood about the 0 matrix in the tangent space. In some sense, this is the main result in the Lie correspondence.

We begin by utilizing two familiar maps which we will define for matrices: the logarithm and the exponential map.

For motivation, recall the Mercator series for the natural logarithm. Given by

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n$$

for all real numbers x with absolute value less than 1, the Mercator series converges to $\ln(1+x)$. The idea then becomes to create something similar for square matrices. Accordingly, throughout the next two lemmas we let $A \in M_l(\mathbb{R})$ for some fixed integer l.

Lemma 13. *The logarithm of a square matrix A, defined by*

$$\log(I + A) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} A^n,$$

converges for all square matrices A with matrix absolute value $|A|$ less than 1. Moreover, it is continuous on the set of all real square matrices A such that $|A| < 1$.

Proof. From our definition, we can write the logarithm as the limit of

$$f_k(A) = \sum_{n=1}^{k} \frac{(-1)^{n+1}}{n} A^n$$

as k approaches infinity. Fix an $\epsilon > 0$, and let us investigate $|f_k(A) - f_m(A)|$ for some integers k and m, assuming without loss of generality that $k > m$.

$$|f_k(A) - f_m(A)| = |\sum_{n=m+1}^{k} \frac{(-1)^{n+1}}{n} A^n| \leq \sum_{n=m+1}^{k} \frac{1}{n}|A|^n \leq \sum_{n=0}^{\infty} |A|^n.$$

Now,

$$\sum_{n=0}^{\infty} |A|^n$$

converges for $|A| < 1$, as it is a geometric series. It follows that for some sufficiently large positive integer N,

$$\sum_{n=N}^{\infty} \frac{1}{n}|A|^n < \epsilon.$$

Thus $|f_k(A) - f_m(A)| < \epsilon$ for all k and m greater than N. The $f_k(A)$ then form a Cauchy sequence. Since the metric space of real matrices is a complete metric space, every Cauchy sequence converges. Thus, the f_k converge to our logarithm, making log a well-defined function for all $|A| < 1$.

To see that it is continuous on the stated set requires a closer look. Since each A^m is a continuous function of A, each f_k is a finite sum of continuous functions, and so is a continuous function. The sequence $\{f_k\}_{k=1}^{\infty}$ converges uniformly to the logarithm on any compact subset of matrices with absolute value less than 1. Therefore, $\log(I + A)$ is a continuous function on the set of all matrices A with $|A| < 1$. □

Additionally, when A is a 1-by-1 real matrix with absolute value less than 1, $\log(1 + A)$

agrees perfectly with the familiar natural logarithm. We now define the exponential function similarly.

Lemma 14. *The exponential of a square matrix A, defined by*

$$\exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!} = I + \sum_{n=1}^{\infty} \frac{A^n}{n!},$$

converges for all real matrices. Moreover, $\exp(A)$ *is a continuous function on the metric space* $M_l(\mathbb{R})$.

Proof. As for the logarithm, we find that the partial sums of $\exp(A)$ form a Cauchy-sequence. Thus, $\exp(A)$ converges for every square matrix, making it a well-defined function. It is continuous as for every compact subset of $M_l(\mathbb{R})$, the partial sums uniformly converge to $\exp(A)$. As seen in the case of the logarithms, the partial sums are continuous. Therefore, $\exp(A)$ is a continuous function on $M_l(\mathbb{R})$.

□

Now, we need a couple of more properties of our two functions to go onward. If $|\exp(A) - I| < 1$ for any square matrix A, then

$$\log(\exp(A)) = \log(I + (\exp(A) - I))$$

$$= \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} (\exp(A) - I)^n$$

$$= A + \sum_{k=2}^{\infty} \frac{A^k}{k!} + \sum_{n=2}^{\infty} \frac{(-1)^{n+1}}{n} (\exp(A) - I)^n.$$

However, if we repeated the process for $\exp(x)$, where x is some real number such that $|\exp(x) - 1| < 1$, we would get a similar result:

$$\log(\exp(x)) = x + \sum_{k=2}^{\infty} \frac{x^k}{k!} + \sum_{n=2}^{\infty} \frac{(-1)^{n+1}}{n} (\exp(x) - 1)^n.$$

If we look at the coefficient of A^k in $\log(\exp(A))$ for each $k > 1$, we see that it must agree with the coefficient of x^k in $\log(\exp(x))$. As $\log(\exp(x)) = x$, the coefficient for x^k must be 0 for any integer k greater than 1. Therefore, $\log(\exp(A))$ must equal A when $|\exp(A) - I| < 1$. A bit more work shows that log and exp are inverses of each other when restricted in such a manner [1, Theorem 2.8].

Additionally, we can show that $\exp(A + B) = \exp(A)\exp(B)$ for any two square matrices A and B which commute. To prove this, first note that

$$\exp(A)\exp(B) = \sum_{n=0}^{\infty} \frac{A^n}{n!} \sum_{m=0}^{\infty} \frac{B^m}{m!}.$$

If we rewrite this expression as a sum of terms of power m, then

$$\exp(A)\exp(B) = \sum_{m=0}^{\infty} \sum_{n=0}^{m} \frac{A^n}{n!} \frac{B^{m-n}}{(m-n)!}$$

$$= \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{n=0}^{m} \frac{m!}{n!(m-n)!} A^n B^{m-n}.$$

As A and B commute,

$$\exp(A)\exp(B) = \sum_{m=0}^{\infty} \frac{1}{m!}(A + B)^m = \exp(A + B),$$

as required.

We next investigate the effect of the exponential function on tangent vectors at I of some matrix Lie group G. Let p be a smooth path in G with $p(0) = I$ so that $p'(0)$ is a tangent vector to G at I. As p is a path, it is continuous. Fix $\frac{1}{2} > \epsilon > 0$. Thus, for large enough positive integer N, we have that $|p(\frac{1}{n}) - I| < \epsilon$ for all integers $n > N$.

By the definition of log, we have

$$\frac{\log(p(\frac{1}{n}))}{\frac{1}{n}} = \frac{p(\frac{1}{n}) - I}{\frac{1}{n}} + \sum_{k=2}^{\infty} \frac{(-1)^{k+1}}{\frac{k}{n}} (p(\frac{1}{n}) - I)^k$$

$$= \frac{p(\frac{1}{n}) - I}{\frac{1}{n}} - \frac{p(\frac{1}{n}) - I}{\frac{1}{n}} \sum_{k=2}^{\infty} \frac{(-1)^k}{k} (p(\frac{1}{n}) - I)^{k-1}.$$

If $n > N$, then

$$\Big| \sum_{k=2}^{\infty} \frac{(-1)^k}{k} (p(\frac{1}{n}) - I)^{k-1} \Big| \leq \sum_{k=2}^{\infty} \frac{|p(\frac{1}{n}) - I|^{k-1}}{k}$$

$$< \sum_{k=2}^{\infty} \frac{\epsilon^{k-1}}{k} < \sum_{k=1}^{\infty} \epsilon^k.$$

The series $\sum_{k=1}^{\infty} \epsilon^k$ converges to $\frac{\epsilon}{1-\epsilon}$, which is less than 2ϵ, and so the limit of $\sum_{k=2}^{\infty} \frac{(-1)^k}{k} (p(\frac{1}{n}) - I)^{k-1}$ as n grows without bound is the 0 matrix.

Thus, the limit of $n \log(p(\frac{1}{n}))$ as n goes to infinity is

$$\lim_{n\to\infty} \frac{p(\frac{1}{n}) - I}{\frac{1}{n}} - \lim_{n\to\infty} \frac{p(\frac{1}{n}) - I}{\frac{1}{n}} \sum_{k=2}^{\infty} \frac{(-1)^k}{k} (p(\frac{1}{n}) - I)^{k-1}$$

$$p'(0) - p'(0) \cdot 0 = p'(0).$$

Consequently $\exp(p'(0))) = \exp(\lim_{n\to\infty} n \log(p(\frac{1}{n}))))$. Since exp is continuous, it follows that

$$\exp(\lim_{n\to\infty} n \log(p(\frac{1}{n}))) = \lim_{n\to\infty} \exp(n \log(p(\frac{1}{n}))),$$

and since the powers of a matrix commute with each other, we have that

$$\lim_{n\to\infty} \exp(n \log(p(\frac{1}{n}))) = \lim_{n\to\infty} (\exp(\log(p(\frac{1}{n}))))^n$$

$$= \lim_{n\to\infty} (p(\frac{1}{n}))^n.$$

This means $\exp(p'(0))$ is a limit of matrices in G. Since the inverse of $\exp(p'(0))$ is $\exp(-p'(0))$ (which is similarly a limit of matrices in G), $\exp(p'(0))$ is nonsingular. Thus $\exp(p'(0))$ must also be contained in G. As $p'(0)$ was an arbitrary element of $T_I(G)$, exp takes every element of $T_I(G)$ to an element of G.

Therefore, our exponential function defined in Lemma 14 maps $T_I(G)$ into G.

To see that log maps a neighborhood of I into $T_I(G)$, we must first define sequential tangent vectors.

Definition 21. *A matrix $X \in M_l(\mathbb{R})$ is a sequential tangent vector to G at I if there is a sequence $\{A_k\}_{k=1}^{\infty}$ of matrices contained in G and a sequence of real numbers $\{\alpha_k\}_{k=1}^{\infty}$, such that the A_k converge to I and*

$$\lim_{k\to\infty} \frac{A_k - I}{\alpha_k} = X.$$

If p is any smooth path with $p(0) = I$, then the sequence $\{p(\frac{1}{n})\}_{n=1}^{\infty}$ converges to I and the sequence $\{\frac{p(\frac{1}{n})-I}{\frac{1}{n}}\}_{n=1}^{\infty}$ converges to $p'(0)$ as we saw above. Thus, every tangent vector is a sequential tangent vector. Surprisingly, the converse is also true, and this fact features heavily in demonstrating log maps a neighborhood about I in G into $T_I(G)$ [7, Chapter 7.3].

Lemma 15. *Every sequential tangent vector X to G at I is a tangent vector to G at I.*

Proof. Let X be any non-zero sequential tangent vector with matrix sequence $\{A_k\}_{k=1}^{\infty}$ and real numbers sequence $\{\alpha_k\}_{k=1}^{\infty}$. It follows that the α_k must converge to 0, as otherwise $\frac{A_k-I}{\alpha_k}$ would converge to 0 instead of X. The sequence $\{\frac{1}{\alpha_k}\}_{k=1}^{\infty}$ must then diverge. Define a new sequence $\{\frac{1}{\beta_k}\}_{k=1}^{\infty}$ where β_k is the closest integer to $\frac{1}{\alpha_k}$. This sequence must also converge to 0, and $\beta_k(A_k - I)$ must converge to X.

Thus, for large enough k

$$\log(A_k^{\beta_k}) = \beta_k \log(A_k) = \beta_k(A_m - I) - \beta_k(A_m - I)\sum_{n=2}^{\infty} \frac{(-1)^n}{n}(A_k - I)^{n-1}.$$

As before, the limit of $\log(A_k^{\beta_k})$ as k increases without bound is $\lim_{k\to\infty} \beta_k(A_m - I) = X$. Thus

$$\exp(X) = \lim_{k\to\infty} \exp(\log(A_k^{\beta_k})) = \lim_{k\to\infty} A_k^{\beta_k}.$$

Then $\exp(X)$ is contained in G. By a similar argument, $\exp(tX)$ is also contained in G for any real number $t \in [-1, 1]$. Thus the function $p = \exp(tX)$ is a smooth path

with $p(0) = I$. Therefore, any sequential tangent vector X is also a tangent vector, as $p'(0) = X\exp(0X) = X$.

□

We now restrict our focus to the neighborhood $N_\delta(I) = \{A \in G \ : |A - I| < \delta\}$ of the identity in G. We will show that there exists some δ such that log maps $N_\delta(I)$ into $T_I(G)$.

Lemma 16. *There exists a $\delta > 0$ such that $\log(N_\delta(I)) \subset T_I(G)$.*

Proof. Begin by supposing the contrary. We can then construct a sequence of matrices $\{A_n\}_{n=1}^{\infty}$ in G that converges to I with $\log(A_n) \notin T_I(G)$. Since G is a group of real l-by-l matrices, $T_I(G)$ is a subspace of $M_l(\mathbb{R})$. Thus, $\log(A_n) = X_n + Y_n$, where $X_n \in T_I(G)$ and Y_n is contained in the orthogonal complement of $T_I(G)$ in $M_l(\mathbb{R})$. As the A_n converge to I, $X_n + Y_n$ must converge to 0. Since both $T_I(G)$ and $T_I(G)^\perp$ are closed subspaces of $M_l(\mathbb{R})$, both X_n and Y_n must converge to 0.

If we observe the sequence of matrices $\frac{Y_n}{|Y_n|}$, each contained in $T_I(G)^\perp$, then we can view it as a sequence of unit vectors in $\mathbb{R}^{l^2}$, all of which lie on the ball of radius 1 about the 0 vector. Since this is a bounded sequence, there must be a convergent subsequence. Let Y be the limit of this convergent subsequence. Thus Y must be a vector of length 1 contained in $T_I(G)^\perp$. Let $\frac{Y_k}{|Y_k|}$ be this subsequence converging to Y.

Now $T_k = \exp(-X_k)A_k$ is an element of G, since $-X_k \in T_I(G)$. Moreover, since the logarithm and exponential functions are inverse functions we see that $A_k = \exp(X_k + Y_k)$. Thus

$$\begin{aligned} T_k &= \exp(-X_k)\exp(X_k + Y_k) \\ &= \exp(-X_k)[\sum_{m=0}^{\infty} \frac{(X_k + Y_k)^m}{m!}] \\ &= \exp(-X_k)[A + \sum_{m=0}^{\infty} \frac{(X_k)^m}{m!} + \frac{(Y_k)^m}{m!}], \end{aligned}$$

where A is a sum of elements that are powers of products of X_k and Y_k. Consequently,

$$T_k = \exp(-X_k)[A + \exp(X_k) + \exp(Y_k)]$$

$$= I + \exp(-X_k)\exp(Y_k) + \exp(-X_k)A$$

$$= I + \exp(-X_k)\sum_{m=0}^{\infty}\frac{(Y_k)^m}{m!} + \exp(-X_k)A$$

$$= I + Y_k + B,$$

where B is again a sum of elements that are powers of products of X_k and Y_k.

Since both X_k and Y_k converge to 0, B converges to 0 as well. Thus

$$\frac{T_k - I}{|Y_k|} = \frac{Y_k + B}{|Y_k|} = \frac{Y_k}{|Y_k|} + \frac{B}{|Y_k|},$$

and the limit of $\frac{T_k - I}{|Y_k|}$ as k approaches infinity must then be Y.

We know that the T_k converge to I as both $\exp(-X_k)$ and A_k converge to I. Thus Y is a sequential tangent vector. However, Y is not in $T_I(G)$, which creates a contradiction. There must then exist some $\delta > 0$ such that $\log(N_\delta(I)) \subset T_I(G)$. □

Therefore, log **maps some neighborhood** $N_\delta(I)$ **into** $T_I(G)$**.**

More importantly, as both log and exp are continuous inverses of each other, they create a continuous bijection between $N_\delta(I)$ and its image in $T_I(G)$.

This brings us to the Campbell-Baker-Hausdorff Theorem. Throughout our discussion, all matrices A, B, C, etc. belong to $M_l(\mathbb{R})$ for some fixed positive integer l. If both $\exp(A)$ and $\exp(B)$ are contained in $N_\delta(I)$ and $\exp(A)\exp(B) = \exp(C)$ for some $A, B, C \in T_I(G)$, we know when A and B commute that $C = A + B$. What can be said about C when A and B do not commute?

To begin, we attempt to find C by considering $\exp(A)\exp(B)$. This expands to

$$I + A + B + AB + \frac{A^2}{2!} + \frac{B^2}{2!} + \cdots + \frac{A^m B^n}{m!n!} + \ldots,$$

where m and n are any positive integers. Taking the logarithm produces

$$C = \log(\exp(A)\exp(B)) = \log(I + A + B + AB + \frac{A^2}{2!} + \frac{B^2}{2!} + \cdots + \frac{A^m B^n}{m!n!} + \ldots)$$

$$= (A + B + AB + \frac{A^2}{2!} + \frac{B^2}{2!} + \cdots + \frac{A^m B^n}{m!n!} + \ldots)$$
$$-\frac{1}{2}(A + B + AB + \frac{A^2}{2!} + \frac{B^2}{2!} + \cdots + \frac{A^m B^n}{m!n!} + \ldots)^2 + \ldots$$
$$= A + B + AB - \frac{1}{2}AB - \frac{1}{2}BA + \ldots$$
$$= A + B + \frac{1}{2}[A, B] + \ldots.$$

Let $F_n(A, B)$ be the homogeneous component of degree n in this expression. Then we may rewrite C as $C = \sum_{n=1}^{\infty} F_n(A, B)$. We wish to show that each F_n is a Lie polynomial, meaning it is a linear combination of Lie brackets in terms of A and B. Visibly, $F_1(A, B)$ and $F_2(A, B)$ are Lie polynomials.

The following proof of the Campbell-Baker-Hausdorff theorem is due to Eichler [7, Chapter 7.7]

Theorem 6. ***Campbell-Baker-Hausdorff Theorem***

For each $n \geq 1$, the polynomial $F_n(A, B)$ is a Lie polynomial.

Proof. Let X, Y, and Z be any square matrices. Since matrix multiplication is associative,

$$[\exp(X)\exp(Y)]\exp(Z) = \exp(X)[\exp(Y)\exp(Z)].$$

Thus

$$W = \log([\exp(X)\exp(Y)]\exp(Z)) = \log(\exp(X)[\exp(Y)\exp(Z)])$$

$$\sum_{i=1}^{\infty} F_i\big(\sum_{j=1}^{\infty} F_j(X,Y),Z\big) = \sum_{i=1}^{\infty} F_i\big(X,\sum_{j=1}^{\infty} F_j(Y,Z)\big).$$

As an induction hypothesis, we assume that F_m is a Lie polynomial for all $m < n$. Now, observe all the homogeneous polynomials of degree n or less in W. If they are of degree less than n, then they are a linear combination of F_m for some $m < n$, making them Lie by the hypothesis. If the homogeneous polynomial is a linear combination of polynomials of the form $F_m(F_k(X,Y),Z)$ or $F_m(X,F_k(Y,Z)))$ on the right side where m and k are integers that sum to n, then they are Lie by the inductive hypothesis. This leaves only the homogeneous polynomials $F_n(X,Y)+F_n(X+Y,Z)$ on the left and $F_n(X,Y+Z)+F_n(Y,Z)$ on the right.

We create a congruence relation from these two polynomials determined by

$$F_n(X,Y)+F_n(X+Y,Z) =_{Lie} F_n(X,Y+Z)+F_n(Y,Z),$$

where $=_{Lie}$ means that the difference of the left and right hand sides is a Lie polynomial.

By manipulating what matrix is where, we can develop several identities using this relation.

If we set $Z=-Y$ where X and Y are any matrices, then

$$F_n(X,Y)+F_n(X+Y,-Y) =_{Lie} F_n(X,0)+F_n(-Z,Z).$$

As $\exp(X)\exp(0)=\exp(X)$ and $\exp(-Z)\exp(Z)=I$,

$$F_n(X,0)+F_n(-Z,Z)=0.$$

Thus we have relation (1),

$$F_n(X,Y) =_{Lie} -F_n(X+Y,-Y).$$

If $X = -Y$ where Z and Y are any matrices, then

$$F_n(-Y, Y) + F_n(0, Z) =_{Lie} F_n(-Y, Y + Z) + F_n(Y, Z)$$

$$0 =_{Lie} F_n(-Y, Y + Z) + F_n(Y, Z)$$

$$F_n(Y, Z) =_{Lie} -F_n(-Y, Y + Z),$$

which we denote by relation (2).

By relation (2) for any matrices X and Y,

$$F_n(X, Y) =_{Lie} -F_n(-X, X + Y).$$

By relation (1),

$$F_n(X, Y) =_{Lie} -(-F_n(-X + X + Y, -X - Y))$$

$$=_{Lie} F_n(Y, -X - Y).$$

From relation (2),

$$F_n(X, Y) =_{Lie} -F_n(-Y, -X).$$

Since every monomial in F_n has degree n, $F_n(-Y, -X) = (-1)^n F_n(Y, X)$. Thus we arrive at relation (3):

$$F_n(X, Y) =_{Lie} (-1)^{n+1} F_n(Y, X).$$

If $Z = \frac{-Y}{2}$ for any matrices X and Y,

$$F_n(X, Y) + F_n(X + Y, \frac{-Y}{2}) =_{Lie} F_n(X, \frac{Y}{2}) + F_n(Y, \frac{-Y}{2}).$$

As Y and $-Y/2$ commute, $F_n(Y, -Y/2) = 0$. Thus

$$F_n(X, Y) =_{Lie} F_n(X, \frac{Y}{2}) - F_n(X + Y, \frac{-Y}{2}),$$

which we denote by relation (4).

If $X = \frac{-Y}{2}$ for any matrices Y and Z,

$$F_n(\frac{-Y}{2}, Y) + F_n(\frac{Y}{2}, Z) =_{Lie} F_n(\frac{-Y}{2}, Y + Z) + F_n(Y, Z)$$

$$F_n(\frac{Y}{2}, Z) =_{Lie} F_n(Y, Z) + F_n(\frac{-Y}{2}, Y + Z).$$

Thus for any two matrices X and Y,

$$F_n(\frac{X}{2}, Y) =_{Lie} F_n(X, Y) + F_n(\frac{-X}{2}, X + Y)$$

$$F_n(X, Y) =_{Lie} F_n(\frac{X}{2}, Y) - F_n(\frac{-X}{2}, X + Y),$$

which we denote by relation (5).

We now observe $F_n(\frac{X}{2}, Y)$ and $F_n(\frac{-X}{2}, X + Y)$ individually, but we first note that

$$F_n(\alpha X, \alpha Y) = \alpha^n F_n(X, Y)$$

for all $\alpha \in \mathbb{R}$. Then by relation (4),

$$F_n(\frac{X}{2}, Y) =_{Lie} F_n(\frac{X}{2}, \frac{Y}{2}) - F_n(\frac{X}{2} + Y, \frac{-Y}{2}).$$

From relation (1), it follows that

$$F_n(\frac{X}{2}, Y) =_{Lie} F_n(\frac{X}{2}, \frac{Y}{2}) + F_n(\frac{X}{2} + \frac{Y}{2}, \frac{Y}{2})$$

$$=_{Lie} 2^{-n}\{F_n(X, Y) + F_n(X + Y, Y)\}.$$

Again from relation (4),

$$F_n(\frac{-X}{2}, X+Y) =_{Lie} F_n(\frac{-X}{2}, \frac{X+Y}{2}) - F_n(\frac{X}{2}+Y, \frac{-X-Y}{2}).$$

By relations (2) and (1),

$$F_n(\frac{-X}{2}, X+Y) =_{Lie} -F_n(\frac{X}{2}, \frac{Y}{2}) + F_n(\frac{Y}{2}, \frac{X+Y}{2})$$

$$=_{Lie} 2^{-n}\{F_n(Y, X+Y) - F_n(X, Y)\}.$$

Thus

$$F_n(X,Y) =_{Lie} 2^{-n}\{F_n(X,Y) + F_n(X+Y,Y) - F_n(Y,X+Y) + F_n(X,Y)\}$$

$$=_{Lie} 2^{-n}\{F_n(X+Y,Y) - F_n(Y,X+Y) + 2F_n(X,Y)\}.$$

From relation (3), it follows that

$$(1-2^{1-n})F_n(X,Y) =_{Lie} 2^{-n}\{F_n(X+Y,Y) + (-1)^n F_n(X+Y,Y)\}$$

$$(1-2^{1-n})F_n(X,Y) =_{Lie} 2^{-n}(1+(-1)^n)F_n(X+Y,Y).$$

We denote this result as relation (6).

If n is odd, then $F_n(X,Y) =_{Lie} 0$ by relation (6). If n is even, then the relation (6) holds for the matrices $X-Y$ and Y.

$$(1-2^{1-n})F_n(X-Y,Y) =_{Lie} 2^{-n}(2)F_n(X,Y)$$

$$-(1-2^{1-n})F_n(X,-Y) =_{Lie} 2^{1-n}F_n(X,Y)$$

$$-F_n(X,-Y) =_{Lie} \frac{2^{1-n}}{(1-2^{1-n})}F_n(X,Y),$$

which is the final relation (7).

Relation (7) holds for the matrices X and $-Y$. Thus by relation (7) applied to $F_n(X, -Y)$,

$$-F_n(X,Y) =_{Lie} \frac{2^{1-n}}{(1-2^{1-n})} F_n(X,-Y)$$

$$-F_n(X,Y) =_{Lie} -\{\frac{2^{1-n}}{(1-2^{1-n})}\}^2 F_n(X,Y)$$

$$0 =_{Lie} F_n(X,Y).$$

Therefore, by induction, $F_n(X,Y) =_{Lie} 0$ for odd or even $n \geq 1$. □

We shall see the utility of this deep theorem below and even more heavily in the next chapter. In the remainder of this chapter, we record (without proof) some fundamental relationships between matrix Lie group homomorphisms $\phi : G \to H$ and Lie algebra homomorphisms $\Phi : T_I(G) \to T_I(H)$.

First, we define what a homomorphism is for matrix Lie groups.

Definition 22. *A matrix Lie group homomorphism is a homomorphism between two matrix Lie groups G and H that is also an infinitely differentiable map.*

We may then state a large result for the Lie correspondence (without proof) from Stillwell [7].

Theorem 7. *For any matrix Lie group homomorphism $\Phi : G \to H$ of matrix Lie groups G and H with Lie algebras $\mathfrak{g} = T_I(G)$ and $\mathfrak{h} = T_I(H)$ respectively, there is a Lie algebra homomorphism $\phi : \mathfrak{g} \to \mathfrak{h}$ such that*

$$\phi(p'(0)) = (\Phi \circ p)'(0)$$

for any smooth path $p(t)$ through I in G.

To create a matrix Lie group homomorphism out of a Lie algebra homomorphism is a

much more difficult matter. In fact, we can only reliably do so for simply connected matrix Lie groups, which we proceed to define.

Definition 23. *A set S is a path connected set, if, for any $A, B \in S$, there is a path $p : [a, b] \to S$ with $p(a) = A$ and $p(b) = B$.*

From there, we can define a simply connected topological space.

Definition 24. *A topological space S is a simply connected space, if S is path connected and, for any two paths p and q that begin at A and end at B, there is a homotopy relative to the endpoints between the two paths. In other words, there is a continuous function $d : [0, 1] \times [0, 1] \to S$ such that $d(0, t) = p(t)$, $d(1, t) = q(t)$, $d(s, 0) = p(0) = q(0)$, and $d(s, 1) = p(1) = q(1)$ for all t and s in $[0, 1]$. (Assuming the paths are on $[0, 1]$).*

If a matrix Lie group G is simply connected, then every element of G is a product of elements in a neighborhood $N_\delta(I)$ about I [7, Chapter 8.6]. Since there is a bijection between a neighborhood about 0 in $T_I(G)$ and $N_\delta(I)$, we can represent any element of G as a product of $\exp(X_i)$, where $X_i \in T_I(G)$.

As such, if we have a Lie algebra homomorphism ϕ between two Lie algebras $\mathfrak{g} = T_I(G)$ and $\mathfrak{h} = T_I(H)$, we may define $\Phi(\exp(X)) = \exp(\phi(X))$ for all $X \in \mathfrak{g}$. Since every element of G is a product of $\exp(X_i)$, where $X_i \in T_I(G)$, we can extend Φ to arbitrary elements of G. Due to the connected properties of G, we can show that Φ is well-defined. The Campbell-Baker-Hausdorff Theorem shows that Φ is a Lie group homomorphism. Additionally, Φ induces ϕ as in Theorem 7. Thus we have the following theorem, also stated from Stillwell [7].

Theorem 8. *If $\mathfrak{g} = T_I(G)$ and $\mathfrak{h} = T_I(H)$ are the Lie algebras of the simply connected matrix Lie groups G and H, and if $\phi : \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism, then there is a matrix Lie group homomorphism $\Phi : G \to H$ that induces ϕ (in the sense of Theorem 7).*

CHAPTER 4

THE MAL'CEV CORRESPONDENCE

In the previous chapter, we explored the logarithm and exponential functions between a matrix Lie group G and its tangent space $T_I(G)$. We defined them as a power series in which the terms were matrices. For most groups G, these two functions are infinite series that do not terminate. However, if we assume that G is a nilpotent group, the logarithm function terminates and is much easier to calculate. A similar assumption occurs for the exponential function when its domain is suitably restricted as we see below.

First, we define a couple of objects. Let $T_n(R)$ denote the group of upper triangular n-by-n matrices with 1s on the diagonal and entries from the ring R. This group is often referred to as the group of unitriangular n-by-n matrices with entries in R. Let $U_n(F)$ be the associative algebra of strictly upper triangular matrices over a field F. For the rest of this section, we concern ourselves solely with $T_n(\mathbb{Q})$ and $U_n(\mathbb{Q})$.

Note that the associative algebra $U_n(\mathbb{Q})$ is nilpotent. In fact, $U_n(\mathbb{Q})$ is nilpotent of class $n-1$. Thus, the Lie algebra $U_n(\mathbb{Q})^-$ is nilpotent by Lemma 2.

Let us go back to our definitions of log and exp. In the previous chapter, we were concerned with convergence. However, since $U_n(\mathbb{Q})$ is nilpotent, the logarithm function converges on $T_n(\mathbb{Q})$ whereas the exponential function converges on $U_n(\mathbb{Q})$. Thus, we can modify the domains of the functions as follows.

Definition 25. *The logarithm of a square matrix $m = I + a \in T_n(\mathbb{Q})$ is defined by*

$$\log(m) = \log(I + a) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} a^k.$$

Since $m \in T_n(\mathbb{Q})$, it follows that $a \in U_n(\mathbb{Q})$. Since $U_n(\mathbb{Q})$ is nilpotent, $\log(m)$ is a finite sum mapping $T_n(\mathbb{Q})$ into $U_n(\mathbb{Q})$.

Definition 26. *The exponential of a square matrix $a \in U_n(\mathbb{Q})$ is defined by*

$$\exp(a) = \sum_{k=0}^{\infty} \frac{a^k}{k!} = I + \sum_{k=1}^{\infty} \frac{a^k}{k!}.$$

Since $a \in U_n(\mathbb{Q})$, the sum $\sum_{k=0}^{\infty} \frac{a^k}{k!}$ terminates after $k = n-1$. Thus $\exp(a) = I + b$, where b is a finite sum of elements in $U_n(\mathbb{Q})$. It follows that exp maps $U_n(\mathbb{Q})$ into $T_n(\mathbb{Q})$. As before, these functions are inverses of each other which can be shown by routine calculations.

We now note some important properties of $T_n(\mathbb{Q})$.

Lemma 17. *The group $T_n(\mathbb{Q})$ is a complete nilpotent torsion-free group. Thus $T_n(\mathbb{Q})$ is a matrix Mal'cev group.*

Proof. It is well known that $T_n(\mathbb{Q})$ is nilpotent and torsion-free, which can be shown by modifying a theorem given in [4, Chapter 5.1].

Let $t \in T_n(\mathbb{Q})$, let k be non-zero positive integer, and set $s = \exp(\frac{1}{k}\log(t))$. Since $\frac{1}{k}\log(t)$ commutes with itself our previous results show that $s^k = \exp(\log(t)) = t$. Thus, t has a kth root in $T_n(\mathbb{Q})$ for every non-zero integer k. Therefore the group $T_n(\mathbb{Q})$ is a complete group and so $T_n(\mathbb{Q})$ is a matrix Mal'cev group by Lemma 3.

□

We now slightly modify our notation in order to improve the readability of the coming results.

Definition 27. *Let G be a matrix Mal'cev group. Define $\flat : G \to U_n(\mathbb{Q})^-$ by $g^\flat = \log(g)$ for all $g \in G$.*

Let L be the image of G under the map $\flat$. Define $\sharp : L \to G$ by $l^\sharp = \exp(l)$ for all $l \in L$.

Let $(\ ,\ )$ denote a commutator in G in order to avoid using $[\ ,\ ]$ for both commutators and Lie products.

These changes allow us to more succinctly state a result of Jennings [6, Remark 2.2.4)].

Lemma 18. *Let G be a matrix Mal'cev group and L its image under $\flat$. Then the Campbell-Baker-Hausdorff Theorem implies that for $g_1, \ldots, g_m \in G$,*

$$(g_1, \ldots, g_m)^\flat = [g_1^\flat, \ldots, g_m^\flat] + \sum_w P_w$$

where each P_w is a rational linear combination of products $[g_{i_1}^\flat, \ldots, g_{i_w}^\flat]$ with $w > m$ and $i_j \in \{1, \ldots, m\}$, such that every integer in $\{1, \ldots, m\}$ appears at least once among the i_j.

We may now state the main result of this chapter which we take from Stewart [6, Theorem 2.2.1].

Theorem 9. *Let G be a matrix Mal'cev group and L its image under $\flat$. Then we have the following results.*

1. *The maps $\flat$ and $\sharp$ are mutual inverses.*
2. *If H is a complete subgroup of G then $H^\flat$ is a Lie subalgebra of L. In particular, L is a Lie subalgebra of $U_n(\mathbb{Q})^-$.*
3. *If M is a subalgebra of L then $M^\sharp$ is a complete subgroup of G.*
4. *If H is a complete normal subgroup of a complete subgroup K of G, then $H^\flat$ is an ideal of $K^\flat$.*
5. *If M is an ideal of a subalgebra N of L, then $M^\sharp$ is a complete normal subgroup of $N^\sharp$.*

The first result follows immediately from the properties of log and exp. The next four results are much more involved, and we prove them here as lemmas for the proof of Theorem 9. Accordingly, we use the notation established in Theorem 9 throughout the remainder of this section.

Lemma 19. *If H is a complete subgroup of G then $H^\flat$ is a Lie subalgebra of L. In particular $L = G^\flat$ is a Lie subalgebra of $U_n(\mathbb{Q})^-$.*

Proof. Suppose we have an expression

$$h^\flat + \sum_j \lambda_j C_j,$$

where the $h \in H$, the $\lambda_j \in \mathbb{Q}$, and the C_j are Lie products of weight $r \geq 1$ or greater in elements of $H^\flat$. Then we can rewrite the expression above as

$$h^\flat + \sum_{i=1}^{s} \mu_i D_i + \sum_k \nu_k E_k,$$

where the D_i are Lie products of weight r, the E_k are Lie products of weight greater than r, and the μ_i and ν_k are scalars.

Now each $D_i = [h_1, \ldots, h_r]$ can be rewritten as $D_i = (h_1, \ldots, h_r)^\flat + \sum_w P_w$ by Lemma 18, where each P_w is a rational linear combination of Lie products in the h_r of weight greater than r. Denote $(h_1, \ldots, h_r)$ by $g_i \in G$. Then for any $q \in \mathbb{Q}$ it follows from the Campbell-Baker-Hausdorff Theorem that

$$(hg^q)^\flat = h^\flat + (g_i^q)^\flat + \sum_l \beta_l A_l$$

$$= h^\flat + q g_i^\flat + \sum_l \beta_l A_l,$$

where the β_l are scalars and the A_l are Lie products of weight at least 2 formed from $h^\flat$ and $g_i^\flat$. As $g_i^\flat = D_i - \sum_w P_w$, it follows that $g_i^\flat$ is a linear combination of Lie products of weight r or greater in $H^\flat$. Thus, each A_l is a linear combination of Lie products of weight greater than r in $H^\flat$. Additionally, as $\mu_i D_i = \mu_i (g_i)^\flat + \sum_w \mu_i P_w$, we have that

$$h^\flat + \mu_i D_i = h^\flat + \mu_i (g_i)^\flat + \sum_w \mu_i P_w$$

$$= (hg_i^{\mu_i})^\flat + \sum_w \mu_i P_w - \sum_l \beta_l A_l.$$

We can then "remove" each D_i and see that

$$h^\flat + \sum_{i=1}^{s} \mu_i D_i + \sum_k \nu_k E_k = (hg_1^{\mu_1} g_2^{\mu_2} \dots g_s^{\mu_s})^\flat + \sum_b \gamma_b B_b,$$

where $\gamma_b \in \mathbb{Q}$ and each B_b is a Lie product of weight greater than r in $H^\flat$.

Note that $hg_1^{\mu_1} g_2^{\mu_2} \dots g_s^{\mu_s}$ is an element of H as H is a complete subgroup of G.

Let L_H be the Lie subalgebra of $U_n(\mathbb{Q})^-$ generated by $H^\flat$. Then every element $l \in L_H$ is a linear combination of Lie products of weight $r \geq 1$, i.e.

$$l = \sum_j \lambda_j C_j = I^\flat + \sum_j \lambda_j C_j,$$

where each λ_j is a rational number and C_j is a Lie product of weight $r \geq 1$.

As shown, there is an $g \in H$ such that

$$l = g^\flat + \sum_i \mu_i D_i,$$

where the μ_i are rational numbers and the D_i are Lie products of weight greater than 1. We may continue to repeat this process until we have

$$l = h^\flat + \sum_k \nu_k E_k,$$

where $h \in H$, the $\nu_k \in \mathbb{Q}$, and the E_k are Lie products of weight greater than $n-1$, the nilpotency class of $U_n(\mathbb{Q})$. Thus every $E_k = 0$.

Therefore, $l = h^\flat$. This shows that $L_H = H^\flat$, and so $H^\flat$ is a Lie algebra. In particular, $L = G^\flat$ is a Lie subalgebra of $U_n(\mathbb{Q})^-$. The image $H^\flat$ of a complete subgroup H of a matrix Mal'cev group G is then a Lie subalgebra of L, as claimed. □

We can now show that any subalgebra of L is mapped to a complete subgroup of G by $\sharp$.

Lemma 20. *If M is a subalgebra of L then $M^\sharp$ is a complete subgroup of G.*

Proof. Let M be a Lie subalgebra of L and H the group generated by $M^\sharp$. Let $a, b \in M$ and $\lambda \in \mathbb{Q}$. By the definition of $\sharp$, $0^\sharp = I$. By the Campbell-Baker-Hausdorff Theorem,

$$(a^\sharp b^\sharp)^\flat = \sum_{i=1}^{\infty} F_i(a, b),$$

where the functions F_i are those that appear in Theorem 6.

As $U_n(\mathbb{Q})^-$ is nilpotent, the $F_i(a, b) \in M$ are all 0 for large enough positive integers i. Thus, $(a^\sharp b^\sharp)^\flat \in M$ and $a^\sharp b^\sharp \in M^\sharp$.

As $(a^\sharp)^\lambda = (\lambda a)^\sharp$, all rational powers of $(a^\sharp)$ are contained in $M^\sharp$. In particular, $(a^\sharp)^{-1} \in M^\sharp$. Therefore, $M^\sharp$ is a complete subgroup of G.

□

The fact that our functions $\flat$ and $\sharp$ send complete subgroups to subalgebras and vice-versa is a very desirable property, but the Mal'cev correspondence does more: it preserves complete normal subgroups and ideals.

Lemma 21. *If H is a complete normal subgroup of a complete subgroup K of a matrix Mal'cev group G, $H^\flat$ is an ideal of $K^\flat$.*

Proof. Consider the Lie products $A_r = [a_1^\flat, \ldots, a_r^\flat]$ where $a_i \in K$ for all i such that $1 \leq i \leq r$ and $a_j \in H$ for at least one j such that $1 \leq j \leq r$. For r greater than the nilpotency class of $G^\flat$, we have that $A_r = 0 \in H^\flat$. We will then argue by reverse induction on r. Assume that $A_k \in H^\flat$ for all $k > r$.

Then by Lemma 18, we have that

$$(a_1, \ldots, a_r)^\flat = A_r + \sum_w P_w,$$

where each P_w is a rational linear combination of Lie products of weight $\lambda > r$ of the form $[a_{i_1}^\flat, \ldots, a_{i_\lambda}^\flat]$. As in Lemma 18, each integer in $\{1, \ldots, r\}$ appears at least once in the i_λ.

By our assumption, each P_w is contained in $H^\flat$. Since H is a normal subgroup of K, $(a_1, \ldots, a_r) \in H$ as long as there is at least one j $(1 \leq j \leq r$ such that $a_j \in H$. Thus

$$A_r = (a_1, \ldots, a_r)^\flat - \sum_w P_w$$

is a linear combination of elements of $H^\flat$. Therefore, $A_r \in H^\flat$. When $r = 2$, we see that $[k, h] \in H^\flat$ for every $k \in K$ and $h \in H$. Thus, $H^\flat$ is an ideal of $K^\flat$.

In particular, if H is a complete normal subgroup of G, then $H^\flat$ is an ideal of $G^\flat$, as claimed.

□

We now arrive at the last stated property of our correspondence.

Lemma 22. *If M is an ideal of a Lie subalgebra N of $U_n(\mathbb{Q})$, then $M^\sharp$ is a complete normal subgroup of $N^\sharp$.*

Proof. Let $m \in M$ and $k \in N$. Then by Lemma 18,

$$(m^\sharp, k^\sharp)^\flat = [m, k] + \sum_w P_w,$$

where each P_w is a rational linear combination of Lie products in m and k of weight greater than 2. Since M is an ideal of N, we have that each $P_w \in M$. Thus, $(m^\sharp, k^\sharp)^\flat \in M$ and $(m^\sharp, k^\sharp) \in M^\sharp$.

Since $(m^\sharp, k^\sharp) = (m^\sharp)^{-1}(k^\sharp)^{-1}m^\sharp k^\sharp$ and $m^\sharp \in M^\sharp$, it follows that $m^\sharp(m^\sharp, k^\sharp) = (k^\sharp)^{-1}m^\sharp k^\sharp$ is an element of $M^\sharp$ for all $k \in N$. As $\sharp$ is a bijection, every element of $N^\sharp$ is of the form $k^\sharp$ for some $k \in N$. Therefore, $M^\sharp$ is a complete normal subgroup of $N^\sharp$.

□

This completes the proof of Theorem 9.

As we remarked earlier, we restrict to matrix Mal'cev groups to mirror the chapter on the Lie correspondence. However, the Mal'cev correspondence can be broadened. Stewart does

so by extending the previous results to complete locally nilpotent torsion-free groups using extended words [6]. Khukhro does so by obtaining a correspondence for arbitrary nilpotent $\mathbb{Q}$-powered groups using the construction of free Lie $\mathbb{Q}$-algebras in the second chapter [2].

CHAPTER 5

CONCLUSION

Throughout this paper, we have given a survey of several important ideas. To begin, we gave a brief overview of basic group and Lie ring properties. We then went on to construct the free Lie rings and $\mathbb{Q}$-algebras, stating the Poincairé-Birkhoff-Witt Theorem and Ado's Theorem as other important theorems discovered along the way. Next, we explored the two correspondences which are heavily related. The Lie correspondence looked at real matrix Lie groups G and the Lie algebras formed from their tangent spaces $T_I(G)$. The Mal'cev correspondence looked at the matrix Mal'cev groups G and the Lie $\mathbb{Q}$-algebras formed from their image under the exponential map.

They both take groups of matrices to Lie algebras over the same field. They both try to find information about either the groups or the Lie algebras given the other. They both can be extended even further given the more advanced machinery of topology and differential geometry. However, the techniques and methods of proving the Lie and Mal'cev correspondences are rather powerful in their own right. Each of the the two main results were originally topological ideas. Demonstrating that they can be studied using mostly algebraic methods is worthwhile.

REFERENCES

[1] Hall, B. C. (2015). *Lie Groups, Lie Algebras, and Representations*. Springer, 2nd edition edition.

[2] Khukhro, E. I. (1998). *p-Automorphisms of Finite p-Groups*. Cambridge University Press.

[3] Munkres, J. R. (2000). *Topology*. Pearson, 2nd edition.

[4] Robinson, D. J. S. (1996). *A Course in the Theory of Groups*. Springer, 2nd edition.

[5] Rudin, W. (1976). *Principles of Mathematical Analysis*. McGraw-Hill, 3rd edition.

[6] Stewart, I. N. (1970). An algebraic treatment of Mal'cev's theorems concerning nilpotent lie groups and their lie algebras. *Compositio Mathematica*, 22(3):289–312.

[7] Stillwell, J. (2008). *Naive Lie Theory*. Springer.

[8] Varadarajan, V. S. (1984). *Lie Groups, Lie Algebras, and Their Representations*. Springer.