INTRODUCTION TO LIE ALGEBRAS, ENGEL'S THEOREM, AND LIE'S THEOREM

AUSTIN FENG

ABSTRACT. This paper discusses elementary properties of Lie algebras, providing lower dimensional and common examples while introducing fundamental concepts such as solvability and nilpotency. It proves two important theorems: Engel's theorem and Lie's theorem. Engel's theorem gives a criterion for Lie algebra nilpotency, and Lie's theorem shows that finite solvable linear Lie algebras have a basis where they are all upper triangular. This paper ends by introducing Killing forms and Cartan's criterion, which are crucial to more advanced studies of Lie algebras.

CONTENTS

1. INTRODUCTION

Lie algebras are a structural extension of vector spaces that provide a vector space with a bracket operation that satisfies certain identities, namely bilinearity, alternativity, and the Jacobi Identity. Plenty of examples arise when analyzing the space of linear transformations on a finite dimensional vector space. It is useful to analyze these spaces in order to derive results such as when the transformations have a common eigenvector or conditions on whether repeated applications of a transformation can yield zero. Lie algebras are important outside of pure mathematics as well; they are used extensively in physics, specifically quantum mechanics and particle physics, and have associations with Lie Groups, which are groups that also function as differentiable manifolds([5]).

This paper discusses elementary properties of Lie algebras, enumerating them in lower dimensions as well as providing common examples, while introducing fundamental concepts such as ideals, adjoint representation, solvability and nilpotency, along with fundamental lemmas and theorems for each. It proves two important theorems: Engel's theorem and Lie's theorem. Engel's theorem gives a criterion for Lie algebra nilpotency, and Lie's theorem shows that finite solvable linear Lie algebras have a basis where they are all upper triangular. This paper ends by introducing Killing forms and Cartan's criterion, which are crucial to more advanced studies of Lie algebras.

2. Basics of Lie algebras

2.1. **Lie algebras, Ideals, and Homomorphisms.** We begin by defining Lie algebras and their basic structures. Before formally defining Lie algebras, we note a few important properties about their underlying fields. Since a Lie algebra contains the structure of a vector space, we must establish a field on which it functions. Although this field may be arbitrary, we will work with the field of complex numbers, denoted $\mathbb{C}$. This field has two crucial properties, it is of characteristic zero[1] and it is algebraically closed.[2]

Definition 2.1. A **Lie algebra** L is a vector space over a field F that is coupled with a binary operation known as the **bracket**: $(x, y) \mapsto [x, y]$ that satisfies the following axioms:

- Bilinearity: For each $x \in L$, the maps $[x, -], [-, x] : L \to L$ are linear
- Alternativity: $[x, x] = 0$ for all $x \in L$
- The Jacobi Identity: $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$ for all x,y,$z \in L$.

From here on out, the symbol L will, by default, denote a finite-dimensional Lie algebra. Thinking about the Lie algebra as a vector space, the most immediate definition that follows is that of a subalgebra, which is informally a closed subspace that contains all key properties that define it.

Definition 2.2. A **subalgebra** S of a Lie algebra L is a subspace of L that is closed under the bracket operation. Formally, if $x, y \in S \subset L$, then $[x, y] \in S$.

Plenty of the structure of Lie algebras depends on the concept of an ideal:

Definition 2.3. An **ideal** $I \subset L$ is a subalgebra of L such that if $x \in I$ and $y \in L$, then $[x, y] \in I$.

Definition 2.4. An **abelian** Lie algebra is a Lie algebra with $[x, y] = 0$ for all $x, y \in L$.

Definition 2.5. A **simple** Lie algebra is one that is not abelian and has no nontrivial ideals. That means that its ideals are 0 and itself, while the bracket operator is nonzero.

---

[1]Fields have a property known as characteristic, which is the number of times that one must add the field's multiplicative identity to get its additive identity, and zero when such number does not exist. Lie algebras can be defined over all fields, however many major properties require the characteristic of the field to be zero.

[2]This means all polynomials with coefficients in $\mathbb{C}$ have roots in $\mathbb{C}$. This is notably not true for $\mathbb{R}$ since the quadratic polynomial $x^2 + 1 = 0$, for example, has only non-real roots.

Example 2.6. Here are some examples of ideals:

- Trivial ideals: The sets 0 and L
- The center: $Z(L) := \{x \in L | [x, L] = 0\}$
- The derived algebra: denoted $[L, L] := \{\sum_{i=1}^{n} [x_i, y_i] | x_i, y_i \in L \text{ and } n \in \mathbb{N}\}$.

Another way to view the derived algebra is the ideal generated by $[a, b]$ for all $a, b \in L$. The derived algebra of L is akin to the commutator subgroup in a group. The commutator subgroup is the group generated by all commutators in the group.

Homomorphisms and isomorphisms in Lie algebras are defined in a way that echoes the definition for standard transformations between vector spaces.

Definition 2.7. A **homomorphism** of two Lie algebras, call them L and L', is a linear map $F : L \to L'$ such that $F([x, y]) = [F(x), F(y)]$.

Definition 2.8. We say a homomorphism $F : L \to L$ is an **isomorphism** if F is bijective. We say F is an **endomorphism** if $L = L'$, and F is an **automorphism** if F is both an endomorphism and an isomorphism.

2.2. **Examples.** For a n-dimensional vector space V, an example of a Lie algebra comes from the space $\operatorname{End} V$ of all linear transformations from V to itself. By choosing a suitable basis of V, an element of $\operatorname{End} V$ can be represented by a $n \times n$ matrix, and we can define the bracket operator as $[x, y] := xy - yx$. It is fairly easy to see that this operation gives $\operatorname{End} V$ a structure of a Lie algebra. To see that the Jacobi Identity holds, note that for any x, y, $z \in \operatorname{End} V$:

$$\begin{aligned}
[x, [y, z]] + [y, [z, x]] + [z, [x, y]] &= [x, (yz - zy)] + [y, (zx - xz)] + [z, (xy - yx)] \\
&= xyz - xzy - yzx + zyx + yzx - yxz - zxy \\
&\quad + xzy + zxy - zyx - xyz + yxz \\
&= 0.
\end{aligned}$$

Notation 2.9. The previously constructed Lie algebra $gl(V)$ is called the General Linear Lie algebra of V. Its elements are linear endomorphisms $V \to V$, and its bracket is the commutator $[x, y] = (xy - yx)$.

The space of $n \times n$ matrices is a good way to study this Lie algebra and converting it to so is considered to be "harmless, and very convenient" by mathematicians such as Humphreys ([1] pg.2). Any subalgebra of such Lie algebras $gl(V)$ is called a **linear Lie algebra**.

2.3. **Lie algebras in Lower Dimensions.** It is easy to enumerate and classify low-dimensional Lie algebras up to isomorphism. One-dimensional Lie algebras are all abelian and hence isomorphic. This is since L =span(v) for any nonzero v, and $[v, v] = 0$. Even though this Lie algebra has no non-trivial ideals, it is not considered simple since it is abelian.

In two dimensions we have two possibilities. In either case L =span(x, y). One possibility is that $[x, y] = 0$. In this case, the Lie algebra is abelian. Otherwise, if $[x, y] \neq 0$, then $[x, y] = a$ for some nonzero element $a \in L$. All nonzero products are multiples of a. So we can simply replace x with some vector in the direction of a to get $[x, y] = cx$ for a constant c. We then re-scale y so that $[x, y] = x$. This means that, up to isomorphism, all non-abelian two-dimensional Lie algebras are isomorphic to the Lie algebra defined by $[x, y] = x$.

2.4. **Subsets of $gl(V)$ and Classical Algebras.** As usual, let V be a vector space of dimension n. Another example of a Lie algebra on V is called the **special linear algebra** $sl(V)$, defined as the set of endomorphisms with trace zero. This space forms a subalgebra of $gl(V)$ because it is closed under the bracket operation: the bracket operation of any two matrices with trace zero is a matrix with trace zero. These matrices are subject to one restriction, which is that the sum of diagonal entries is zero. Considering the diagonal, the first $n-1$ terms uniquely determine the diagonal. The last element is confined to be the value which makes the whole diagonal sum to zero. Since all non-diagonal entries are independent of each other and the diagonal, the dimension of this Lie algebra is n^2-1.

Other examples of Lie algebras that are subalgebras of $gl(V)$ include upper triangular matrices, strictly upper triangular matrices, and diagonal matrices. Checking that each one is closed under the bracket operation for $gl(V)$, one can verify that these are indeed Lie algebras. For upper triangular matrices the dimension is $\frac{n(n+1)}{2}$ and for strictly upper triangular ones it is $\frac{n(n-1)}{2}$. Diagonal matrices have dimension n.

The **symplectic Lie algebra** (only defined in spaces where $\dim V := n = 2{*}l$ for some $l \in \mathbb{N}$) is the algebra of endomorphisms A that satisfy $s(Av,w)+s(v,Aw)=0$ for s a skew-symmetric, bilinear form on V with matrix:

$$\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

where I has dimension $l = \frac{n}{2}$. A skew-symmetric matrix is one whose transpose equals its negative. What remains to be checked is that the bracket is closed under such a space. This follows from it being linear in both arguments by definition of bilinear.

We can now consider the dimensions of a symplectic Lie algebra. We can verify a matrix x is symplectic if $s(xv,w)+s(v,xw)=0 \leftrightarrow sx = -x^T s$. Writing the matrix x as $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$, we see $B = B^T$, $C = C^T$ and $A = -D^T$ in order for the equation to hold. Note how this forces the trace of the matrix to be zero, since the diagonal of the upper left will be the negative of the diagonal of the bottom right. From this relationship we can get constraints that allow us to compute the Lie algebra's dimension. Using this construction we can see that the dimension of this space is $2l^2+l$, which can be verified by using the above description of the matrix and a similar analysis method as for the special linear algebra.

2.5. **Derivations and adjoint representation.** For a vector space V endowed with a binary operation, a **derivation** is a map M from V to itself satisfying the product rule, $M(ab) = aM(b) + bM(a)$. Define $Der(V)$ as the space of all derivations on V. The binary operation for Lie algebras is automatically the bracket operation.

In the definition of derivation, V can be any Lie algebra. It is natural to define a derivation taking $y \to [x,y]$ for a fixed x; this is called $\operatorname{ad} x$. To verify that it is indeed a derivation one must verify the product rule. Since the Jacobi Identity can be written as $[x,[y,z]] = [[x,y],z] + [y,[x,z]]$, we can see this to be true by taking a to be y, b to be z, and $M(t) = [x,t]$.

Definition 2.10. The map $ad : L \to Der(L)$ sending each element x to its adjoint $\operatorname{ad} x$ is called the **adjoint representation** of L.

3. Solvability

We now begin to look at important features that Lie algebras can have. Two key ones are solvability and nilpotency, defined in similar but distinct ways.

Definition 3.1. The **derived series** of a Lie algebra L is defined inductively as follows: $L^{(0)} = L, L^{(1)} = [L, L], L^{(i)} = [L^{(i-1)}, L^{(i-1)}]$ for all $i \in \mathbb{N}$.

Lemma 3.2. *The elements of the derived series are ideals.*

Proof. We show $L^{(n)}$ is an ideal by examining $[x, y]$ for $x \in L$ and $y \in L^{(n)}$ and proving that it is in $L^{(n)}$.

We prove this by induction. To verify the base case, we observe that $L^{(0)}$ equals L and is hence an ideal, and that $L^{(1)}$ is trivially an ideal, because any element $v \in [L, L]$ has to be an element in L, and thus $[v, x]$ for any other $x \in L$ is in $[L, L]$.

For any integer i, we assume $L^{(i)}$ is an ideal and examine $L^{(i+1)}$. For any element $a \in L^{(i)}$ and element $b \in L^{(i)}$, $[a, b] \in [L^{(i)} L^{(i)}]$, so $[a, b] \in L^{(i+1)}$ by definition of derived series. Now consider an element $x \in L$. Any element $l \in L^{(i+1)}$ is the sum of elements $[a, b]$ for some a, $b \in L^{(i)}$. We are allowed to consider only one term due to linearity, so assume that $l = [a, b]$. Now, $[x, l] = [x, [a, b]] = [a, [x, b]] + [b, [a, x]]$ by the Jacobi Identity. By the induction hypothesis, $[a, x]$, $[x, b] \in L^{(i)}$ and thus $[x, l] \in L^{(i+1)}$. □

Definition 3.3. A Lie algebra (or ideal) L is **solvable** if $L^{(n)} = 0$ for some $n \in \mathbb{N}$.

Examples of solvable matrices include diagonal and strictly upper triangular matrices.

We must ultimately reach the definition of the radical of L, the largest solvable ideal of L. However, it benefits us to first prove theorems about quotient spaces and homomorphisms.

Definition 3.4. For a Lie algebra L and an ideal I we can define the **quotient space** L/I as the space of equivalence classes $[x + I]$ given by the equivalence relation $x\ y$ if (and only if) $x - y \in I$. Also, define $(x + I) + (y + I) := (x + y) + I$. Here, $[(x+I), (y+I)] := [x, y]+I$. Both of these are well-defined. For example the latter is unambiguous since if one switches out x, y with x', y' with $x' = x+u, u \in I$ and $y' = y+v, v \in I$, $[x', y'] = [x, y]+[u, v]+[u, y]+[x, v]$ and since the latter three terms are in I, $[x', y'] + I = [x, y] + I$ and thus they are the same element in L/I.

There is a **canonical map** $\pi : L \to L/I$ given by $\pi(x) = x + I$ for $x \in L$.

For Lie algebras L and L', suppose $\phi : L \to L'$ is a homomorphism, meaning that $\phi([x, y]) = [\phi(x), \phi(y)]$. It should be evident that the kernel of ϕ denoted $\operatorname{Ker}\phi$ is an ideal: for any $x \in \operatorname{Ker}\phi$, $\phi(x) = 0$, $\phi([x, y]) = [\phi(x), \phi(y)] = 0$ so $[x, y] \in \operatorname{Ker}\phi$. To an ideal I there is a one-to-one correspondence between ideals containing I and the ideals of L/I, with the canonical map $\pi : L \to L/I$. And there is a one-to-one correspondence between maps ϕ and kernels $\operatorname{Ker}\phi$ in L.

Lemma 3.5. *Let $\phi : L \to L'$ be a homomorphism. There exists an isomorphism from $L/(\operatorname{Ker}\phi)$ to $\operatorname{Im}\phi$. Additionally, if I is an ideal included in $(\operatorname{Ker}\phi)$ and π is the canonical map from L to L/I, then there is a unique map $M : L/I \to L'$ whose composition with π equals ϕ.*

Proof. Observe that for any $x \in L$ where $\phi(x) = y$ and any element $k \in \operatorname{Ker}\phi$ we have $\phi(x + k) = y$. Therefore, if we map an element $\alpha \in L/(\operatorname{Ker}\phi)$ to $\phi(x)$ for any representative $x + \operatorname{Ker}\phi = \alpha$ then the entire image must be represented. The map is unambiguous due to the above argument, and no two different elements in $L/(\operatorname{Ker}\phi)$ can map to the same point in the image.

Additionally, for the map M, define it in a similar way: since I is in $\operatorname{Ker}\phi$ we can define maps from L/I to L' unambiguously. Therefore, for any $k \in L/I$ we take a representative element x and map k to $\phi(x)$. It is easy to check that this map is unambiguous and that its composition with π commutes with ϕ, which means they will produce the same outcome for any inputs. Furthermore, such map is unique, as if any different map existed, they must map some element $k \in L/I$ differently to L'. But the fact that it must commute makes that impossible, since every element x' inside the equivalence class defined by k must be mapped to $\phi(x)$ for the representative x. □

Lemma 3.6. *If I and J are ideals such that $I \subset J$, then J/I is an ideal of L/I and $(L/I)/(J/I)$ is isomorphic to L/J.*

Proof. We first verify J/I is an ideal by taking $x + I \in L/I$ and $y + I \in J/I$ and verifying that $[x + I, y + I] \in J/I$. By definition $[x + I, y + I] = [x, y] + I$. Since $y \in J$ and J is an ideal $[x, y] \in J$, thus $[x, y] + I \in J/I$.

To show $(L/I)/(J/I)$ is equivalent to L/J we can take an obvious map from $L/I \to L/J$ and prove that its kernel is J/I. Specifically we define a map ϕ from $L/I \to L/J$ as $\phi(x + I) = x + J$. This is unambiguous since $I \subset J$ and its kernel is J/I since all elements that differ by an element of J are mapped to the same thing. Thus, $(L/I)/(J/I)$ is equivalent to L/J by Lemma 3.5.

□

Lemma 3.7. *If I and J are ideals then there is an isomorphism between $I/(I \cap J)$ and $(I + J)/J$.*

Proof. The proof is similar to the above. Specifically we map each element in I to itself plus J in $(I + J)/J$, written as $\phi(x) = (x + J)/J$. The transformation yields zero if and only if x is contained in I as well as J, so in short it needs to be in $I \cap J$. So there exists an isomorphism between $I/(I \cap J)$ and $(I + J)/J$.

□

The above three statements, Lemmas 3.5-3.7, are commonly known as the three isomorphism theorems and will be very useful later on. Now, back to the topic of radical of L.

Lemma 3.8.
a) If L is solvable, all subalgebras and homomorphic images are solvable.
b) If I is a solvable ideal and L/I is solvable, then so is L.
c) If I and J are solvable ideals then $I + J$ is solvable.

Proof. a) If K is a subalgebra of L, then $K^{(i)} \subset L^{(i)}$ for all $i \in \mathbb{N}$. Hence if L is solvable, so is any subalgebra K. If $\phi : L \to M$ is a homomorphism then induction on i shows that $\phi((L^{(i)})) \subset M^{(i)}$. Thus $\phi(L)$ is solvable.

b) Say that $(L/I)^{(n)} = 0$. Now note that using part (a) on π, the canonical map $L \to L/I$, $\pi(L^{(n)}) = 0$. Thus, $L^{(n)}$ is in the kernel of π. Since π is the canonical map, the kernel is precisely I. So since $I^{(m)} = 0$ for some integer m, $L^{(m+n)} = 0$.

c) To prove this, note that $I/(I \cap J)$ is solvable as a homomorphic image of I. But that is isomorphic to $(I + J)/J$ by Lemma 3.8, which must be solvable. Then we get $(I + J)$ is solvable since J is solvable using part (b). □

In light of Lemma 3.9(c) we may defined the radical of a Lie algebra L.

Definition 3.9. The **radical** of L, denoted $\operatorname{Rad} L$, is the largest solvable ideal of L.

Such radical is unique (there cannot be two equally sized solvable ideals where there are none of strictly larger size). This is obtained by Lemma 3.8(c), since if there are two such non-identical largest radicals, their sum must be an even larger radical, which is a contradiction. Furthermore, any solvable ideal is a subset of the radical, because if it weren't, then the union of that solvable ideal and the radical will be larger than the radical and be a solvable ideal, which is a contradiction. The existence of a unique radical allows us to define semisimple Lie algebras.

Definition 3.10. A **semisimple** Lie algebra L is one where $\operatorname{Rad} L = 0$, i.e. its maximal solvable ideal is 0.

A simple algebra L has no ideals except the trivial one and itself, so its maximal solvable ideal is 0. This is because L itself is nonsolvable, since the derived series will always be mapping to L. As a result, the only other ideal it has, the zero ideal, must be the maximal solvable ideal.

If $L = 0$, then L is semisimple. For any L, $L/\operatorname{Rad} L$ is semisimple, since any solvable ideal in that space would contradict the fact that $\operatorname{Rad} L$ is the largest solvable ideal.

4. Nilpotency

We have seen the definition of solvability, and nilpotency is defined in a similar way, but using a different series than the derived series, called the lower central series.

Definition 4.1. The **normalizer** of a subalgebra K of a Lie algebra L is defined as the subalgebra of L $N_L(K) := \{x \in L | [x, K] \subseteq K\}$. This is a subalgebra because of the Jacobi Identity: Taking an arbitrary $x \in K$, y,$z \in N_L(K)$ we have:

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

Since the right-hand side is in K by definition of normalizer, the left hand side must be as well, meaning $[[y, z], K] \subset K$ so $[y, z] \in N_L(K)$. Basically, one can imagine the normalizer as the largest subalgebra of L to include K as an ideal. One can easily imagine a scenario where $N_L(K)$ is the smallest possible, i.e. K itself. If this happens K is called **self-normalizing**.

Definition 4.2. The **lower central series** is defined as follows: $L^0 = L$ and $L^{i+1} = [L, L^i]$ for $i \in \mathbb{N}$, $i > 0$.

Definition 4.3. A Lie algebra L is **nilpotent** if $L^n = 0$ for some $n \in \mathbb{N}$. Furthermore, x is **ad-nilpotent** if $\operatorname{ad} x^n(y) = 0$ for all $y \in L$.

Every nilpotent Lie algebra is solvable. This follows since the elements of the derived series is a subset of the corresponding element in the lower central series, hence if the lower central series reaches zero, the derived series reaches zero. The converse is not true, however. The most common example of a nilpotent Lie algebra is the space of strictly upper triangular matrices.

Definition 4.4. Let $Z(L)$ denote the **center** of L, which means all elements $z \in Z(L)$ satisfy $[z, k] = 0$ for all $k \in L$.

Three immediate lemmas follow:

Lemma 4.5.
a) All subspaces and homomorphisms of a nilpotent algebra are nilpotent.
b) If $L/Z(L)$ is nilpotent, then so is L.
c) If L is nilpotent and nonzero, then $Z(L)$ is nonzero.

Proof. The proof of the first is similar to the one for solvability, i.e. Lemma 3.8(a). The third follows from the fact that the last nonzero element in the lower central series must contain only elements that map to zero under the bracket with any other element in L, so that the next element in the lower central series is zero. Hence, that means the center must have nonzero elements.

For b), if $L/Z(L)$ is nilpotent, since $Z(L)$ is the 0 element in $L/Z(L)$, L^n is a subset of $Z(L)$. This means in the Lie algebra L itself, $L^{n+1} \subset 0$ by definition of $Z(L)$. □

4.1. **Properties of Ad-nilpotency.** Another way to express that a Lie algebra is nilpotent in terms of elements $x_1, x_2, ..., y$ in L is that $[x_1, [x_2, [\ldots, [x_n, y] \ldots]]]$ reaches 0 for all choices of x_n and y. A consequence of this is that $ad(x)^n(y) = 0$ for all $x, y \in L$. This is achieved by selecting x_i to be the same value x for all integers $i < n$.

Clearly, the above logic implies that if a Lie algebra is nilpotent, all of the elements of it are ad-nilpotent. In the next section, we will prove that the converse is also true.

5. Engel's Theorem

Theorem 5.1. *Engel's Theorem: A Lie algebra is nilpotent if and only if all its elements are ad-nilpotent.*

We already proved the"only if" direction, but the other direction is the meat of the theorem.

A useful example to keep in mind is strictly upper triangular matrices. These are nilpotent because successive elements of the lower central series "push" the triangle of nonzero values more and more upward until there are none left. Engel's theorem asserts that since the space of upper triangular matrices is nilpotent, each one is ad-nilpotent. This follows easily in the upper triangular case since for x strictly upper triangular, $\operatorname{ad} x^n$ is just the zero matrix and it sends everything to zero.

In fact, as we shall soon see, a critical part of the proof is finding an upper triangular representation in some basis for elements of L.

Lemma 5.2. *Let L be a Lie subalgebra in $gl(V)$ of nilpotent endomorphisms. If V is nonzero and finite-dimensional, then V must contain a nonzero element v where $l(v) = 0$ for all $l \in L$.*

Proof. We proceed using induction on $\dim L$. The case $\dim L = 1$ is obvious. All elements in L are in $span(l)$ for some element l. If L is nilpotent, some nonzero element in V is mapped to zero by a certain l in L, and since $L = span(l)$ we must have $l'(v) = 0$ for all $l' \in L$.

Take L a Lie algebra of dimension $\dim L > 1$. The first step is constructing an ideal with codimension one. Since any one-dimensional subalgebra is a proper subalgebra, a proper subalgebra exists. Suppose $K \subset L$ is a proper subalgebra.

We can conclude that K is, through applying *ad* on its elements, a Lie algebra of nilpotent transformations (since L consists of nilpotent elements) on L, and thus also on L/K. Since $\dim K < \dim L$, the induction hypothesis implies there is a nonzero vector, a certain $x + K \in L/K$, that K sends to zero. Since K is the zero space in L/K, we can add elements of K to x without affecting this identity. Translating this to L, we see that there is some $x \notin K$ such that $[k, x] \in K$ for all $k \in K$. Thus K is a proper subset of $N_L(K)$. If K is a maximal proper subalgebra of L, we will show that the codimension of K is one. From this, we can assert $N_L(K) = L$ since it has to be a subalgebra larger than K, of which the only option is L.

Armed with this assumption that K is maximal, we assert $\dim L/K = 1$. We proceed by contradiction: If $\dim L/K > 1$, we take H a one-dimensional subalgebra of L/K and consider the canonical map $L \to L/K$. Specifically we look at the preimage of $H + K$ which is a subalgebra that is strictly larger than K, but properly contained in L. This is a contradiction, because K is supposed to the the maximal subalgebra, so there cannot exist a subalgebra properly containing it properly contained by L. Thus $\dim L/K = 1$. Thus, $N_L(K) = L$ and K is an ideal of L of codimension one, as required.

Let $W := \{v \in V | Kv = 0\}$. By the induction hypothesis, W is nonzero. Additionally W is stable under L, because for $x \in L, y \in K, w \in W$:

$$yxw = xyw - [x, y]w = 0.$$

Since K is an ideal and $[x, y] \in K$, both terms on the right involve applying something in K to w, and are therefore zero.

Finally, let $z \in L - K$ and let it act on W (since W is stable under L, z is a map $W \to W$. It is nilpotent, so it has an eigenvector, some nonzero $v \in W$ for which $zv = 0$. So anything in $L - K$ as well as anything in K is mapped to zero, so the subtheorem is proven. □

From there Engel's theorem is evident.

Proof. Proof of Engel's theorem

Suppose L is a Lie algebra comprised of ad-nilpotent elements. By Lemma 5.2 applied to the Lie algebra $\operatorname{ad} L$ there exists a nonzero vector $x \in L$ for which $[L, x] = 0$ (since $\operatorname{ad} L = 0$). Then take $L/Z(L)$ where $Z(L)$ is the center. From there we can see that $Z(L)$ contains x so $L/Z(L)$ clearly has smaller dimension than L.

With the base case again being evident for $\dim L = 0$ we use induction to find that $L/Z(L)$ is nilpotent. By Lemma 4.4 (b), if $L/Z(L)$ is nilpotent then so is L, completing the proof. □

6. Lie's theorem

Lie's theorem has many parallels with Engel's theorem, but is about solvable subalgebras.

Definition 6.1. A **flag** is a chain of subspaces in V, $V_0 = 0, V_1 \subset V_2 \subset ... \subset V_n = V$ where $\dim V_i = i$.

Theorem 6.2. *Lie's Theorem: Let L be a solvable subalgebra of $gl(V)$, $\dim V < \infty$. Then L stabilizes some flag of V. In other words the matrices of L relative to some basis in V are upper triangular.*

It should be easy to see why the two statements of Lie's Theorem are equivalent. This is because we can take a basis of V where the first i vectors form a basis for V_i in the flag that it stabilizes. In the matrix representation, these will be elementary vectors. The fact that the matrix stabilizes the flag implies that for any subspace V_i in the flag, it sends that basis to something that can be represented using that basis and nothing more. Hence, the matrix in the first i columns outside of the first i rows (the part which maps the i basis vectors in V_i to other dimensions) must be zero, i.e. the matrix must be upper triangular when one looks at all i.

To prove the theorem one must invoke a theorem analogous to Lemma 5.2 as used in the proof of Engel's theorem:

Theorem 6.3. *Common Eigenvector Theorem: Let L be a solvable subalgebra of $gl(V)$, $\dim V < \infty, V \neq 0$. Then V contains some common eigenvector of elements in L.*

Proof. Similarly to the proof of Lemma 5.2, we use induction on $\dim L$ with the base case $\dim L = 1$ being trivial. Now assume the induction hypothesis that a common eigenvector exists when $\dim L < n$.

Let L be a subalgebra of $gl(V)$ with $\dim L = n$. First we find an ideal of codimension one. Because L is solvable, L properly includes $[L, L]$ and $L/[L, L]$ is abelian (any two elements in that space map into $[L, L]$ since any two elements in L map into $[L, L]$). So any subspace of $L/[L, L]$ is an ideal, as any subspace of an abelian Lie algebra is an ideal. This follows since any subspace contains zero and any operation in the Lie algebra results in zero. If we take a subspace of codimension one in $L/[L, L]$, which is an ideal, and when we look at its inverse image, we have a codimension one ideal $K \subset L$.

By the first statement in Lemma 3.8, we deduce that K is solvable. Since $\dim L > 1$, $\dim K > 0$, so K is clearly nonzero. By the induction hypothesis, a common eigenvector for all elements in K exists, and shall be called v. For $x \in K$, $xv = \lambda(x)v$, where $\lambda(x)$ is the eigenvalue corresponding to x and is by definition a linear function in x.

We can define by W the space of vectors w for which the property that holds for v holds. Formally,

$$W := \{w \in V |\, xw = \lambda(x)w \,\forall\, x \in K\}.$$

Since we already know that a nonzero element v is in W, we know that W is nonzero. We must show that L leaves the space W invariant. To prove this we take $x \in L$ and $w \in W$ and want to show that $xw \in W$. By definition of W, it suffices to show that for an arbitrary $y \in K$, $yxw = \lambda(y)xw$. However, by the definition of bracket operator,

$$yxw = xyw - [x, y]w = \lambda(y)xw - \lambda([x, y])w. \tag{6.4}$$

We thus only need to show that $\lambda([x, y]) = 0$ for all $y \in K$. Define n to be the smallest natural number where w, xw, x^2w, ... , $x^n w$ are linearly dependent. Define by W_i the subset of V spanned by w, xw, x^2w, ... , $x^{i-1}w$. By definition, $\dim W_n = n$ and we set $W_{n+i} = W_n$ for $i \in \mathbb{N}$. By definition of W_n, x leaves W_n invariant. The transformation y leaves W_i invariant because plugging $y = y$ and $x = x^j$, $j < i$ into 6.4, we have $yx^j w = \lambda(y)x^j w - \lambda([x^j, y])w$, both of which are in W_i because one is a multiple of w and one is a multiple of $x^j w$. And since y preserves the basis, it must preserve the overall space.

We claim that y is an upper triangular matrix in the basis W_n. We can prove this by proving the congruence

$$yx^i w \equiv \lambda(y)x^i w \mod W_i$$

for all i. When one sees a modulus one may replace it with $yx^i w = \lambda(y)x^i w + t$, $t \in W_i$. We prove the equivalence relation by induction (once again the base case is left to be checked easily):

$$yx^i w = yx^{i-1}xw = xyx^{i-1}w - [x,y]x^{i-1}w.$$

By induction, $yx^{i-1}w = \lambda(y)x^{i-1}w + t$, $t \in W_{i-1}$. We substitute into the first term of the rightmost expression:

$$xyx^{i-1}w - [x,y]x^{i-1}w = x\lambda(y)x^{i-1}w + t - [x,y]x^{i-1}w.$$

And since $[x,y] \in K$ by K being an ideal, the expression on the right is an element in W_i, so the congruence holds for i. By induction the congruence holds for all i.

Now we are ready to prove that y is a upper triangular matrix in that basis. We do so by examining what y does to each W_i:

$$y(w) = \lambda(y)w.$$

By the above equivalence,

$$y(xw) = \lambda(y)xw + aw$$

and generally

$$y(x_i w) = \lambda(y)x_i w + a(w + xw + \ldots + x_{i-1}w).$$

Since the rightmost part is a basis of W_i. So taking w, xw, ... as a basis,

$$y = \begin{pmatrix} \lambda(y) & a & \ldots \\ 0 & \lambda(y) & a \\ 0 & 0 & \lambda(y) \end{pmatrix}$$

constructing the matrix using the above equations for y of $x_i w$. Clearly the trace of y is equal to $\lambda(y) * n$. If an element $k \in K$ equals $[x,y]$, since x and y both stabilize W_n it is the commutator of two endomorphisms in W_n so its trace is zero. Hence, $\lambda(k) = 0$.

Finally, we construct the common eigenvector. Let $L = K + cz$ where $c \in \mathbb{C}$. By the last step, z stabilizes W. So $z|_W : W \to W$ has an eigenvector in W. This eigenvector is a common eigenvector of elements in L since it is both an eigenvector of elements in K and elements of form cz. Therefore, we are done. □

Proof. Proof of Lie's theorem: It follows from Theorem 6.3. Once again we use induction but this time in the dimension of V. For zero dimensions there is nothing to check (and the inductive step will still work).

The inductive step is to choose a basis for V with the common eigenvector v as $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ . \\ . \\ . \end{pmatrix}$ and then look at the matrix of any transformation in that basis. Evidently, it must be of form $\begin{pmatrix} \lambda(x) & ... \\ 0 & ... \\ 0 & ... \\ 0 & ... \\ . & ... \\ . & ... \\ . & ... \end{pmatrix}$ for the eigenvector v to actually be an eigenvector. By induction on the dimension of the space V/v, the rest of the matrix that operates on this smaller dimensional space is upper triangular (or nothing, in the case of $\dim V = 1$, justifying the base case). This proves the upper triangular matrix statement of Lie's Theorem. □

7. Conclusion

There are further results along this path, a major one of which is Cartan's criterion, which depends on constructing the Killing Form:

Definition 7.1. The **Killing form** is an operation, specifically a symmetric ad-invariant bilinear form, given by $K(x, y) := Tr(ad(x) \circ ad(y))$.

The Killing form is bilinear. Bilinearity is easy to see since the matrix of an adjoint representation of a multiple or sum is a scalar multiple or sum of the matrices of the individual elements. This of course stems from the fact that the bracket itself is bilinear. It is also symmetric, because the trace of two transformation matrices commutes with multiplication. Additionally, it is ad invariant, which means $K(x, [y, z]) = K([x, y], z)$. This is proved algebraically and uses the fact that $\operatorname{ad} x$ is a Lie algebra homomorphism and the fact that $Tr(AB) = Tr(BA)$:

$$\begin{aligned} K([x,y],z) - K(x,[y,z]) &= Tr(\operatorname{ad}[x,y] \circ \operatorname{ad} z - \operatorname{ad} x \circ \operatorname{ad}[y,z]) \\ &= Tr([\operatorname{ad} x, \operatorname{ad} y] \circ \operatorname{ad} z - \operatorname{ad} x \circ [\operatorname{ad} y, \operatorname{ad} z]) \\ &= Tr(\operatorname{ad} x \operatorname{ad} y \operatorname{ad} z - \operatorname{ad} y \operatorname{ad} x \operatorname{ad} z - \operatorname{ad} x \operatorname{ad} y \operatorname{ad} z \\ &\quad + \operatorname{ad} x \operatorname{ad} z \operatorname{ad} y) \\ &= Tr(\operatorname{ad} x \operatorname{ad} z \operatorname{ad} y - \operatorname{ad} y \operatorname{ad} x \operatorname{ad} z) \\ &= 0. \end{aligned}$$

Thus, the Killing form is an ad-invariant inner product. This definition allows us to state a powerful criterion for whether a finite-dimensional linear lie algebra is solvable.

Theorem 7.2. ***Cartan's Criterion***

Let V be finite-dimensional and L a sub-Lie algebra of $gl(V)$. Suppose $Tr(xy) = 0$ for all $x \in [L, L], y \in L$. Then L is solvable. In other words a Lie algebra is solvable if its Killing form is 0.

Acknowledgments

Thank you to Professor Peter May for organizing the program, Daniil Rudenko for giving the Apprentice Course lecture, and Katie Gallagher for mentoring me and revising this paper. Your guidance helped me traverse a previously foreign subject, and made it possible for me to write this paper.

References

[1] James E Humphreys. Introduction to Lie algebras and Representation Theory. https://www.math.uci.edu/ brusso/humphreys.pdf
[2] B. Zweibach. Notes in Lie algebras. https://home.gwu.edu/ hgrie/lectures/math-methods-script/Zwiebach-MIT8_324F10_lie_algebras3.pdf
[3] Paniz Imani. Introduction to Lie algebras. https://www.math.uni-hamburg.de/home/wockel/teaching/data/cohomology_of_lie_algebras_imani.pdf
[4] Alexei Skorobogatov. Lie algebras. https://www.ma.imperial.ac.uk/ anskor/LIE.PDF
[5] What is a Lie Group? https://aimath.org/E8/liegroup.html