CLASSICAL THEOREMS IN LIE ALGEBRA REPRESENTATION THEORY
—
A BEGINNER'S APPROACH.

Andrew P. Whitman, S.J.
Vatican Observatory
Vatican City

with

John A. Lutts (editor)
University of Massachusetts/Boston

2013

****

REAL AND COMPLEX REPRESENTATIONS OF
SEMISIMPLE REAL LIE ALGEBRAS

Andrew P. Whitman, S.J.
Vatican Observatory
V-00120 Vatican City

****

If any reader finds an error in this work (and there are some that have been left there, in part, to challenge you and the known ones are marked off by lines of +'s) or if anyone wishes to suggest improvements in the work's presentation) the author and the editor would be most appreciative of your help. They may be contacted via email, respectively, at

apwsj1951@gmail.com

and

john.lutts@umb.edu

****

Table of Contents

BIBLIOGRAPHY

[B] Bourbaki, N. : *Groupes et algèbres de Lie*, Springer-Verlag, 2007

[FH] Fulton, W. and Harris, J. : *Representation Theory: A First Course*, Springer, 1991

[H] Helgason, S. : *Differential Geometry, Lie Groups and Symmetric Spaces*, Academic Press, 1978.

[J] Jacobson, N. : *Lie Algebras*, Interscience, 1962

[K] Knapp, A. W. : *Lie Groups Beyond An Introduction*, 2nd ed, Birkhauser, 2002

[N] Nomizu, K. and Kobayashi, S. : *Foundations of Differential Geometry Vols. I. and II* Wiley Classics Library, 1966

[R] Ruelle, D. : *The Mathematician's Brain*, Princeton Univ. Press, 2007

1.1 Comments to the Reader – Why These Notes?

I was the first graduate student of Prof. Katsumi Nomizu back in 1961. He and Prof. Soshichi Kobayashi had just finished their wonderful book: *Foundations of Differential Geometry* and for my thesis I was immersed in Connection Theory. My career over the next 35 years, however, was essentially a career in undergraduate teaching . And my love for mathematics grew more and more over these years. Besides the book of Kobayashi and Nomizu I was constantly referring to the remarkable book of Helgason: *Differential Geometry, Lie Groups and Symmetric Spaces.* When I reached 70 years, I knew my days of teaching undergraduates were coming to an end. Back in 1981 I took a sabbatical at the Vatican Observatory, and I remained in contact with that remarkable group of observational and theoretical astronomers. Thus, when I retired from teaching it was natural to become one of the staff at the Vatican Observatory. That was 16 years ago, and in that period I began to write about Lie Algebras.

When I was in the first year of graduate studies, I can remember my attempts to read some standard treatments of certain areas of mathematics. The word "read", of course, took on a completely different significance. I can remember spending hours trying to understand just one page. Too much background was presumed by the author [which of course he/she had a perfect right to assume in the context in which he/she was writing] which background I simply did not yet have. I suppose one can say that this has been the story of my mathematical career — filling in backgrounds in many areas. However I had an ideal in mind: could not I fill in the background for the reader as I developed this area of mathematics? Mathematics is exciting, full of pleasure, and very satisfying. Indeed it is one of the most truly satisfying experiences of our human condition. And I wanted to expose in this manner this beautiful episode that the title describes.

Of course, I cannot start from the "beginning ", wherever and whatever that elusive concept may mean. And the scourge of verbosity is always lurking around ominously. Indeed Ruelle [R] identifies parsimony as one of the necessary traits of a mathematician. But there are many layers in that word "mathematician", and I do hope that the layer on which I am focusing I will be generous with my parsimony for my reader.

I basically studied three books: Jacobson [J], Fulton and Harris [F&H], and Knapp [K] — giant and classical treatises and textbooks on Lie Algebras. (I also dipped a little into Bourbaki [B]). While immersing myself in these books, I began to write the following notes. My main purpose was to be as explicit as I could be. This led to a verbosity which is intolerable in advanced level mathematics textbooks. Phrases such as: "it is easily shown" or "this result follows easily from" masked many difficult statements. And thus I started "filling in these gaps" in order better to understand Lie Algebras, and for me it was the only way I was to understand the material. Thus if anyone who already understands this material should start reading these notes, he/she can easily skip them. But maybe many students would appreciate these expanded explanations and analyses. Certainly they were very helpful to me.

And thus these notes for a good while just remained as something very personal to me. But along came the internet and it was suggested by my colleagues at the Vatican Observatory that I make them available on that medium. Of course, no editor or reviewer would tolerate such verbosity. However I asked a mathematician friend of mine, Jack Lutts, to join me in this project. His career was very similar to mine. He taught all his professional life at a not a very high level institution, where many students were adults holding down jobs. And he did a good degree at the University of Pennsylvania under C.T. Yang. His primary jobs in our collaboration have been to be a proof reader and an editor. Thus he has relentlessly checked the accuracy of my mathematics, making corrections where needed, and often suggesting rewordings when my phraseology was not clear or was too verbose or ungrammatical. Lutts adds that he has been a learner as well, rising to frequent challenges to review what he once knew better than he does now and discovering relationships that he never had a chance to learn during his math teaching career.

Now with our endeavors on the Web Page of the Vatican Observatory the whole world can access them [but not change them] and where struggling young mathematicians might prosper with the verbosity of the exposition of these notes. Also we are certain that any mistakes or ambiguities in our exposition and any gaps or errorts in any proofs would be eagerly pointed out. and we welcome any suggestions from our readers. In a sense we are putting out these notes for anyone who wishes to review and suggest clarifications and emendations.

But our hope is that we can lead a reader to savor a rather substantial piece of mathematics with the same pleasure that we had while thinking through and writing this exposition. And we hope that assuming a substantial course in undergraduate Linear Algebra will not be too onerous for our readers. [Of course, we will make explicit just what facts and results we will be using from Linear Algebra, but we will not attempt to verify or justify them. We do, however, give some brief explanations of the terms we use in several appendices located at the end of this work.]

Thus we welcome our readers to many hours of pleasurable mathematics.

2.1 Our starting point: What is a Lie algebra?

The Structure of a Lie Algebra.

The structure of a Lie algebra gives a set $\hat{g}$ three operations:

(1) A binary operation of addition:

$$\begin{array}{ccc} \hat{g} \times \hat{g} & \longrightarrow & \hat{g} \\ (a,b) & \longmapsto & a+b \end{array}$$

with the properties that give $\hat{g}$ the structure of an abelian group. [The identity of the group is, of course, written as 0.]

(2) A binary multiplication operation called a *bracket product*:

$$\begin{array}{ccc} \hat{g} \times \hat{g} & \longrightarrow & \hat{g} \\ (a,b) & \longmapsto & [a,b] \end{array}$$

that is anticommutative

$$[b,a] = -[a,b]$$

and satisfies the Jacobi identity:

$$[a,[b,c]] + [c,[a,b]] + [b,[c,a]] = 0$$

[Thus a Lie algebra is non-associative.]

(3) A scalar multiplication operation:

$$\begin{array}{ccc} \mathbb{F} \times \hat{g} & \longrightarrow & \hat{g} \\ (\alpha,c) & \longmapsto & \alpha c \end{array}$$

where $\mathbb{F}$ is a field.

Now, these three operations are tied into one another in the following manner. Bracket multiplication distributes over addition on the left and on the right [i.e., it is bilinear with respect to addition], giving us the structure of a non-commutative, non-associative ring. Thus

$$[a, b + c] = [a, b] + [a, c] \quad \textit{and} \quad [a + b, c] = [a, c] + [b, c]$$

Scalar multiplication combines with addition to give the structure of a linear space over $\mathbb{F}$. [Several things should be pointed out here. First, the only fields of scalars that we are interested in will be the real numbers and the complex numbers. Second, it should be noted out that we are interested in only finite dimensional linear spaces over these fields.. These limiting conditions are very important and our readers, we hope, will forgive us if we often remind them of these points.] Scalar multiplication is bilinear with respect to bracket multiplication, giving to the entire structure that of an algebra. Thus

$$\alpha[a, b] = [\alpha a, b] = [a, \alpha b]$$

The structure that we will be primarily interested in is that of a semisimple real Lie algebra. However the concept of "semisimple" does not depend on the field of scalars. Thus for real Lie algebras and complex Lie algebras we have the same definition. It depends on the concept of the solvable radical of a Lie algebra. Thus we are led to the notion of a solvable Lie algebra, and along with this notion, to that of a nilpotent Lie algebra. These depend on two natural series of subspaces of any Lie algebra — the derived series and the lower central series. All these terms will be defined and exemplified in what follows.

2.2 The Derived Series and the Lower Central Series

2.2.1 Ideals. Before we identify these objects, however, we need to define some other important objects of a Lie algebra. The symbol $[\hat{s}, \hat{t}]$ means the linear space generated by taking all possible bracket products $[a, b]$, $a \in \hat{s}$ and $b \in \hat{t}$, where $\hat{s}$ and $\hat{t}$ are subspaces of $\hat{g}$. A *subalgebra* is a subspace $\hat{s}$ in which the brackets close, i.e., where $[\hat{s}, \hat{s}] \subset \hat{s}$. If $\hat{s}$ is a subspace of $\hat{g}$, then $\hat{s}$ is an *ideal* if $[\hat{s}, \hat{g}] \subset \hat{s}$. [Thus every ideal is also a *subalgebra*.] A Lie algebra is *abelian* if $[\hat{g}, \hat{g}] = 0$, that is, if all brackets in $\hat{g}$ are 0. The *center* $\hat{z}$ of a Lie algebra $\hat{g}$ is the subspace $\hat{z}$ such that $[\hat{z}, \hat{g}] = 0$. Since $0 \in \hat{z}$, this means that the center is an ideal. We remark that every Lie algebra $\hat{g}$ has two ideals, which are called *improper ideals*: $\hat{g}$ itself (since $[\hat{g}, \hat{g}] \subset \hat{g}$), and the subspace 0 [since $[0, \hat{g}] = 0$, as is true in any ring].

Now given a Lie algebra $\hat{g}$, we define the *derived series* as:

$$\begin{aligned} D^0\hat{g} &:= \hat{g} \\ D^1\hat{g} &:= [\hat{g}, \hat{g}] \\ D^2\hat{g} &:= [D^1\hat{g}, D^1\hat{g}] \end{aligned}$$

$$
\begin{gathered}
D^3\hat{g} := [D^2\hat{g}, D^2\hat{g}] \\
\cdots \\
D^{k+1}\hat{g} := [D^k\hat{g}, D^k\hat{g}] \\
\cdots
\end{gathered}
$$

We define the *lower central series* as

$$
\begin{gathered}
C^0\hat{g} := \hat{g} \\
C^1\hat{g} := [\hat{g}, \hat{g}] = D^1\hat{g} \\
C^2\hat{g} := [C^1\hat{g}, \hat{g}] \\
C^3\hat{g} := [C^2\hat{g}, \hat{g}] \\
\cdots \\
C^{k+1}\hat{g} := [C^k\hat{g}, \hat{g}] \\
\cdots
\end{gathered}
$$

2.2.2 The Lower Central Series. We should make some remarks at this point. The lower central series is just the process of repeating the bracket multiplication in $\hat{g}$. Thus $C^3\hat{g} = [C^2\hat{g}, \hat{g}] = [[[\hat{g}, \hat{g}], \hat{g}], \hat{g}]$ means that we are just taking all expressions generated by three products from $\hat{g}$: $[[[a_0, a_1], a_2], a_3]$ for a_0, a_1, a_2, a_3 in $\hat{g}$. We emphasize that, because of a lack of associativity, we have chosen to perform these multiplications by multiplying the next element always on the right. We remark that we can define an analogous series of multiplications in any associative algebra. However, as we shall see, the non-associativity of a Lie algebra gives in the derived series special information which is not part of an analogous concept in associative algebras. More about this later.

Also, it does not matter if we would define the lower central series by bracketing $\hat{g}$ on the right (as above) or by bracketing $\hat{g}$ on the left in the following manner:

$$
\begin{gathered}
C^0\hat{g} := \hat{g} \\
C^1\hat{g} := [\hat{g}, \hat{g}] = D^1\hat{g} \\
C^2\hat{g} := [\hat{g}, C^1\hat{g}] \\
C^3\hat{g} := [\hat{g}, C^2\hat{g}] \\
\cdots \\
C^{k+1}\hat{g} := [\hat{g}, C^k\hat{g}] \\
\cdots
\end{gathered}
$$

Since we have anticommutativity in a Lie algebra, i.e., since $[b, a] = -[a, b]$, and since we are constructing linear spaces, which means that every element and its negative must appear in the linear space, both sets of products given above give the same elements when all the elements generated are collected into a set. Thus there is no ambiguity in the definition of the symbol $C^{k+1}\hat{g} = [C^k\hat{g}, \hat{g}] = [\hat{g}, C^k\hat{g}]$.

It is immediate from induction that $C^k\hat{g}$ is an ideal for all $k >= 0$. We have $C^0\hat{g} = \hat{g}$, and we know that $\hat{g}$ is an ideal. Let us assume that $C^{k-1}\hat{g}$ is an ideal. Then $[C^k\hat{g}, \hat{g}] = [[C^{k-1}\hat{g}, \hat{g}], \hat{g}]$. But by assumption $C^{k-1}\hat{g}$ is an ideal. Thus $[[C^{k-1}\hat{g}, \hat{g}], \hat{g}] \subset [C^{k-1}\hat{g}, \hat{g}] = C^k\hat{g}$. We can conclude that $[C^k\hat{g}, \hat{g}] \subset C^k\hat{g}$, which says that $C^k\hat{g}$ is an ideal. We observe that this also means that $C^k\hat{g} \subset C^{k-1}\hat{g}$.

2.2.3 The Derived Series. It is also immediate that $D^1\hat{g}$ is an ideal since $D^1\hat{g} = C^1\hat{g}$. We also can assert that $D^k\hat{g}$ is an ideal for all $k >= 0$, but we need the Jacobi identity to prove this. (This indicates that the concept of the derived series is deeper and more fundamental than the concept of the lower central series. We shall see this theme develop in the following pages.) In fact we can prove that *if $\hat{s}$ is an ideal, then $[\hat{s}, \hat{s}]$ is also an ideal.* Thus we need to show that $[[\hat{s}, \hat{s}], \hat{g}] \subset [\hat{s}, \hat{s}]$. Now let $a_1, a_2 \in \hat{s}$ and $c \in \hat{g}$. We have $[[\hat{s}, \hat{s}], \hat{g}]$ generated by elements of the form $[[a_1, a_2], c]$. Using the Jacobi identity, we have $[[a_1, a_2], c] = [[a_1, c], a_2] + [[c, a_2], a_1]$. But $\hat{s}$ is an ideal. Thus $[a_1, c] \in \hat{s}$ and $[c, a_2] \in \hat{s}$. Thus $[[a_1, c], a_2] \in [\hat{s}, \hat{s}]$ and $[[c, a_2], a_1] \in [\hat{s}, \hat{s}]$. Since $[\hat{s}, \hat{s}]$ is a subspace, the sum of two elements in $[\hat{s}, \hat{s}]$ is also in $[\hat{s}, \hat{s}]$. We thus have $[[a_1, a_2], c] \in [\hat{s}, \hat{s}]$. We conclude that $[[\hat{s}, \hat{s}], \hat{g}] \subset [\hat{s}, \hat{s}]$, which says that $[\hat{s}, \hat{s}]$ is an ideal. Now let us assume that $D^{k-1}\hat{g}$ is an ideal. But $D^k\hat{g} = [D^{k-1}\hat{g}, D^{k-1}\hat{g}]$. By induction it is now immediate that $D^k\hat{g}$ is an ideal. We also observe that since every ideal is also a subalgebra, we have that $D^k\hat{g} \subset D^{k-1}\hat{g}$.

2.2.4 Solvable Lie Algebras. Nilpotent Lie Algebras. Now we can define a solvable Lie algebra and a nilpotent Lie algebra. A *solvable* Lie algebra $\hat{s}$ is one such that for some $k >= 0$, $D^k\hat{s} = 0$. Trivially 0 is a solvable Lie algebra since $D^0 0 = 0$. A *nilpotent* Lie algebra $\hat{n}$ is one such that for some $k >= 0$, $C^k\hat{n} = 0$. Trivially 0 is also a nilpotent Lie algebra since $C^0 0 = 0$.

More observations can be made. If $\hat{a} \neq 0$ is an abelian Lie algebra, then $\hat{a}$ is a nilpotent Lie algebra, since $C^1\hat{a} = [\hat{a}, \hat{a}] = 0$; but also $\hat{a}$ is solvable since $D^1\hat{a} = C^1\hat{a} = 0$. If $C^k\hat{g} = 0$ but $C^{k-1}\hat{g} \neq 0$, then $\hat{g}$ has non zero center, for $[C^{k-1}\hat{g}, \hat{g}] = C^k\hat{g} = 0$ means that $C^{k-1}\hat{g} \subset \hat{z}$, where $\hat{z}$ is the center of $\hat{g}$. Thus every nonzero nilpotent Lie algebra has a nonzero center. Also if $\hat{s} \neq 0$ is a solvable Lie algebra, then for some $k > 0$, $D^k\hat{s} = 0$ and $D^{k-1}\hat{s} \neq 0$. But $D^k\hat{s} = [D^{k-1}\hat{s}, D^{k-1}\hat{s}] = 0$ says that every nonzero solvable Lie algebra has an nonzero abelian ideal, which is nilpotent.

2.3 The Radical of a Lie Algebra We now want to prove a most important observation, that every Lie algebra $\hat{g}$ has a maximal solvable ideal

$\hat{r}$ in the sense that any other solvable ideal in $\hat{g}$ must be contained in $\hat{r}$.[Recall that we are treating only finite dimensional Lie algebras.] We begin this proof by letting $\hat{r}$ be a solvable ideal of maximum dimensionality, and by letting $\hat{s}$ be any other solvable ideal. We are then led to consider the linear space $\hat{r} + \hat{s}$.

2.3.1 The Sum and Intersection of Ideals are Ideals. Thus we consider two linear subspaces $\hat{r}$ and $\hat{s}$ of a Lie algebra $\hat{g}$, and their sum $\hat{r}+\hat{s}$. This, of course, means that we are taking the sum of every two elements a and b, a in $\hat{r}$ and b in $\hat{s}$. But also we want to take the bracket product of all these sums and generate a Lie subalgebra. This means that if a_1 and a_2 are in $\hat{r}$ and b_1 and b_2 are in $\hat{s}$, then $[a_1 + b_1, a_2 + b_2]$ is in $[\hat{r}, \hat{r}] + [\hat{s}, \hat{s}] + [\hat{r}, \hat{s}]$. Now if $\hat{r}$ and $\hat{s}$ are subalgebras, then we can affirm that $[a_1 + b_1, a_2 + b_2]$ is in $\hat{r}+\hat{s}+[\hat{r}, \hat{s}]$. Only if either $\hat{r}$ or $\hat{s}$ is an ideal can we affirm that $[a_1+b_1, a_2+b_2]$ is in $\hat{r} + \hat{s}$, which means that $[\hat{r} + \hat{s}, \hat{r} + \hat{s}] \subset \hat{r} + \hat{s}$, and thus under these conditions is $\hat{r} + \hat{s}$ a subalgebra.

But we can assert more, namely that if $\hat{r}$ and $\hat{s}$ are ideals, then this linear space $\hat{r} + \hat{s}$ is an ideal in $\hat{g}$. For let a be in $\hat{r}$; b be in $\hat{s}$; and c be in $\hat{g}$. Now $[a + b, c] = [a, c] + [b, c]$. But since $\hat{r}$ and $\hat{s}$ are ideals, $[a, c] \in \hat{r}$ and $[b, c] \in \hat{s}$, and thus $[a, c] + [b, c] \in \hat{r} + \hat{s}$, from which we can conclude that $\hat{r} + \hat{s}$ is an ideal. As a corollary, we can also assert that $\hat{r}$ and $\hat{s}$ are ideals in $\hat{r} + \hat{s}$, for $[\hat{r}, \hat{r}+\hat{s}] \subset [\hat{r}, \hat{g}] \subset \hat{r}$ since $\hat{r}$ is an ideal in $\hat{g}$; and likewise $[\hat{s}, \hat{r}+\hat{s}] \subset [\hat{s}, \hat{g}] \subset \hat{s}$ since $\hat{s}$ is an ideal in $\hat{g}$.

Now we want to show that $\hat{r} \cap \hat{s}$ is also an ideal in $\hat{g}$ if $\hat{r}$ and $\hat{s}$ are ideals in $\hat{g}$. It is obviously a subspace. For let a be in $\hat{r} \cap \hat{s}$. Thus a is in $\hat{r}$ and a is in $\hat{s}$. Now let c be in $\hat{g}$. Then $[a, c]$ is in $[\hat{r} \cap \hat{s}, \hat{g}]$. But $[a, c]$ is in $[\hat{r}, \hat{g}]$, and $[a, c]$ is in $[\hat{s}, \hat{g}]$. We conclude that $[a, c] \in [\hat{r}, \hat{g}] \cap [\hat{s}, \hat{g}]$. But because $\hat{r}$ and $\hat{s}$ are ideals, then $[a, c] \in \hat{r} \cap \hat{s}$. We conclude that $[\hat{r} \cap \hat{s}, \hat{g}] \subset \hat{r} \cap \hat{s}$, and thus that $\hat{r} \cap \hat{s}$ is an ideal in $\hat{g}$. As a corollary we conclude also that $[\hat{r} \cap \hat{s}, \hat{r}] \subset [\hat{r} \cap \hat{s}, \hat{g}] \subset \hat{r} \cap \hat{s}$, and $[\hat{r} \cap \hat{s}, \hat{s}] \subset [\hat{r} \cap \hat{s}, \hat{g}] \subset \hat{r} \cap \hat{s}$, which means that $\hat{r} \cap \hat{s}$ is an ideal in $\hat{r}$ and $\hat{r} \cap \hat{s}$ is an ideal in $\hat{s}$.

2.3.2 Homomorphisms of Lie Algebras and Quotient Lie Algebras.

We also need to examine homomorphisms between Lie algebras $\hat{g}$ and $\hat{h}$. Let $\phi : \hat{g} \longrightarrow \hat{h}$ be a linear map between $\hat{g}$ and $\hat{h}$. This map ϕ is also a *homomorphism of Lie algebras* $\hat{g}$ and $\hat{h}$ if for every a and b in $\hat{g}$, $\phi[a, b] = [\phi(a), \phi(b)]$, i.e., ϕ preserves brackets. The properties of the map ϕ — being a surjective map, an injective map, a bijective map — carry over to the usual terms of surjective homomorphism, injective homomorphism, and isomorphism. If the target space $\hat{h}$ is equal to the domain space $\hat{g}$, the vocabulary changes from homomorphism to endomorphism, and from isomorphism to automorphism.

Obviously the homomorphic image of a Lie algebra is a Lie subalgebra of the target Lie algebra. Equally obvious is that subalgebras map to subalgebras, and ideals of surjective homomorphisms map to ideals. If we have a homomorphism, then the kernel of the map is an ideal in the domain Lie algebra. It would be good to see this proved. Our map again is $\phi : \hat{g} \longrightarrow \hat{h}$. Now $\ker(\phi) = \{c \in \hat{g} | \phi(c) = 0\}$. The kernel of a linear map is always a subspace. Let c be in $\ker(\phi)$ and a be any element in $\hat{g}$. Then $\phi[c, a] = [\phi(c), \phi(a)] = [0, \phi(a)] = 0$. Thus $[c, a]$ is in $\ker(\phi)$, which says that $[\ker(\phi), \hat{g}] \subset \ker(\phi)$, and thus $\ker(\phi)$ is an ideal in $\hat{g}$.

Now let us consider a Lie algebra $\hat{g}$ and an ideal $\hat{s}$ in $\hat{g}$. We can form the quotient linear space $\hat{g}/\hat{s}$. We assert that we can define a Lie bracket in this quotient space if $\hat{s}$ is an ideal. We define this bracket as follows. Let $a + \hat{s}$ and $b + \hat{s}$ be two elements in $\hat{g}/\hat{s}$. Then $[a + \hat{s}, b + \hat{s}] := [a, b] + \hat{s}$. To verify the validity of this definition we choose arbitrary elements s_1 and s_2 in $\hat{s}$ and calculate $[a + s_1, b + s_2] = [a, b] + [a, s_2] + [s_1, b] + [s_1, s_2]$. Since $\hat{s}$ is an ideal, we know that $[a, s_2]$ is in $\hat{s}$, $[s_1, b]$ is in $\hat{s}$, and, of course, $[s_1, s_2]$ is in $\hat{s}$. Thus $[a + s_1, b + s_2]$ is in the coset $[a, b] + \hat{s}$. It is immediate that all the properties of a Lie algebra are valid in $\hat{g}/\hat{s}$ using this definition of bracket product in $\hat{g}/\hat{s}$.

Applying this definition to a surjective homomorphism $\phi : \hat{g} \longrightarrow \hat{h}$ of Lie algebras, we know that $\hat{g}/\ker(\phi)$ is a Lie algebra, and indeed is isomorphic to $\phi(\hat{g}) = \hat{h}$. This latter statement reflects at the level of Lie algebras the crucial dimension theorem of linear algebra:

$$\dim(\hat{g}) = \dim(\ker\phi) + \dim(\text{image}(\phi))$$

or

$$\dim(\hat{g}) - \dim(\ker\phi) = \dim(\text{image}(\phi))$$

giving immediately the isomorphism of the *Lie algebras*

$$\hat{g}/\ker(\phi) \cong \text{image}(\phi) = \hat{h},$$

where we are, of course, assuming ϕ to be a surjective homomorphism.

We can also assert that the homomorphic image of a solvable Lie algebra $\hat{g}$ is also solvable. This means that if we have the homomorphism $\phi : \hat{g} \longrightarrow image(\phi)$, we can assert that the image(ϕ) is solvable when $\hat{g}$ is solvable. Since $\hat{g}$ is solvable, we have a k such that $D^k\hat{g} \neq 0$ and $[D^k\hat{g}, D^k\hat{g}] = D^{k+1}\hat{g} = 0$.

Since ϕ is a homomorphism, a simple induction shows that for all l, $\phi(D^l\hat{g}) = D^l(\phi(\hat{g}))$. Thus $\phi(D^{k+1}\hat{g}) = D^{k+1}(\phi(\hat{g})) = 0$ for some $k >= 0$.

We remark that if $D^k\hat{g}$ is in the kernel of ϕ, this k may not be the smallest positive integer with the property that $D^k(\phi(\hat{g})) \neq 0$ and $D^{k+1}(\phi(\hat{g})) = 0$. But we do know that the derived series for $\phi(\hat{g})$ does arrive at 0, which is enough to affirm that $\phi(\hat{g})$ is solvable.

We now can return to the proof that every Lie algebra $\hat{g}$ has a maximal solvable ideal. [Recall that we are only treating finite dimensional Lie algebras.] It will be done in stages. We begin this proof by letting $\hat{r}$ be a solvable ideal of maximum dimensionality, and by letting $\hat{s}$ be any other solvable ideal. We are then led to consider the linear space $\hat{r} + \hat{s}$. We know that this space is Lie subalgebra of $\hat{g}$ and indeed an ideal in $\hat{g}$. We want to show that this ideal is also solvable.

Now we know that $\hat{r} \cap \hat{s}$ is also an ideal in $\hat{s}$, and by assumption $\hat{s}$ is a solvable ideal in $\hat{g}$, and thus $\hat{s}$ is a solvable Lie algebra. We thus have the Lie algebra $\hat{s}/\hat{r} \cap \hat{s}$ which is the homomorphic image of a solvable Lie algebra. Thus we conclude that $\hat{s}/\hat{r} \cap \hat{s}$ is a solvable Lie algebra.

We now use the other fundamental dimension theorem of finite dimensional linear spaces, which states

$$\dim(\hat{r} + \hat{s}) = \dim(\hat{r}) + \dim(\hat{s}) - \dim(\hat{r} \cap \hat{s})$$

or

$$\dim(\hat{r} + \hat{s}) - \dim(\hat{s}) = \dim(\hat{r}) - \dim(\hat{r} \cap \hat{s})$$

This translates immediately into the isomorphism theorem

$$(\hat{r} + \hat{s})/\hat{s} \cong \hat{r}/\hat{r} \cap \hat{s}$$

It is important to observe that these are isomorphic *Lie algebras* since $\hat{s}$ is an ideal in $\hat{r} + \hat{s}$ and $\hat{r} \cap \hat{s}$ and $\hat{s}$ an ideal in $\hat{r}$, and thus the quotient linear spaces are quotient Lie algebras. We can therefore conclude now that $(\hat{r} + \hat{s})/\hat{s}$ is a solvable Lie algebra since it is isomorphic to a solvable Lie algebra $\hat{r}/\hat{r} \cap \hat{s}$.

2.3.3 Homomorphism Theorem for Solvable Lie Algebras. We now assert that $\hat{r} + \hat{s}$ is also solvable. This will be true if we can show that if $\hat{l}$ is a Lie algebra which contains a solvable ideal $\hat{s}$ and $\hat{l}/\hat{s}$ is solvable, then $\hat{l}$ is solvable.

First, we have the homomorphism $\phi : \hat{l} \longrightarrow (\hat{l}/\hat{s})$ and since $\phi(\hat{l}) = (\hat{l}/\hat{s})$, and $\hat{l}/\hat{s}$ is solvable, we know that there exists a k such that $D^k(\hat{l}/\hat{s}) \neq 0$ and $D^{k+1}(\hat{l}/\hat{s}) = 0$. Thus $D^k(\phi(\hat{l})) \neq 0$ and $D^{k+1}(\phi(\hat{l})) = 0$.

This says that $\phi(D^{k+1}(\hat{l})) = 0$. Thus $D^{k+1}(\hat{l}) \subset \ker(\phi) = \hat{s}$. Knowing that $D^{k+1}(\hat{l}) = [D^k(\hat{l}), D^k(\hat{l})] \subset \ker(\phi) = \hat{s}$, we conclude that $[D^{k+1}(\hat{l}), D^{k+1}(\hat{l})] = D^{k+2}(\hat{l}) \subset [\hat{s}, \hat{s}] \subset \hat{s}$.

Since $\hat{s}$ is solvable, we know that there is a p such that $D^p(\hat{s}) \neq 0$ and $D^{p+1}(\hat{s}) = 0$. Now if $k \geq p$, then $D^{k+1}(\hat{l}) = 0$. If $k < p$, then $D^{p+1}(\hat{l}) \subset D^{p+1}(\hat{s}) = 0$. Thus, in either event, $\hat{l}$ is solvable.

Note that his proof is easy to conceptualize. Since the target space is solvable, this means that after some k,$D^k(\hat{l}) \subset \hat{s}$, since $\phi(D^k(*)) = D^k(\phi(*))$. Once the derived series for $\hat{l}$ is in $\hat{s}$, then since $\hat{s}$ is solvable, the derived series for $\hat{l}$ will arrive at 0 as well.

2.3.4 The Existence of the Radical of a Lie Algebra. Finally, we arrive at what we have been seeking. In $\hat{g}$ let $\hat{r}$ be the solvable ideal of maximum dimension. [Recall that since we are in the context of finite dimensional Lie algebras, and since 0 is a solvable ideal, this ideal always exists.] Now let $\hat{s}$ be any other solvable ideal. We have just shown that $\hat{r} + \hat{s}$ is also a solvable ideal. Thus the $\dim(\hat{r} + \hat{s}) \leq \dim(\hat{r})$. But $\hat{r} \subset (\hat{r} + \hat{s})$, and this means that $\hat{s} \subset \hat{r}$. And thus every Lie algebra of finite dimension contains a maximal solvable ideal which contains all other solvable ideals. This maximal solvable ideal is called the *radical* of the Lie algebra $\hat{g}$.

2.4 Some Remarks on Semisimple Lie Algebras (1)

We can now give the definition of a semisimple Lie algebra which we have been seeking. A *semisimple Lie algebra* is a Lie algebra whose radical is trivial, *i.e.*, one whose radical is 0.

There are some properties of a semisimple Lie algebra which are immediate from the definition. First, let us take an arbitrary finite dimensional Lie algebra $\hat{g}$. We know that $\hat{g}$ possesses a radical, which we denote by $\hat{r}$. Since this radical is an ideal, we can form the Lie algebra $\hat{g}/\hat{r}$. It is interesting that we can prove that this quotient Lie algebra is semisimple. Let us represent the homomorphism by $\phi : \hat{g} \longrightarrow \hat{g}/\hat{r}$. Now we take any solvable ideal $\hat{s}$ in $\hat{g}/\hat{r}$ and look at its pre-image $\phi^{-1}(\hat{s})$ in $\hat{g}$. We know that this pre-image is an ideal in $\hat{g}$. And by using the same reasoning as was used above, we can assert that this ideal is solvable. Its image $\hat{s} = \phi^{-1}(\hat{s})/\hat{r}$ is solvable by assumption. And since $\hat{r}$ is the radical of $\hat{g}$, $\hat{r}$ also is solvable. Thus again we have the situation where we have a Lie algebra $\hat{l}$ and a homomorphism of $\hat{l}$ onto a solvable Lie algebra, the kernel of which homomorphism is also solvable. As above we can conclude that the Lie algebra $\hat{l}$ is also solvable. Thus $\phi^{-1}(\hat{s})$ is a solvable Lie algebra of $\hat{g}$, and we can conclude that $\phi^{-1}(\hat{s}) \subset \hat{r}$, the radical, which is the kernel of the map ϕ. But this means that the image of $\phi^{-1}(\hat{s})$, which is $\hat{s}$, is equal to 0. Thus any solvable ideal in $\hat{g}/\hat{r}$ must be the 0 ideal, which means that $\hat{g}/\hat{r}$ is semisimple.

Thus we have the short exact sequence:

$$0 \longrightarrow \hat{r} \longrightarrow \hat{g} \longrightarrow \hat{g}/\hat{r} \longrightarrow 0$$

The question to ask now is: does this short exact sequence split? If it does, then there is an injective isomorphism of Lie algebras

$$\hat{g}/\hat{r} \longrightarrow \hat{k} \subset \hat{g}$$

[thus making $\hat{k}$ a semisimple Lie subalgebra of $\hat{g}$] such that $\hat{g}$ is a direct sum of $\hat{r}$ and $\hat{k}$.

$$\hat{g} = \hat{k} \oplus \hat{r}$$

There is a famous theorem of Levi that gives a positive answer to this question [See 2.16 for its proof.]. The subalgebra $\hat{k}$ is called a Levi factor of $\hat{g}$. [We might again remark that this theorem is only valid in the case that the field of scalars is of characteristic 0. But since we are only interested in the fields of real numbers and complex numbers, the conclusion is valid in our situation.] Here the direct sum is only that of linear spaces, not of Lie algebras. This says that $[\hat{k}, \hat{r}]$ is not necessarily 0. We describe this situation by saying the $\hat{g}$ is a *semi-direct product* of the Lie algebras $\hat{k}$ and $\hat{r}$.

In fact we can make the following observations. Since $\hat{r}$ is an ideal, we know that $[\hat{k}, \hat{r}] \subset \hat{r}$. But we also have $[\hat{k}, \hat{r}] \subset [\hat{g}, \hat{g}] = D^1\hat{g}$. Thus $[\hat{k}, \hat{r}] \subset D^1\hat{g} \cap \hat{r}$. It is interesting to remark at this point that we have the important identity $[\hat{g}, \hat{r}] = D^1\hat{g} \cap \hat{r}$, but its proof depends on the Levi Decomposition Theorem. Certainly $[\hat{g}, \hat{r}] \subset [\hat{g}, \hat{g}] = D^1\hat{g}$ and $[\hat{g}, \hat{r}] \subset \hat{r}$ since $\hat{r}$ is an ideal. Thus $[\hat{g}, \hat{r}] \subset D^1\hat{g} \cap \hat{r}$. Now using the Levi Decomposition Theorem, we let $c_1 = a_1 + b_1$ and $c_2 = a_2 + b_2$ be in $\hat{g}$, with a_1 and a_2 in $\hat{k}$ and b_1 and b_2 in $\hat{r}$. Then $[c_1, c_2] = [a_1 + a_2, b_1 + b_2]$ is in $D^1\hat{g}$. Also $[a_1 + b_1, a_2 + b_2] = [a_1, a_2] + [a_1, b_2] + [b_1, a_2] + [b_1, b_2]$ with $[a_1, a_2]$ in $\hat{k}$ and $[a_1, b_2] + [b_1, a_2] + [b_1, b_2]$ in $\hat{r}$, since $\hat{r}$ is an ideal. But by hypothesis $[a_1 + b_1, a_2 + b_2]$ is also in $\hat{r}$. Since we have a direct sum of linear spaces, this means that $[a_1, a_2] = 0$
+++
That $[a_1, a_2] = 0$ is questionable at present for the reason given does not hold. Do you see why? Can you fix the proof? Hint: see 2.16.2 on p. 200.
+++
and $[a_1, b_2] + [b_1, a_2] + [b_1, b_2]$ is in $[\hat{g}, \hat{r}]$. We conclude that $[\hat{g}, \hat{r}] = D^1\hat{g} \cap \hat{r}$. Of course, since $\hat{r}$ is an ideal, we know that $[\hat{g}, \hat{r}] \subset \hat{r}$, but the above shows exactly where $[\hat{g}, \hat{r}]$ lies in $\hat{r}$. [Of course, once again, this result is only valid for

Lie algebras of characteristic 0, since it depends on the Levi decomposition theorem.]

We do not want to present the proof of this theorem of Levi at the present moment. But we do wish to ask: How unique is this Levi factor? The answer is very interesting. Let $\hat{g} = \hat{k}_1 \oplus \hat{r}$ and $\hat{g} = \hat{k}_2 \oplus \hat{r}$ be any two Levi decompositions of $\hat{g}$. Then there exist a very special kind of automorphism A of $\hat{g}$ such that $A(\hat{k}_2) = \hat{k}_1$. And these automorphisms will play a large role in the theme of this exposition of the real representations of semisimple real Lie algebras. Again at this moment we do not want to present a proof of this statement. However it is imperative that we identify these automorphisms. And this takes us on an exploration of another trail in this trek through Lie algebras, namely, that of nilpotent Lie algebras.

2.5 Nilpotent Lie Algebras (1)

2.5.1 Nilpotent Lie Algebras Are Solvable Lie Algebras. Let us review and examine more carefully the information found in the lower central series of a Lie algebra. The first few spaces in this series are

$$\begin{gathered} C^0\hat{g} = \hat{g} \\ C^1\hat{g} = [C^0\hat{g}, \hat{g}] \\ C^2\hat{g} = [C^1\hat{g}, \hat{g}] \\ C^3\hat{g} = [C^2\hat{g}, \hat{g}] \end{gathered}$$

We observe immediately that $C^1\hat{g}$ is merely the space generated by all possible brackets of $\hat{g}$. Thus for any a_1 and a_2 in $\hat{g}$, $[a_1, a_2]$ is in $C^1\hat{g}$. And for any a_3 in $\hat{g}$, $[[a_1, a_2], a_3]$ is in $C^2\hat{g}$. Continuing we have for any a_4 in $\hat{g}$, $[[[a_1, a_2], a_3], a_4]$ is in $C^3\hat{g}$. Thus the lower central series just iterates the bracket product in the Lie algebra, but chooses it in a definite manner because we do not have an associative algebra (where no matter how one groups the products, the answer is the same). We see that, because of the manner in which we have defined the series, it begins the iteration from the left and proceeds to the right. Thus any element of the form $[[[\cdots[a_1, a_2], \cdots], a_k], a_{k+1}]$ is in $C^k\hat{g}$. Now if $\hat{g} \neq 0$ is nilpotent, we know that there exists a k such that $C^{k-1}\hat{g} \neq 0$ but $C^k\hat{g} = 0$. Thus a nilpotent Lie algebra has the amazing property that any Lie bracket with k factors bracketed in this manner is 0. Obviously this is a rather strong property.

We can now show an interesting relationship between the derived series and the lower central series. If we examine again the derived series for $\hat{g}$, we see

$$D^0\hat{g} = \hat{g}$$

$$D^1\hat{g} = [\hat{g}, \hat{g}]$$
$$D^2\hat{g} = [[\hat{g}, \hat{g}], [\hat{g}, \hat{g}]]$$
$$D^3\hat{g} = [[[\hat{g}, \hat{g}], [\hat{g}, \hat{g}]], [[\hat{g}, \hat{g}], [\hat{g}, \hat{g}]]]$$

Let us write out these products in terms of elements of the Lie algebra.

$$[a_1, a_2] \text{ is in } D^1\hat{g}$$
$$[[a_1, a_2], [a_3, a_4]] \text{ is in } D^2\hat{g}$$
$$[[[a_1, a_2], [a_3, a_4]], [[a_5, a_6], [a_7, a_8]]] \text{ is in } D^3\hat{g}$$

Now using the Jacobi identity, we have

$$[a_1, a_2] \text{ is in } D^1\hat{g} = C^1\hat{g}$$
$$[[a_1, a_2], [a_3, a_4]] = -[a_3, [[a_1, a_2], a_4]] - [a_4, [a_3, [a_1, a_2]] =$$
$$-[[[a_1, a_2], a_3], a_4] + [[[a_1, a_2], a_4], a_3] \text{ is in } D^2\hat{g} \subset C^3\hat{g}$$

If we try to analyze $D^3\hat{g}$ in this manner, the calculations become unwieldy, but we would like to be able to conclude that $D^k\hat{g} \subset C^{2^k-1}\hat{g}$. Let us look a little more closely at the calculations involved. $D^1\hat{g}$ involves one product with two factors. $D^2\hat{g}$ involves four factors, which means, using the Jacobi identity, it really is a product of four factors with three products, with the products starting from the left and continuing to the right, i.e., it is contained in $C^3\hat{g}$. Now $D^2\hat{g} = [D^1\hat{g}, D^1\hat{g}]$, and $D^1\hat{g} = C^1\hat{g}$. Thus $D^2\hat{g} = [C^1\hat{g}, C^1\hat{g}]$, and we can conclude that $[C^1\hat{g}, C^1\hat{g}] \subset C^3\hat{g}$, where $3 = 1 + 1 + 1 = 2^2 - 1$.

Examining $D^3\hat{g} = [D^2\hat{g}, D^2\hat{g}]$, we see that we have 8 factors $[8 = 2^3]$, and if by using the Jacobi identity we could rearrange them so that we would have 7 products starting from the left and continuing to the right, we would be in $C^7\hat{g}$. From what is said above we know that $D^2\hat{g} \subset C^3\hat{g}$, and thus $D^3\hat{g} \subset [C^3\hat{g}, C^3\hat{g}]$. Now if $[C^3\hat{g}, C^3\hat{g}] \subset C^7\hat{g}$, we see that our numbers are correct since $3 + 3 + 1 = 7 = 2^3 - 1$.

These observations lead us to prove that $[C^i\hat{g}, C^j\hat{g}] \subset C^{i+j+1}\hat{g}$ by induction. For the base case we know that for all i, $[C^i\hat{g}, C^0\hat{g}] = C^{i+0+1}\hat{g}$ by definition. We assume that for some $j \geq 0$ and for all i $[C^i\hat{g}, C^j\hat{g}] \subset C^{i+j+1}\hat{g}$, and we prove that $[C^i\hat{g}, C^{j+1}\hat{g}] \subset C^{i+j+1+1}\hat{g}$. Of course, we will use the Jacobi identity. $[C^i\hat{g}, C^{j+1}\hat{g}] = [C^i\hat{g}, [C^j\hat{g}, \hat{g}]] \subset [[C^i\hat{g}, C^j\hat{g}], \hat{g}]] + [C^j\hat{g}, [\hat{g}, C^i\hat{g}]] \subset [C^{i+j+1}\hat{g}, \hat{g}] + [C^j\hat{g}, [C^i\hat{g}, \hat{g}]] \subset C^{i+j+1+1}\hat{g} + [C^j\hat{g}, C^{i+1}\hat{g}] \subset C^{i+j+1+1}\hat{g} + C^{i+1+j+1}\hat{g} \subset C^{i+j+1+1}\hat{g}$.

To conclude we once more use induction. The base case is obvious. It says that $D^0 \subset C^0$ [which says $\hat{g} \subset \hat{g}$]. Then we assume that for some $k \geq 0$ that $D^k\hat{g} \subset C^{2^k-1}\hat{g}$, and therefore $D^{k+1}\hat{g} = [D^k\hat{g}, D^k\hat{g}] \subset [C^{2^k-1}\hat{g}, C^{2^k-1}\hat{g}] \subset C^{(2^k-1)+(2^k-1)+1}\hat{g} = C^{2(2^k)-1}\hat{g} = C^{2^{k+1}-1}\hat{g}$. This is our desired relationship.

With this information we can now conclude that if a Lie algebra is nilpotent, then it must also be solvable, for if we take down the lower central series to 0, then the derived series will also be pulled down to 0. (See 2.2.4 for the relevant definitions.) Conversely, however, a solvable Lie algebra is not necessarily nilpotent. We will soon have an example of this situation. We will see that for a given n, all upper triangular matrices with arbitrary diagonal elements form a solvable Lie subalgebra in the Lie algebra of all nxn matrices, but they are not nilpotent, since all of these upper triangular matrices which form a nilpotent Lie subalgebra must have a zero diagonal.

2.5.2 The Existence of a Maximal Nilpotent Ideal — The Nilradical. We can now also show the existence of a maximal nilpotent ideal for any finite dimensional Lie algebra. All that needs to be done is to show that the sum of two nilpotent ideals is also a nilpotent ideal. If we repeat the proof for the solvable Lie algebras, we can indeed show that the homomorphic image of a nilpotent Lie algebra $\hat{g}$ is nilpotent. Since $\hat{g}$ is nilpotent, we have an l such that $C^l\hat{g} \neq 0$ and $[C^l\hat{g}, \hat{g}] = C^{l+1}\hat{g} = 0$. Now we have the homomorphism $\phi : \hat{g} \longrightarrow image(\phi)$, and we want to assert that the image(ϕ) is nilpotent. The key fact in the proof for solvable Lie algebras was that $\phi(D^l\hat{g}) = D^l(\phi(\hat{g}))$. A simple induction indeed shows that $\phi(C^l\hat{g}) = C^l(\phi(\hat{g}))$. Thus $\phi(C^{l+1}\hat{g}) = C^{l+1}(\phi(\hat{g})) = 0$. -

However we do not have an isomorphism theorem for nilpotent Lie algebras. Analyzing the proof for the solvable case, we see that after some k, $C^k(\hat{l}) \subset \hat{s}$, since $\phi(C^k(*)) = C^k(\phi(*))$. Now nilpotency demands that we bracket $\hat{s}$ with $\hat{g}$, but we only know that $\hat{s}$ is nilpotent. This means that we would have to bracket $\hat{s}$ with $\hat{s}$ if we wanted to use the nilpotency of $\hat{s}$. And thus we cannot get information on $[\hat{s}, \hat{g}]$.

Indeed the example that was given above shows the phenomenon mentioned above. We have for a given n a homomorphism of all upper triangular matrices with arbitrary diagonal [a solvable Lie algebra] to the diagonal matrices [an abelian Lie algebra, thus a nilpotent Lie algebra] by moding out the upper triangular matrices with zero diagonal [a nilpotent Lie algebra]. Thus, this is a counterexample to an isomorphism theorem for nilpotent Lie algebras since the upper triangular matrices with arbitrary diagonal is a solvable Lie algebra but not a nilpotent one. Thus we are reduced to finding another method of proving the sum of two nilpotent ideals is also a nilpotent ideal.

We already know that the sum of two ideals is an ideal. Let us take some brackets and examine the developing pattern. We take $\hat{k}$ and $\hat{l}$ to be two nilpotent ideals and we want to show their sum $\hat{k} + \hat{l}$ is also nilpotent. We take a_i in $\hat{k}$ and b_j in $\hat{l}$. Then $[a_1 + b_1, a_2 + b_2]$ is in $[\hat{k} + \hat{l}, \hat{k} + \hat{l}] = C^1(\hat{k} + \hat{l}])$. Now

$$[a_1+b_1,a_2+b_2]=[a_1,a_2]+[a_1,b_2]+[b_1,a_2]+[b_1,b_2].$$

We see that $[a_1,a_2]$ is in $C^1\hat{k}$, and because $\hat{k}$ is an ideal, we have $[a_1,b_2]$ is in $C^0\hat{k}$. Likewise we have $[b_1,b_2]$ is in $C^1\hat{l}$, and because $\hat{l}$ is an ideal, we have $[b_1,a_2]$ is in $C^0\hat{l}$. Also

$$[[a_1+b_1,a_2+b_2],a_3+b_3]=[[a_1,a_2]+[a_1,b_2]+[b_1,a_2]+[b_1,b_2],a_3+b_3]=$$
$$[[a_1,a_2],a_3]+[[a_1,b_2],a_3]+[[b_1,a_2],a_3]+[[b_1,b_2],a_3]+$$
$$[[a_1,a_2],b_3]+[[a_1,b_2],b_3]+[[b_1,a_2].b_3]+[[b_1,b_2],b_3]$$

is in $[[\hat{k}+\hat{l},\hat{k}+\hat{l}],\hat{k}+\hat{l}]=C^2(\hat{k}+\hat{l}])$. We see that $[[a_1+b_1,a_2+b_2],a_3+b_3]$ has 8 brackets. We choose 4 of these brackets, $[[a_1,a_2],a_3]$, $[[a_1,b_2],a_3]$, $[[b_1,a_2],a_3]$ and $[[a_1,a_2],b_3]$. Each of these products has at least two a_i factors. Now because $\hat{k}$ is an ideal, we have immediately that $[[a_1,a_2],a_3]$, $[[a_1,b_2],a_3]$, $[[b_1,a_2],a_3]$ are in $C^1\hat{k}$. Also since $C^1\hat{k}$ is an ideal, we have $[[a_1,a_2],b_3]$ is in $C^1\hat{k}$. Thus all 4 of these terms are in $C^1\hat{k}$. Likewise the other 4 terms $[[b_1,b_2],a_3]$, $[[a_1,b_2],b_3]$, $[[b_1,a_2].b_3]$, $[[b_1,b_2],b_3]$ are in $C^1\hat{l}$. We take one more bracket in $\hat{k}+\hat{l}$, namely $[[[a_1+b_1,a_2+b_2],a_3+b_3],a_4+b_4]$, which is in $C^3(\hat{k}+\hat{l}])$. Now

$$[[[a_1+b_1,a_2+b_2],a_3+b_3],a_4+b_4]=$$
$$[[[a_1,a_2],a_3],a_4]+[[[a_1,b_2],a_3],a_4]+[[[b_1,a_2],a_3],a_4]+[[[b_1,b_2],a_3],a_4]+$$
$$[[[a_1,a_2],b_3],a_4]+[[[a_1,b_2],b_3],a_4]+[[[b_1,a_2].b_3],a_4]+[[[b_1,b_2],b_3],a_4]+$$
$$[[[a_1,a_2],a_3],b_4]+[[[a_1,b_2],a_3],b_4]+[[[b_1,a_2],a_3],b_4]+[[[b_1,b_2],a_3],b_4]+$$
$$[[[a_1,a_2],b_3],b_4]+[[[a_1,b_2],b_3],b_4]+[[[b_1,a_2],b_3],b_4]+[[[b_1,b_2],b_3],b_4]$$

Thus we see that we have 16 terms, 8 of which have 2 or a greater number of factors a_i, and 8 of which have 2 or a greater number of factors b_j. The 8 with a_i are:

$$[[[a_1,a_2],a_3],a_4],[[[a_1,a_2],a_3],b_4],[[[a_1,a_2],b_3],a_4],\ [[[a_1,b_2],a_3],a_4],$$
$$[[[b_1,a_2],a_3],a_4],[[[b_1,a_2],b_3],a_4],[[[a_1,b_2],b_3],a_4],[[[a_1,a_2],b_3],b_4]$$

[Some of the terms with 2 a_i's and 2 b_j's can be treated either in $\hat{k}$ or in $\hat{l}$.] Now $[[[a_1,a_2],a_3],a_4]$ is in $C^3\hat{k}$; $[[[a_1,b_2],a_3],a_4]$ is in $C^2\hat{k}$, using the ideal structure of $\hat{k}$; $[[[b_1,a_2],a_3],a_4]$ is in $C^2\hat{k}$; $[[[a_1,a_2],b_3],a_4]$ is in $C^2\hat{k}$, using the ideal structure of $C^1\hat{k}$; $[[[a_1,b_2],b_3],a_4]$ is in $C^1\hat{k}$; $[[[b_1,a_2],b_3],a_4]$ is in $C^1\hat{k}$; $[[[a_1,b_2],a_3],b_4]$ is in $C^1\hat{k}$; $[[[a_1,a_2],a_3],b_4]$ is in $C^2\hat{k}$, using the ideal structure of $C^2\hat{k}$; $[[[a_1,a_2],b_3],b_4]$ is in $C^1\hat{k}$. The pattern is easy to recognize. Any bracket, starting from the left and moving to the right [which order we have chosen, you will recall, since we do not have associativity], with two or more factors a_i in <u>any</u> position, is in $C^1\hat{k}$ if there are 2 a_i's; in $C^2\hat{k}$ if there are 3 a_i's; and is in $C^3\hat{k}$ if there are 4 a_i's. Likewise we obtain the same conclusion for the 8 terms which have 2 or greater number of factors b_j, except now they pertain to the ideal $\hat{l}$. Thus $C^3(\hat{k}+\hat{l}]) \subset C^1\hat{k}+C^1\hat{l}$ since $C^3\hat{k} \subset C^2\hat{k} \subset C^1\hat{k}$, and likewise for $\hat{l}$.

Now if we assume that we could do an induction on $C^{i+1}(\hat{k}+\hat{l}]) \subset C^{i-1}\hat{k}+C^{i-1}\hat{l}$, we are unfortunately led to a dead end. The induction step would be

$$\begin{gathered} C^{i+2}(\hat{k}+\hat{l}]) = \\ [C^{i+1}(\hat{k}+\hat{l}]),(\hat{k}+\hat{l})] = \\ [C^{i+1}(\hat{k}+\hat{l}]),\hat{k}]+[C^{i+1}(\hat{k}+\hat{l}]),\hat{l}] \subset \\ [C^{i-1}\hat{k}+C^{i-1}\hat{l},\hat{k}]+[C^{i-1}\hat{k}+C^{i-1}\hat{l},\hat{l}] \subset \\ [C^{i-1}\hat{k},\hat{k}]+[C^{i-1}\hat{l},\hat{k}]+[C^{i-1}\hat{k},\hat{l}]+[C^{i-1}\hat{l},\hat{l}] \subset \\ C^{i}\hat{k}+C^{i-1}\hat{l}+C^{i-1}\hat{k}+C^{i}\hat{l} \subset \\ C^{i-1}\hat{k}+C^{i-1}\hat{l} \end{gathered}$$

and we see that we do not get the desired conclusion that $C^{i+2}(\hat{k}+\hat{l}]) \subset C^{i}\hat{k}+C^{i}\hat{l}$.

But a more perceptive analysis of these two examples leads us to the solution. All we have to do is observe that any term that has i a_r's as factors can be shown, by using the ideal structure of $\hat{k}$ and $C^j\hat{k}$, to belong to $C^{i-1}\hat{k}$. Thus we see from the above example that

$$\begin{gathered} [[[a_1,a_2],a_3],a_4] \text{ is in } C^3\hat{k} \\ [[[a_1,a_2],a_3],b_4] \text{ is in } C^2\hat{k} \\ [[[a_1,a_2],b_3],a_4] \text{ is in } C^2\hat{k} \\ [[[a_1,b_2],a_3],a_4] \text{ is in } C^2\hat{k} \\ [[[b_1,a_2],a_3],a_4] \text{ is in } C^2\hat{k} \\ [[[a_1,b_2],b_3],a_4] \text{ is in } C^1\hat{k} \\ [[[a_1,b_2],a_3],b_4] \text{ is in } C^1\hat{k} \\ [[[a_1,a_2],b_3],b_4] \text{ is in } C^1\hat{k} \end{gathered}$$

With these patterns before us we can see the general situation. Thus let us start with $C^i(\hat{k}+\hat{l})$ and i even; and $C^j(\hat{k}+\hat{l})$ with j odd. The $i-$th product will have $i+1$ factors, an odd number of factors; and the $j-$th product will have $j+1$ factors, an even number of factors. If a_r is in $\hat{k}$, and b_s is in $\hat{l}$, then $C^i(\hat{k}+\hat{l})$ will have 2^{i+1} terms and $C^j(\hat{k}+\hat{l})$ will have 2^{j+1} terms, each term of which will be a string of a_r's and b_s's as factors in arbitrary order. For $C^i(\hat{k}+\hat{l})$, $\frac{1}{2}(2^{i+1}) = 2^i$ of the terms, each with $i+1$ factors, will have $\frac{1}{2}(i)+1$ or more a_r's and $\frac{1}{2}(i)$ or less b_s's; or the contrary, $\frac{1}{2}(i)+1$ or more b_s's and $\frac{1}{2}(i)$ or less a_r's. For $C^j(\hat{k}+\hat{l})$, $\frac{1}{2}(2^{j+1}) = 2^j$ of the terms, each with $j+1$ factors, with $\frac{1}{2}(j+1)$ or more a_r's and $\frac{1}{2}(j+1)$ or less b_s's; or vice versa, $\frac{1}{2}(j+1)$ or more b_s's and $\frac{1}{2}(j+1)$ or less a_r's.

Thus, in the case where i is even, the first separation gives each term at least $(\frac{1}{2}(i)+1)$ a_r's, with each other term increasing the number of a_r's until all $i+1$ factors have only a_r's. In the second separation each term has at least $(\frac{1}{2}(i)+1)$ b_s's, with each other term increasing the number of b_s's until all $i+1$ factors have only b_s's. [In the example above, $C^2(\hat{k}+\hat{l})$ with $i=2$, each product has $2+1=3$ factors, and we have a total of $2^{2+1}=8$ terms. The first separation contained 4 terms

$$[[a_1,a_2],b_3],[[a_1,b_2],a_3],[[b_1,a_2],a_3],[[a_1,a_2],a_3]$$

while the second separation contained the other 4 terms

$$[[b_1,b_2],a_3],[[b_1,a_2].b_3],[[a_1,b_2],b_3],[[b_1,b_2],b_3].$$

In the first separation each term contains at least $\frac{1}{2}(i)+1=\frac{1}{2}(2)+1=2$ a_r's and $\frac{1}{2}(2)=1$ or less b_s's; and it continues adding a_r's until all the factors are a_r's, that is, until all 3 factors are a_r's. In the second separation the a_r's and the b_s's exchange roles.]

For the case $C^j(\hat{k}+\hat{l})$ where j is odd, the first separation gives each term at least $\frac{1}{2}(j+1)$ a_r's, with each other term increasing the number of a_r's until all $j+1$ factors have only a_r's. In the second separation each term has at least $\frac{1}{2}(j+1)$ b_s's with each other term increasing the number of b_s's until all $j+1$ factors have only b_s's. [In the example above, $C^3(\hat{k}+\hat{l})$ with $j=3$, each product has $3+1=4$ factors, and we have a total of $2^{3+1}=16$ terms. The first separation contains 8 terms

$$[[[a_1,a_2],b_3],b_4],[[[a_1,b_2],b_3],a_4],[[[b_1,a_2],b_3],a_4],$$
$$[[[a_1,a_2],a_3],b_4],[[[a_1,a_2],b_3],a_4,[[[a_1,b_2],a_3],a_4],[[[b_1,a_2],a_3],a_4],$$
$$[[[a_1,a_2],a_3],a_4]$$

while the second separation contains the other 8 terms

$$[[[b_1,b_2],a_3],a_4],[[[b_1,a_2],a_3],b_4],[[[a_1,b_2],a_3],b_4],$$
$$[[[b_1,b_2],b_3],a_4],[[[b_1,b_2],a_3],b_4],[[[b_1,a_2],b_3],b_4],[[[a_1,b_2],b_3],b_4],$$
$$[[[b_1,b_2],b_3],b_4]$$

In the first separation each term contains at least $\frac{1}{2}(j+1)=\frac{1}{2}(3+1)=2$ a_r's; and $\frac{1}{2}(3+1)=2$ or less b_s's; and continues adding a_r's until all the factors are a_r's, that is, until all 4 factors are a_r's. In the second separation the a_r's and the b_s's exchange roles. We might also add that when we have the same number of a_r's and b_s's, it makes no difference into what separation we place these terms.]

From these results it is immediate that $\hat{k}+\hat{l}$ is a nilpotent ideal if $\hat{k}$ and $\hat{l}$ are nilpotent ideals. In both even and odd cases we have $C^i(\hat{k}+\hat{l})\subset C^{j_1}(\hat{k})+$

$C^{j_2}(\hat{l})$, where $j_1 = j_2 = \frac{1}{2}(i) + 1$ in the even case; and $j_1 = j_2 = \frac{1}{2}(i+1)$ in the odd case. Thus all we have to do is take i sufficiently large so that both $C^{j_1}\hat{k}$ and $C^{j_2}\hat{l}$ are brought down to 0, which insures that $C^i(\hat{k} + \hat{l})$ will also be brought down to 0. Thus we can conclude that the sum of two nilpotent ideals is also a nilpotent ideal.

And now just as we have argued in the case of the existence of the radical — the maximal solvable ideal which contains any other solvable ideal — that any finite dimensional Lie algebra possesses, we can conclude to the existence of the *nilradical* — the maximal nilpotent ideal which contains any other nilpotent ideal — that any finite dimensional Lie algebra possesses.

2.6 Some First Remarks on Representations of Lie Algebras

We recall that we are seeking an automorphism A such that $A(\hat{k}_2) = \hat{k}_1$, where $\hat{k}_1$ and $\hat{k}_2$ are two Levi factors in the decomposition of an arbitrary Lie algebra $\hat{g}$ into $\hat{g} = \hat{k}_1 \oplus \hat{r}$ and $\hat{g} = \hat{k}_2 \oplus \hat{r}$, where $\hat{r}$ is the radical of $\hat{g}$. Later in our exposition we will show that A is obtained by integrating a nilpotent Lie algebra. [Now this process of integration in this case turns out to be an algebraic process and not an analytic process, and thus we remain in the context of algebra.] This would lead us into an exploration of nilpotent Lie algebras. But another reason for exploring nilpotent Lie algebras is that they are related to the concept of solvable Lie algebras, and this is the next topic that we will explore. But first we must examine the concept of a representation of a Lie algebra. Since the word *representation* appears in the title of this document, one would suspect that this concept is central to our exposition.

2.6.1 The Set of Endomorphisms of V: End(V); $\widehat{\mathbf{gl}}(\mathbf{v})$. Recall that we are now in the context of a finite dimensional linear space V over a field $\mathbb{F}$, which is either the field of real numbers $\mathbb{R}$ or the field of complex numbers $\mathbb{C}$. Let $End(V)$ stand for the set of all linear transformations of V, that is, the set of all endomorphisms of V. Thus an element ϕ of $End(V)$ is a function $\phi : V \longrightarrow V$ which preserves the addition and scalar multiplication in V:

$$\begin{gathered}
\phi : V \longrightarrow V \\
\phi : u \longmapsto \phi(u) \\
\phi : u + v \longmapsto \phi(u+v) = \phi(u) + \phi(v) \\
\phi : \alpha u \longmapsto \phi(\alpha u) = \alpha\phi(u)
\end{gathered}$$

Now $End(V)$ has the structure of an *associative algebra over* $\mathbb{F}$. In an associative algebra we again have a field $\mathbb{F}$ and a set with three operations:

addition, multiplication, and scalar multiplication. The operation of addition gives the set the structure of an abelian group; scalar multiplication combines with addition to give the set the structure of a linear space over $\mathbb{F}$; multiplication distributes over addition both on the left and on the right, giving the ring structure; and scalars are bilinear with respect to multiplication, giving the whole structure that of an algebra. Thus the set $End(V)$ has three operations: addition, multiplication, and scalar multiplication over $\mathbb{F}$. Since each endomorphism ϕ in $End(V)$ is a function, addition and scalar multiplication on $End(V)$ are pointwise addition and pointwise scalar multiplication of functions:

$$
\begin{gathered}
\phi_1 + \phi_2 : V \longrightarrow V \\
\phi_1 + \phi_2 : u \longmapsto (\phi_1 + \phi_2)(u) := \phi_1(u) + \phi_2(u) \\
\alpha\phi : V \longrightarrow V \\
\alpha\phi : u \longmapsto (\alpha\phi)(u) := \alpha\phi(u)
\end{gathered}
$$

It is trivial that these operations define an addition and a scalar multiplication on $End(V)$ over $\mathbb{F}$. The zero linear transformation is the zero of the addition operation; while the additive inverse is the negative of the linear transformation:

$$
\begin{gathered}
0 : V \longrightarrow V \\
0 : u \longmapsto 0(u) := 0 \\
-\phi : V \longrightarrow V \\
-\phi : u \longmapsto (-\phi)(u) := -(\phi(u))
\end{gathered}
$$

Again since an endomorphism is a function with the same domain and target sets, we define the multiplication operation in $End(V)$ as the composition of functions:

$$
\begin{gathered}
\phi_1\phi_2 : V \longrightarrow V \\
V \xrightarrow{\phi_2} V \xrightarrow{\phi_1} V \\
u \longmapsto \phi_2(u) \longmapsto \phi_1(\phi_2(u)) = (\phi_1 \circ \phi_2)(u) := (\phi_1\phi_2)(u)
\end{gathered}
$$

This multiplication distributes over addition on the left and on the right, giving the necessary ring property:

$$
\begin{gathered}
\phi_1(\phi_2 + \phi_3) = \phi_1\phi_2 + \phi_1\phi_3 \\
(\phi_1(\phi_2 + \phi_3))(u) = \phi_1((\phi_2 + \phi_3)(u)) = \phi_1(\phi_2(u) + \phi_3(u)) = \\
\phi_1(\phi_2(u)) + \phi_1(\phi_3(u)) = (\phi_1\phi_2)(u) + (\phi_1\phi_3)(u) = (\phi_1\phi_2 + \phi_1\phi_3))(u)
\end{gathered}
$$

Likewise

$$(\phi_1 + \phi_2)\phi_3 = \phi_1\phi_3 + \phi_2\phi_3$$

It is evident that this multiplication is bilinear with respect to the scalars and thus

$$\begin{gathered}\alpha(\phi_1\phi_2) = (\alpha\phi_1)\phi_2 = \phi_1(\alpha\phi_2) \\ (\alpha(\phi_1\phi_2))(u) = (\phi_1\phi_2)(\alpha u) = \phi_1(\phi_2(\alpha u)) \\ ((\alpha\phi_1)\phi_2)(u) = (\alpha\phi_1)(\phi_2(u)) = \phi_1(\alpha(\phi_2(u)) = \phi_1(\phi_2(\alpha u)) \\ (\phi_1(\alpha\phi_2))(u) = \phi_1((\alpha\phi_2)(u)) = \phi_1(\phi_2(\alpha u))\end{gathered}$$

Thus $End(V)$ exhibits the structure of an algebra over the field $\mathbb{F}$.

We observe that since we have chosen to write functions on the left of the element on which they are acting, the second listed function in a composition acts first while the first listed function acts second. But since we are just defining a binary operation of multiplication in $End(V)$, the notation is consistent. Note, however, that function composition is not a commutative operation.

Since composition of functions is an associative operation, our multiplication also associates. Thus $End(V)$ is a non-commutative, associative algebra over $\mathbb{F}$. This algebra differs from a Lie algebra in this multiplication operation, since in a Lie algebra we have the Jacobi identity on the bracket product replacing associativity. Another consequence of the associative multiplication is that it has a natural multiplicative identity. The identity function in $End(V)$ is an identity with respect to composition. And thus every associative algebra with a multiplicative identity has a subgroup which contains all the elements which have multiplicative inverses. In the case of $End(V)$ this is just the group of non-singular linear transformations of V or the group of invertible transformations in $End(V)$ or the group of automorphisms of V. It is given the symbol $Aut(V)$.

Finally, we come to a most surprising structure. Every associative algebra over $\mathbb{F}$ can be made into an Lie algebra over $\mathbb{F}$: the Lie bracket of two elements is what is called the commutator of these two elements. Thus for any associative algebra $\mathcal{A}$, and for a and b in $\mathcal{A}$, we define

$$[a, b] := ab - ba$$

[We remark that if the associative algebra $\mathcal{A}$ is commutative, the commutator is always 0. Thus, in general, the commutator measures the lack of commutativity in the multiplication structure of the algebra $\mathcal{A}$.]

It is trivial that $[b, a] = ba - ab = -[a, b]$. The Jacobi identity holds as well because for any three elements a, b, and c in $\mathcal{A}$, we have:

$$
\begin{aligned}
[[a,b],c] &= [a,b]c - c[a,b] = (ab-ba)c - c(ab-ba) = abc - bac - cab + cba \\
[[c,a],b] &= [c,a]b - b[c,a] = (ca-ac)b - b(ca-ac) = cab - acb - bca + bac \\
[[b,c],a] &= [b,c]a - a[b,c] = (bc-cb)a - a(bc-cb) = bca - cba - abc + acb
\end{aligned}
$$

and thus, after combining, we have

$$[[a,b],c] + [[c,a],b] + [[b,c],a] = 0$$

Also, scalar multiplication is bilinear with respect to the Lie bracket:

$$
\begin{aligned}
\alpha[a,b] &= [\alpha a, b] = [a, \alpha b] \\
\alpha[a,b] &= \alpha(ab - ba) = \alpha(ab) - \alpha(ba) \\
[\alpha a, b] &= (\alpha a)b - b(\alpha a) = \alpha(ab) - \alpha(ba) \\
[a, \alpha b] &= a(\alpha b) - (\alpha b)a = \alpha(ab) - \alpha(ba)
\end{aligned}
$$

Obviously the linear space structure of $\mathcal{A}$ remains the same in the Lie algebra structure. Moreover, this Lie bracket distributes over addition both on the left and on the right, that is,

$$[a, b+c] = a(b+c) - (b+c)a = ab + ac - ba - ca = (ab - ba) + (ac - ca) = [a,b] + [a.c].$$

Likewise, we have

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

Thus any associative algebra $\mathcal{A}$ can be given the structure of a Lie algebra by means of the commutator. In particular $End(V)$ with this bracket takes on the structure of a Lie algebra over $\mathbb{F}$.

When we wish to treat $End(V)$ as a Lie algebra, we use the notation $End_L(V)$ or $\widehat{gl}(V)$ [this latter notation will become clear in a moment]. Also quite frequently we use $GL(V)$ for $Aut(V)$.

There is one more important and natural associative algebra over $\mathbb{F}$. If we are given a finite dimensional linear space V over $\mathbb{F}$ of dimension n, we know that bases exist for V. If we fix a basis $\mathcal{B}$ for V, then it is well known that we have a bijective linear transformation from V to $\mathbb{F}^n$, where we are using the canonical basis in $\mathbb{F}^n$. Also we can represent an element of $\mathbb{F}^n$ as an nx1 <u>column</u> matrix in $M_{nx1}(\mathbb{F})$ using this same basis. Then any linear transformation of V to V, that is, any endomorphism ϕ in $End(V)$, has a representation as an nxn matrix A over $\mathbb{F}$. The following commutative diagram illustrates this situation:

$$\begin{array}{ccc} V & \xrightarrow{\ \phi\ } & V \\ \mathcal{B}\downarrow & & \downarrow\mathcal{B} \\ M_{n\text{x}1}(\mathbb{F}) & \xrightarrow{\ A\ } & M_{n\text{x}1}(\mathbb{F}) \end{array}$$

We know how to add square matrices, to scalar multiply square matrices, and to multiply square matrices [row-by-column multiplication], which operations give the square matrices $M_{n\text{x}n}(\mathbb{F})$ the structure of an associative algebra over $\mathbb{F}$ isomorphic to the structure of $End(V)$. The multiplicative identity is the identity matrix I_n, and the group of non-singular square matrices is the general linear group $GL(n, \mathbb{F})$. [It is for this reason that the notation $GL(V)$ is used for $Aut(V)$]. Also when we give $M_{n\text{x}n}(\mathbb{F})$ the structure of a Lie algebra by defining the bracket as the commutator, we use the notation $\widehat{gl}(n, \mathbb{F})$, which denotes the Lie algebra of the Lie group $GL(n, \mathbb{F})$.

Even though these are concepts and relations which we do not wish to explore at the present moment, we still want to point out that the notation $\widehat{gl}(V)$ is also used for $End_L(V)$. Also, when there is a fixed basis being used in an exposition, we frequently do not distinguish between $\widehat{gl}(V)$ and $\widehat{gl}(n, \mathbb{F})$, and we just say that the elements of $\widehat{gl}(V)$ are matrices.

2.6.2 The Adjoint Representation of a Lie Algebra. Now we can define the key concept of a representation. A *representation of Lie algebra $\hat{g}$ on a linear space* V is a homomorphism ρ of the Lie algebra $\hat{g}$ into the Lie algebra $\widehat{gl}(V)$, that is, it is a linear transformation such that brackets are preserved, that is, $\rho([a, b]) = [\rho(a), \rho(b)] = \rho(a)\rho(b) - \rho(b)\rho(a)$. It is easy to see the motivation for looking at representations— we study an unknown object $\hat{g}$ as it realizes itself in a very well known object $\widehat{gl}(V)$ or $\widehat{gl}(n, \mathbb{F})$. This exposition, then, explores certain aspects of this representation structure.

It is remarkable that there exists a very natural representation of a Lie algebra $\hat{g}$ on its own linear space $\hat{g}$. This representation is called the *adjoint representation* and depends on the bracket product. It is one of the most important objects that we have for studying the structure of a Lie algebra. We give it the name *ad* and it is defined as follows:

$$\begin{array}{lcll} \hat{g} & \xrightarrow{ad} & \widehat{gl}(\hat{g}) & \\ a & \longmapsto & ad(a): & \hat{g} \longrightarrow \hat{g} \\ & & & b \longmapsto ad(a)(b) := [a, b] \end{array}$$

First, we see that $ad(a)$ is a linear map and thus is in $\widehat{gl}(\hat{g})$:

$$ad(a)(b_1 + b_2) = [a, b_1 + b_2] = [a, b_1] + [a, b_2] = ad(a)(b_1) + ad(a)(b_2)$$
$$ad(a)(\alpha b) = [a, \alpha b] = \alpha[a, b] = \alpha(ad(a)(b))$$

Next, we see that ad is a linear map:

$$(ad(a_1 + a_2))(b) = [(a_1 + a_2), b] = [a_1, b] + [a_2, b] = ad(a_1)(b) + ad(a_2)(b) =$$
$$(ad(a_1) + ad(a_2))(b)$$
$$(ad(\alpha a))(b) = [\alpha a, b] = \alpha[a, b] = \alpha(ad(a)(b)) = (\alpha ad(a))(b)$$

We remark how we have used the bilinearity of addition and scalar multiplication distribution in the Lie algebra to effect these calculations.

Finally, we show that ad preserves brackets, which is nothing more than another way of expressing the Jacobi identity. [Recall that brackets in $\widehat{gl}(\hat{g})$ are commutators.]

$$(ad[a_1, a_2])(b) = [[(a_1, a_2)], b] = -[[b, a_1], a_2] - [[a_2, b], a_1] =$$
$$-[a_2, [a_1, b]] + [a_1, [a_2, b]] = -ad(a_2)([a_1, b]) + ad(a_1)([a_2, b]) =$$
$$-ad(a_2)(ad(a_1)(b)) + ad(a_1)(ad(a_2)(b)) =$$
$$(ad(a_1)ad(a_2))(b) - (ad(a_2)ad(a_1))(b) =$$
$$((ad(a_1)ad(a_2) - (ad(a_2)ad(a_1))(b) = [ad(a_1), ad(a_2)](b)$$

Thus we have in ad a homomorphism of Lie algebras and a representation of $\hat{g}$ in $\hat{g}$.

We pause a moment to make some observations. The meaning of the symbol $\widehat{gl}(\hat{g})$ is clear. It is the set of linear endomorphisms of the linear space $\hat{g}$, but with a Lie algebra structure defined by commutation. There is a subset of these, the non-singular linear endomorphisms, or automorphisms, which form a group $Aut(\hat{g})$ or $GL(\hat{g})$. But since $\hat{g}$ is more than just a linear space, for it has a bracket product, we can ask whether these endomorphisms and automorphisms respect this bracket product. Thus an ambiguity can arise with respect to these structures. We thus will conform to the following usage. An element ϕ of $GL(\hat{g})$ which also is an automorphism of the Lie algebra structure of $\hat{g}$ will be said to belong to $Aut(\hat{g})$. An element ϕ of $\widehat{gl}(\hat{g})$ which also preserves the brackets in $\hat{g}$ unfortunately will be given no special name or symbol, and it will have to be described fully every time that it occurs. Finally, there is a set of endomorphisms of $\hat{g}$, thus belonging to $\widehat{gl}(\hat{g})$, which occur significantly in this structure. They are called *derivations*. They act on the bracket product by a Leibniz rule. Thus an endomorphism D in $\widehat{gl}(\hat{g})$ is a derivation if $D([a, b]) = [D(a), b] + [a, D(b)]$. It evidently cannot be an endomorphism of Lie algebra structure on $\hat{g}$, since by definition it does not preserve brackets. Surprisingly $ad(a)$, coming from the representation ad, is a derivation in $\widehat{gl}(\hat{g})$. We will return to this point later.

We make some preliminary observations on this adjoint representation. First, it is highly non-surjective. The dimension of the domain space is the dimension n of the Lie algebra $\hat{g}$. The target space is $\widehat{gl}(\hat{g})$, which has the dimension n^2. Next , we see that the kernel of the adjoint map is the center of the Lie algebra since $ad(a) = 0$ means that for all b in $\hat{g}$, $ad(a)(b) = [a, b] = 0$. This says that a is in the center $\hat{z}$ of the Lie algebra.

2.7 Nilpotent Lie algebras (2)

Nilpotent Lie Algebras Determine Nilpotent Linear Transformations. We will need to examine the concept of a nilpotent Lie algebra in depth. In fact, with the adjoint representation defined, we can give a rather complete picture of a nilpotent Lie algebra. Thus we now consider a nilpotent Lie algebra $\hat{n}$ and ask the question: How does information about the lower central series, which terminates at 0 in this case, transfer over by the adjoint representation? We begin by taking an a_1 in $\hat{n}$ and mapping it over to $\widehat{gl}(\hat{n})$ by the adjoint:

$$\begin{aligned} a_1 \quad \longmapsto \quad ad(a_1) : \quad & \hat{n} \longrightarrow \hat{n} \\ & b \longmapsto ad(a_1)(b) = [a_1, b] \end{aligned}$$

We do the same with a_2 in $\hat{n}$, but this time we act on $ad(a_1)(b) = [a_1, b]$ in $\hat{n}$:

$$\begin{aligned} a_2 \quad \longmapsto \quad ad(a_2) : \quad & \hat{n} \longrightarrow \hat{n} \\ & ad(a_1)(b) \longmapsto ad(a_2)(ad(a_1)(b)) = [a_2, [a_1, b]] \end{aligned}$$

We continue in this manner:

$$\begin{aligned} a_3 \quad \longmapsto \quad ad(a_3) : \quad & \hat{n} \longrightarrow \hat{n} \\ & ad(a_2)(ad(a_1)(b)) \longmapsto ad(a_3)(ad(a_2)(ad(a_1)(b))) = \\ & \qquad [a_3, [a_2, [a_1, b]]] \end{aligned}$$

$$\vdots$$

$$\begin{aligned} a_k \quad \longmapsto \quad ad(a_k) : \quad & \hat{n} \longrightarrow \hat{n} \\ & ad(a_{k-1})(\cdots(ad(a_3)(ad(a_2)(ad(a_1)(b))))\cdots) \\ & \longmapsto ad(a_k)(ad(a_{k-1})(\cdots(ad(a_3)(ad(a_2)(ad(a_1)(b))))\cdots)) = \\ & \qquad [a_k, [a_{k-1}, [\cdots[a_3, [a_2, [a_1, b]]]\cdots]]] \end{aligned}$$

Now this means that the last term in each of the above maps can also be written, as we have already displayed:

$$
\begin{aligned}
ad(a_1)(b) &= [a_1, b] \\
ad(a_2)([a_1, b]) &= [a_2, [a_1, b]] \\
ad(a_3)([a_2, [a_1, b]]) &= [a_3, [a_2, [a_1, b]]] \\
&\cdot \\
&\cdot \\
&\cdot \\
ad(a_k)([a_{k-1}, [\cdots[a_3, [a_2, [a_1, b]]]) &= \\
&[a_k, [a_{k-1}, [\cdots[a_3, [a_2, [a_1, b]]] \cdots]]]
\end{aligned}
$$

But this immediately reveals how the adjoint map treats the lower central series, namely

$$
\begin{aligned}
ad(\hat{n})(C^0(\hat{n})) &= C^1(\hat{n}) \\
ad(\hat{n})(C^1(\hat{n})) &= C^2(\hat{n}) \\
ad(\hat{n})(C^2(\hat{n})) &= C^3(\hat{n}) \\
&\cdot \\
&\cdot \\
&\cdot \\
ad(\hat{n})(C^{k-1}(\hat{n})) &= C^k(\hat{n})
\end{aligned}
$$

[We again make the observation that we are defining the lower central series with brackets moving from right to left. But as we remarked above, this does not matter.]

Thus we have this beautiful conclusion. Let $\hat{n}$ be a nilpotent Lie algebra. Suppose $C^{k-1}(\hat{n}) \neq 0$ and $C^k(\hat{n}) = 0$. Then any product $ad(a_1) \cdots ad(a_k)$ of k factors is the zero transformation in $\widehat{gl}(\hat{n})$. In particular $ad(a)^k = 0$ for any a in $\hat{n}$, which means that

$ad(a)$ is a nilpotent <u>linear</u> transformation in $\widehat{gl}(\hat{n})$ for every a in $\hat{n}$.

Indeed this is one of the conclusions that we are seeking.

But we can assert more than this. Knowing that $C^{k-1}(\hat{n}) \neq 0$ and $C^k(\hat{n}) = 0$, we can relate the index k to the dimension of $\hat{n}$. Let us say that the dimension of $\hat{n}$ is l. First we claim that $\dim(C^r(\hat{n})) > \dim(C^{r+1}(\hat{n}))$ with the bound $(r+1) < k$. We have $C^{r+1}(\hat{n}) = [C^r(\hat{n}), \hat{n}]$ and $C^{r+1}(\hat{n}) \subset C^r(\hat{n})$. Now if $C^{r+1}(\hat{n}) = C^r(\hat{n})$, then $[C^r(\hat{n}), \hat{n}] = C^r(\hat{n})$ and the lower central series would stall at $C^r(\hat{n})$ and never reach 0. But our Lie algebra $\hat{n}$ is nilpotent. Thus $C^{r+1}(\hat{n}) \subset C^r(\hat{n})$ properly and $\dim(C^r(\hat{n})) > \dim(C^{r+1}(\hat{n}))$. Thus starting with $\hat{n} = C^0(\hat{n})$, we have

$$
\begin{gathered}
\hat{n} = C^0(\hat{n}) \supset C^1(\hat{n}) \supset C^2(\hat{n}) \supset \cdots \supset C^{k-1}(\hat{n}) \supset C^k(\hat{n}) = 0 \\
\dim(\hat{n}) = \dim(C^0(\hat{n})) > \dim(C^1(\hat{n})) > \dim(C^2(\hat{n})) > \cdots > \dim(C^{k-1}(\hat{n})) > \\
\dim(C^k(\hat{n})) = 0
\end{gathered}
$$

Since the $\dim(\hat{n}) = l$, we can conclude that, if the above decreases the dimension by only one in each step, the largest non-zero index r in $C^r(\hat{n})$ that can occur is $l-1$. Thus $k \leq l$.

2.8 Engel's Theorem

2.8.1 Statement of Engel's Theorem. In the 19th century when these ideas were first being explored, the idea of a Lie algebra was linked to transformations of a linear space V of dimension n. We know that $End(V)$ can be made into a Lie algebra by defining the bracket as the commutator in $End(V)$, and with this structure we can rename the set $\widehat{gl}(V)$. Thus a Lie subalgebra in $\widehat{gl}(V)$ is a well defined object in $\widehat{gl}(V)$.

Now the classical *Engel's Theorem* is the following:

> Let $\hat{g}$ be a Lie subalgebra of $\widehat{gl}(V)$. Suppose that every $X \in \hat{g}$ is a nilpotent <u>linear</u> transformation in $\widehat{gl}(V)$. Then there exist a non-zero vector v in V which is a simultaneous eigenvector with eigenvalue 0 for all X in $\hat{g}$.

It is easy to see how abstract nilpotent Lie algebras fit into this context. For any nilpotent Lie algebra $\hat{n}$ we know that for every a in $\hat{n}$, $ad(a)^k = 0$ for some k, that is, $ad(a)$ is a nilpotent <u>linear</u> transformation in $\widehat{gl}(\hat{n})$. Also we know the homomorphic image of $\hat{n}$ by ad gives a Lie subalgebra $ad(\hat{n})$ of $\widehat{gl}(\hat{n})$. We see that this satisfies the supposition of the above theorem. Using the theorem, we conclude that there exist a non-zero vector v in $\hat{n}$ which is a simultaneous eigenvector with eigenvalue 0 for all $ad(a)$ in $\widehat{gl}(\hat{n})$. Note that we are able to come to this conclusion for the [abstract] nilpotent Lie algebra above by using the fact that it has a non-zero center. And we remark that $(ad(a))v = [a, v] = 0$ for all a in $\hat{n}$ says that v belongs to the center of $\hat{n}$.

Now using this simultaneous eigenvector v with eigenvalue 0 for all the linear transformations in $\hat{g}$ to build a quotient space, we can find a basis in which all the transformations of $\hat{g}$ are represented by upper diagonal matrices with a zero diagonal. We give here the details of this construction.

Thus we want to give a basis for V and represent any X in $\hat{g}$ by the matrix A with respect to that basis. We let the first vector v_1 in this basis be v. Then $X(v_1) = 0$ gives the first column of the matrix A as the zero column matrix:

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ . \\ . \\ . \\ 0 \end{bmatrix}$$

But knowing that v is a <u>simultaneous</u> eigenvector with eigenvalue 0 for all Y in $\hat{g}$, then we have for any other Y in $\hat{g}$ that the matrix representation B of Y with respect to this first basis vector v_1 will also be the zero column matrix.

We now form the quotient linear space $V/sp(v_1)$ and note that $\hat{g}$ induces a linear transformation in this linear space. Let X be an element of $\hat{g}$. Without changing the symbol of this transformation, we have

$$X : V/sp(v_1) \longrightarrow V/sp(v_1)$$
$$u + sp(v_1) \longmapsto X(u + sp(v_1)) := X(u) + sp(v_1)$$

Now $X(u + sp(v_1)) = X(u) + X(sp(v_1))$. But $X(sp(v_1)) \subset (sp(v_1)$ [in fact, $(X(sp(v_1)) = 0]$. Thus we can conclude that

$$X(u + sp(v_1)) = X(u) + sp(v_1)$$

and thus we have a valid definition of how X operates on $V/sp(v_1)$.

[To be complete in our argument, we should show that cosets go into cosets. But the above relation implies this. For if we take two elements in the same coset, $u + sp(v_1)$ and $u_2 + sp(v_1)$, we can show that they map into the same coset. This means that the difference of their images maps into the zero coset, i.e., into $sp(v_1)$. Since $u_1 + sp(v_1)$ and $u_2 + sp(v_1)$ are in the same coset, their difference must be in the zero coset $sp(v_1)$. Thus $(u_1 + sp(v_1)) - (u_2 + sp(v_1)) = (u_1 - u_2) + sp(v_1) \subset sp(v_1)$ Thus

$$X((u_1 + sp(v_1)) - (u_2 + sp(v_1))) \subset X(sp(v_1)) \subset sp(v_1)$$
$$X((u_1 + sp(v_1)) - X((u_2 + sp(v_1)) \subset sp(v_1)$$

We conclude that we have a valid definition.]

But we can also assert that for all X in $\hat{g}$, X acts as a nilpotent linear transformation on $V/sp(v_1)$. We know that for some r, $X^r = 0$ since X is a nilpotent linear transformation in V. Thus for any $u+sp(v_1)$ in $V/sp(v_1)$, we have $X^r(u+sp(v_1)) \subset X^r(u)+sp(v_1) = 0+sp(v_1) = sp(v_1)$, which, of course, is 0 in $V/sp(v_1)$. We can therefore apply Engel's Theorem again and assert that there exist a non-zero $v_2 + sp(v_1)$ in $V/sp(v_1)$ which is a simultaneous eigenvector with eigenvalue 0 for all X in $\hat{g}$.

$$X(v_2 + sp(v_1)) = sp(v_1).$$

Thus $X(v_2)$ gives some element in $sp(v_1)$. This fact gives the second column of the matrix A:

$$\begin{bmatrix} 0 & * \\ 0 & 0 \\ 0 & 0 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 0 & 0 \end{bmatrix}$$

But the simultaneity of the eigenvector with eigenvalue 0 for all X in $\hat{g}$ means that for any other Y in $\hat{g}$, the matrix representation B of Y with respect to the basis vectors (v_1, v_2) will also give a column matrix of the same form [with, of course, a different value for $*$ in the $(1, 2)$ place in the matrix]. Thus all the matrix representations of $\hat{g}$ begin with the same configuration for the first two columns of the matrix.

Now obviously this process can be continued until we find a basis $(v_1, \cdot\cdot\cdot, v_n)$ in V (where, recall, the dimension of V is n) such that all the matrices coming from any X in $\hat{g}$ take the form:

$$\begin{bmatrix} 0 & * & * & \cdot & \cdot & \cdot & * \\ 0 & 0 & * & \cdot & \cdot & \cdot & * \\ 0 & 0 & 0 & * & \cdot & \cdot & * \\ & & & \cdot & & & \\ & & & \cdot & & & \\ & & & \cdot & & & \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 \end{bmatrix}$$

With these results we can make an important observation in Representation Theory. It concerns what is called an <u>irreducible</u> representation. We say a representation is irreducible if it has no proper invariant subspaces. If, as an example, we have a representation of a 4-dimensional Lie algebra $\hat{g}$ with basis (e_1, e_2, e_3, e_4) by ϕ into the $4x4$ matrices:

$$\hat{g} \longrightarrow M_{4\text{x}4}(\mathbb{F})$$

a 2-dimensional reducible representation would take the form

$$\begin{bmatrix} * & * & * & * \\ * & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & * & * \end{bmatrix}$$

since

$$\begin{bmatrix} * \\ * \\ 0 \\ 0 \end{bmatrix} \longrightarrow \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & * & * \end{bmatrix} \cdot \begin{bmatrix} * \\ * \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} * \\ * \\ 0 \\ 0 \end{bmatrix}$$

Thus a representation is irreducible if it cannot be blocked as follows:

$$\begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$$

Now from the form of the above matrices coming from the X in $\hat{g}$ we see immediately that the only irreducible representations of a nilpotent Lie algebra are one-dimensional, i.e., V must be of dimension one.

2.8.2 Nilpotent Linear Transformations Determine Nilpotent Lie Algebras. But there is one more important remark that we can make at this point. Suppose we start with an abstract Lie algebra $\hat{g}$, and map it over to $\widehat{gl}(\hat{g})$ by the adjoint map and we further assume that for every x in $\hat{g}$, $ad(x)$ in $\widehat{gl}(\hat{g})$ is a nilpotent <u>linear</u> transformation. Then Engel's Theorem and its consequences state that $ad(\hat{g})$ is the set of linear transformations in $\widehat{gl}(\hat{g})$ which can be represented as upper triangular matrices with a zero diagonal. As the example (which we will give after the proof of Engel's Theorem) shows, these matrices form a nilpotent Lie subalgebra of $\widehat{gl}(\hat{g})$. Now we can obtain the converse of the theorem, namely that if $\hat{g}$ is a nilpotent Lie algebra, then for each x in $\hat{g}$ the transformation $ad(x)$ in $\widehat{gl}(\hat{g})$ is a nilpotent linear transformation. Thus, if we assume for a Lie algebra $\hat{g}$ that for each x in $\hat{g}$ the transformation $ad(x)$ in $\widehat{gl}(\hat{g})$ is a nilpotent linear transformation, Engel's Theorem then asserts that $ad(\hat{g})$ can be represented as a set of upper triangular matrices each with a zero diagonal. Thus we know that for some r, every product of r-factors of the form $ad(x_1)\cdots ad(x_r)$ has the property that $ad(x_1)\cdots ad(x_r) = 0$. If we let this product act on an arbitrary element y in $\hat{g}$ then $(ad(x_1)\cdots ad(x_r))(y) = 0$. But this translates to $[x_1, [x_2, [\cdots, [x_r, y]\cdots]]] = 0$. We can conclude that $C^r(\hat{g}) = 0$, which says that $\hat{g}$ is nilpotent.

2.8.3 Proof of Engel's Theorem. Starting with $\hat{g}$ as a Lie subalgebra of $\widehat{gl}(V)$, in which every element of $\hat{g}$ is a nilpotent linear transformation in $End(V)$, we can build up an element which becomes zero in its lower central series. We can do this since we know how to calculate the bracket product in $\hat{g}$ because it is nothing but the commutator of the associative algebra $End(V)$.

Thus what we want to do is calculate a series of brackets in $\hat{g}$:

$$[X,[X,[X,[\cdots,[X,Y]\cdots]]]] \text{ for } X,Y \text{ in } \hat{g}$$

until the result is zero, on the hypothesis that every element X in $\hat{g}$ is a nilpotent linear transformation. First we rewrite the above in the adjoint notation:

$$ad(X)ad(X)ad(X)\cdots ad(X)(Y) \text{ for } X,Y \text{ in } \hat{g}$$

and thus we see that we are just repeating the adjoint representation of $\hat{g}$ on $\hat{g} \subset \widehat{gl}(V)$ since

$$X \longmapsto ad(X) : \hat{g} \longrightarrow \hat{g}$$
$$Y \longmapsto (ad(X))Y = [X,Y]$$

Since $\hat{g}$ is a Lie subalgebra of $\widehat{gl}(V) = End(V)$, we know that the brackets close in $\hat{g}$, and thus we know that $(ad(X))Y = [X,Y]$ is in $\hat{g}$. We remark that now we are considering $ad(X)$ to be in $End(\hat{g}) = \widehat{gl}(\hat{g}) \subset \widehat{gl}(End(V))$, which in a matrix representation would give an n^2 x n^2 matrix, where n is the dimension of V. And finally, we observe that if we can show that these brackets do terminate in zero, then we can conclude that every $ad(X)$ in $ad(\hat{g})$ is a nilpotent linear transformation in $End(\hat{g})$. Thus we need to show that $ad(\hat{g})$, a Lie subalgebra of $\widehat{gl}(\hat{g})$, has the property that for all $ad(X)$ in $ad(\hat{g})$, $ad(X)$ is a nilpotent linear transformation in $\widehat{gl}(\hat{g})$, on the hypothesis that every element X in $\hat{g}$ is a nilpotent linear transformation in $End(V)$. [Obviously we use the commutator as the definition of the bracket in $\hat{g}$.]

Thus we need to show that for any $ad(X)$, we can find an s such that $(ad(X))^s = 0$ when it operates on $\hat{g}$. This says that for any Y in $\hat{g}$, $((ad(X))^s)Y = 0$, and thus $ad(X)$ is a nilpotent <u>linear</u> transformation in $End(\hat{g})$. Now $((ad(X))^s)Y = ad(X)(\cdots(ad(X)((ad(X)Y)))\cdots)$ for s repetitions of $ad(X)$. But $ad(X)(\cdots(ad(X)((ad(X)Y)))\cdots) = [X,\cdots[X,[X,Y]]\cdots]$. However since X and Y are matrices in $\hat{g} \subset \widehat{gl}(V)$, these are now not just abstract brackets. We can calculate these brackets, because they are commutators: $[X,Y] = XY - YX$. Thus we have:

$$(ad(X))(Y) = [X,Y] = XY - YX$$
$$((ad(X)^2)(Y) = [X, XY - YX] = XXY - XYX - XYX + YXX =$$
$$XXY - 2XYX + YXX$$
$$((ad(X)^3)(Y) = [X, XXY - 2XYX + YXX] =$$
$$XXXY - XXYX - 2XXYX + 2XYXX + XYXX - YXXX =$$
$$XXXY - 3XXYX + 3XYXX - YXXX$$
$$((ad(X)^4)(Y) = [X, XXXY - 3XXYX + 3XYXX - YXXX] =$$
$$XXXXY - XXXYX - 3XXXYX + 3XXYXX + 3XXYXX -$$

$$3XYXXX - XYXXX + YXXXX =$$
$$XXXXY - 4XXXYX + 6XXYXX - 4XYXXX + YXXXX$$
$$\vdots$$

> Thus we observe that the power of X as a linear transformation in $\widehat{gl}(V) = End(V)$ is increasing either on the left or on the right of Y. We know that since X is a nilpotent linear transformation in $End(V)$, for some r, $X^r = 0$ and thus it is just a matter of combinatorics to see how large s must be so that each power of X is equal to or greater than r in every term. When this s is reached, we then have our conclusion that $(ad(X))^s = 0$.

But we can make the observation now that this sets up the hypothesis of Engel's Theorem in this situation. Our linear space of interest is $\hat{g}$ in $End(V)$ and we have the Lie algebra of commutators in $\widehat{gl}(\hat{g})$. Our Lie subalgebra of $\widehat{gl}(\hat{g})$ is $ad(\hat{g})$. [Since ad is a homomorphism of Lie algebras, and our $\hat{g}$ is a Lie algebra, $ad(\hat{g})$ is a Lie subalgebra of $\widehat{gl}(\hat{g})$.] Now we have that for each X in $\hat{g}$, $ad(X)$ in $ad(\hat{g})$ is a nilpotent linear transformation in $\widehat{gl}(\hat{g})$. Thus we satisfy the conditions of Engel's Theorem in this situation. Engel's Theorem would conclude that there exists a nonzero linear transformation Y in $\hat{g}$ which is a simultaneous eigenvector with eigenvalue zero for all transformations $ad(X)$ in $ad(\hat{g})$: $ad(X) \cdot Y = [X, Y] = 0$. We remark that we have, in this processThus with this idenidentified a non-zero element Y in the center of $\hat{g}$.

Now to prove Engel's Theorem we still need to identify a vector v in V which is a simultaneous eigenvector with eigenvalue zero for all X in $\hat{g}$.

The theorem is true for $\hat{g}$ of dimension one. This can be seen as follows. In this case $\hat{g}$ in $End(V)$ is generated by any nonzero X in $\hat{g}$. But by hypothesis $X \neq 0$ is a nilpotent linear transformation in $End(V)$. Thus it has an eigenvector $v \neq 0$ in V with eigenvalue 0, i.e., $X(v) = 0$. Now any other Y in $\hat{g}$ is a scalar multiple of X, i.e., $Y = cX$ for c in the scalar field. Thus $Y(v) = (cX)(v) = c(X(v)) = c(0)) = 0$. And thus we have the simultaneous eigenvector $v \neq 0$ with eigenvalue $= 0$.

We consider the following. Let $\dim(\hat{g})$ be greater than 1 and let $\hat{h}$ be any proper subalgebra of $\hat{g}$. [Of course, such a subalgebra exists. For example choose any one-dimensional subspace of $\hat{g}$. Note that this is an abelian subalgebra of $\hat{g}$.] We form the quotient linear space $\hat{g}/\hat{h}$. Since $\hat{h}$ is a proper subalgebra of $\hat{g}$, we know that the dimension of $\hat{g}/\hat{h}$ is greater than or equal to one. Now we can define an action of $ad(Z)$ on $\hat{g}/\hat{h}$ for each Z in $\hat{h}$:

$$ad(Z) : \hat{g}/\hat{h} \longrightarrow \hat{g}/\hat{h}$$
$$X + \hat{h} \longmapsto ad(Z)(X + \hat{h}) := ad(Z)(X) + \hat{h}$$

We see that $ad(Z)(X + \hat{h}) = ad(Z)(X) + ad(Z)(\hat{h}) \subset ad(Z)(X) + \hat{h}$, since $ad(Z)(\hat{h}) = [Z, \hat{h}]$. Now $\hat{h}$ is a subalgebra and Z is in $\hat{h}$. Thus $ad(Z)(\hat{h})$ is in $\hat{h}$, and we have a valid definition [according to 2.7.2]. We also have a subalgebra in $\widehat{gl}(\hat{g}/\hat{h})$ since $[adZ_1, adZ_2] = ad[Z_1, Z_2]$, which is in $ad(\hat{h})$.

From the proof given above we know that $ad(\hat{g})$, which is a Lie subalgebra of $\widehat{gl}(\hat{g})$, has the property that for all $ad(X)$ in $ad(\hat{g})$, $ad(X)$ is a nilpotent linear transformation in $End(\hat{g})$. Thus $ad(\hat{h})$ has the same property. And we see from the above definition that this same property holds when $ad(\hat{h})$ acts on $\hat{g}/\hat{h}$. But we can now make this observation. $\hat{h}$ is a Lie algebra of dimension less that the dimension of $\hat{g}$. Suppose we now assume, by induction, that Engel's theorem is valid for all Lie algebras of dimension less than that of $\hat{g}$ where the linear space V can be any fixed linear space. In particular suppose this linear space is $\hat{g}/\hat{h}$, and our Lie algebra is $ad(\hat{h})$. Since $\dim \hat{h} < \dim \hat{g}$, we know that the $\dim ad(\hat{h}) \leq \dim \hat{h} < \dim \hat{g}$. Thus we can use the induction hypothesis and we can assert that there exist a $\overline{Y}$ in $\hat{g}/\hat{h}$ not equal to zero such that it is a simultaneous eigenvector with eigenvalue zero for all $ad(Z)$ in $ad(\hat{h})$.

$$\overline{Y} \longmapsto ad(Z)(\overline{Y}) = 0 \qquad \text{for all } ad(Z) \text{ in } ad(\hat{h})$$

Unwinding this quotient, we have a $Y \neq 0$ in $\hat{g}$ but not in $\hat{h}$ such that for all $ad(Z)$ in $ad(\hat{h})$, $ad(Z)(Y) = [Z, Y]$ is in $\hat{h}$.

Now we consider the subspace $\hat{h} \oplus sp(Y) \subset \hat{g}$. We know that this subspace has dimension one higher than $\hat{h}$. Using the above relation $ad(Z)(Y) = [Z, Y]$ in $\hat{h}$ for all Z in $\hat{h}$, we know that this subspace is actually a subalgebra of $\hat{g}$.

$$[\hat{h} \oplus sp(Y), \hat{h} \oplus sp(Y)] \subset [\hat{h}, \hat{h}] + [\hat{h}, sp(Y)] + [sp(Y), sp(Y)] \subset \hat{h} + \hat{h} + 0 = \hat{h}$$

Now if $\hat{h} \oplus sp(Y)$ is still a proper subalgebra of $\hat{g}$, then we could have used this subalgebra in our induction step. What we are saying is that in the induction step we should be using a proper subalgebra which is a maximal proper subalgebra. Then we can conclude that $\hat{h} \oplus sp(Y)$ is actually equal to $\hat{g}$.

But we can say more. The above calculation shows in this case that $\hat{h}$ is actually a maximal proper ideal in $\hat{g}$.

$$[\hat{h}, \hat{g}] = [\hat{h}, \hat{h} \oplus sp(Y)] \subset [\hat{h}, \hat{h}] + [\hat{h}, sp(Y)] \subset \hat{h} + \hat{h} = \hat{h}$$

And thus in our induction step above we could have started with $\hat{h}$ being a maximal proper ideal.

At this point we can make some interesting remarks. It may have been observed above that we did not define an action of $ad(\hat{g})$ on $\hat{g}/\hat{h}$. We could not do this if $\hat{h}$ were only a subalgebra. This would have meant that for X in $\hat{g}$, we would have been calculating $[X, \hat{h}]$, which would not necessarily have been in $\hat{h}$. However now that we have an $\hat{h}$ that is an ideal, then $[X, \hat{h}]$ would be in $\hat{h}$. Thus we can define an action of $ad(\hat{g})$ on $\hat{g}/\hat{h}$. We know that for every $ad(X)$ in $ad(\hat{g})$, $ad(X)$ will be a nilpotent linear transformation on $\hat{g}/\hat{h}$. Thus we have all the conditions for the application of Engel's Theorem. But we remark that the induction step, after unwinding the quotient, only gave a $Y \neq 0$ in $\hat{g}$ which may not be in $\hat{h}$ but is such that $ad(Z)(Y)$ is in $\hat{h}$ for all Z in $\hat{h}$. If we calculate $ad(X)(Y)$ for all X in $\hat{g}$, we would only know that $ad(X)(Y)$ is in $\hat{g}$ even though we know that $Y \neq 0$ is not in $\hat{h}$. We would have to do more work to show that there exists an $A \neq 0$ in $\hat{g}$ such that $ad(X)(A) = 0$ for all X in $\hat{g}$. But we will do exactly this in the context of the original expression of Engel's Theorem.

Thus with this identification of the transformation Y, we can now show that there exists a vector v in V such that for all X in $\hat{g}$, v is a simultaneous eigenvector with eigenvalue 0. We again work by induction on the dimension of $\hat{g}$. We know that $\hat{h}$ has one dimension less that $\hat{g}$, and that for every Z in $\hat{h}$, Z is nilpotent linear transformation in V. Thus there exists a vector $v_1 \neq 0$ in V such that for every Z in $\hat{h}$, $Z(v_1) = 0$. We have now produced a transformation Y in $\hat{g}$ and a vector $v_1 \in V$ and we know that $\hat{g} = \hat{h} \oplus sp(Y)$. Now for every Z in $\hat{h}$, $Z(v_1) = 0$. If $Y(v_1) = 0$, we are finished, since for every X in $\hat{g}$, $X(v_1) = 0$. But of course we cannot assert that $Y(v_1) = 0$. However we can do the following. Let W be the subspace of V such that for every Z in $\hat{h}$, $Z(W) = 0$. Obviously $v_1 \in W$. Now what is remarkable is that for any $w \in W$, $Y(w)$ is in W, i.e., Ystabilizes W. We show this by taking the bracket $[Z, Y]$ in $\hat{g}$, where Z is any element in $\hat{h}$. This gives $Z(Y(w)) = [Z, Y](w) + Y(Z(w))$. Now $\hat{h}$ is an ideal in $\hat{g}$, and thus $[Z, Y] \in \hat{h}$ and we conclude that $[Z, Y](w) = 0$. Also $Z(w) = 0$, and thus $Y(Z(w)) = 0$. Now we have $Z(Y(w)) = 0$ and we thus we have $Y(w) \in W$. But this says that $Y(W) \subset W$. But we know that Y acts nilpotently on V. Thus there exists an r such that $Y^r = 0$ but $Y^{(r-1)} \neq 0$. Thus for some k between 0 and $r-1$ we have that $v = Y^k(v_1) \neq 0$ with $(Y(Y^k))(v_1) = Y(v) = 0$. Since v_1 is in W and Y stabilizes W, any iterate of v_1 by Y is in W. Thus v is in W. Now for any $X = Z + cY$ in $\hat{g}$, with Z in $\hat{h}$ and c a scalar, we have $X(v) = (Z + cY)(v) = Z(v) + (cY)(v) = 0 + 0 = 0$. Thus we have found a $v \neq 0$ in V such that for all X in $\hat{g}$, $X(v) = 0$, which is the conclusion of Engel's Theorem.

2.8.4 Examples. Some computed examples are enlightening. Let us use a V of dimension 4 and give it a basis (v_1, v_2, v_3, v_4). Then $End(V) = \widehat{gl}(V)$ is the set of 4x4 matrices. We look at the upper triangular matrices with a zero diagonal:

$$\begin{bmatrix} 0 & a_{12} & a_{13} & a_{14} \\ 0 & 0 & a_{23} & a_{24} \\ 0 & 0 & 0 & a_{34} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Given these matrices, we see immediately that v_1 is a simultaneous eigenvector with eigenvalue zero whose existence Engel's Theorem affirms. However what we would like to do is to follow the proof of the theorem in this case and see how the proof identifies this vector to be the vector that we are seeking.

It is evident that the set of all such matrices is a 6-dimensional subspace of the 16-dimensional space of matrices. But this subspace is also an associative subalgebra [but without an identity], for we have closure of products:

$$\begin{bmatrix} 0 & a_{12} & a_{13} & a_{14} \\ 0 & 0 & a_{23} & a_{24} \\ 0 & 0 & 0 & a_{34} \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & b_{12} & b_{13} & b_{14} \\ 0 & 0 & b_{23} & b_{24} \\ 0 & 0 & 0 & b_{34} \\ 0 & 0 & 0 & 0 \end{bmatrix} =$$
$$\begin{bmatrix} 0 & 0 & a_{12}b_{23} & a_{12}b_{24} + a_{13}b_{34} \\ 0 & 0 & 0 & a_{23}b_{34} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

From the shape of the product XY, it is clear that $XY - YX$ is also an upper triangular $4x4$ matrix with 0's along the diagonal and thus we also have a Lie subalgebra. We call this Lie subalgebra $\hat{g}$. This is the kind of Lie subalgebra of $End(V)$ spoken of in Engel' s Theorem. The only condition on this subalgebra is that each element of it is also a linear nilpotent transformation in $End(V)$. We observe that these matrices do indeed have the property of linear nilpotency. For any X in $\hat{g}$, we see that we have effectively already calculated $X \cdot X$ above. We now continue the iteration of multiplication by X:

$$\begin{bmatrix} 0 & 0 & a_{12}a_{23} & a_{12}a_{24} + a_{13}a_{34} \\ 0 & 0 & 0 & a_{23}a_{34} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & a_{12} & a_{13} & a_{14} \\ 0 & 0 & a_{23} & a_{24} \\ 0 & 0 & 0 & a_{34} \\ 0 & 0 & 0 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & 0 & 0 & a_{12}a_{23}a_{34} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & 0 & a_{12}a_{23}a_{34} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & a_{12} & a_{13} & a_{14} \\ 0 & 0 & a_{23} & a_{24} \\ 0 & 0 & 0 & a_{34} \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus after four products we obtain the zero transformation.

Thus we have all the conditions for the application of Engel's Theorem. What we are seeking to do is expose in this specific example the various elements that we find in the proof of Engel's Theorem. But before we do this, let us calculate brackets in $\hat{g}$.

We choose the standard basis for $End(V)$. We let E_{ij} be the matrix with 1 in the i-th row and j-th column, and with 0 in all the other 15 entries. This gives us 16 basis vectors. The basis for the subalgebra $\hat{g}$ is $(E_{12}, E_{13}, E_{14}, E_{23}, E_{24}, E_{34})$. Thus we have 15 distinct brackets:

$$\begin{array}{ccc} [E_{12}, E_{13}] = 0 & [E_{12}, E_{14}] = 0 & [E_{12}, E_{23}] = E_{13} \\ [E_{12}, E_{24}] = E_{14} & [E_{12}, E_{34}] = 0 & [E_{13}, E_{14}] = 0 \\ [E_{13}, E_{23}] = 0 & [E_{13}, E_{24}] = 0 & [E_{13}, E_{34}] = E_{14} \\ [E_{14}, E_{23}] = 0 & [E_{14}, E_{24}] = 0 & [E_{14}, E_{34}] = 0 \\ [E_{23}, E_{24}] = 0 & [E_{23}, E_{34}] = E_{24} & [E_{24}, E_{34}] = 0 \end{array}$$

This calculates $C^1\hat{g}$. Continuing, to calculate $C^2\hat{g}$, we have only 7 distinct brackets:

$$\begin{array}{ccc} [E_{12}, E_{13}] = 0 & [E_{12}, E_{14}] = 0 & [E_{12}, E_{24}] = E_{14} \\ [E_{13}, E_{14}] = 0 & [E_{13}, E_{24}] = 0 & [E_{14}, E_{24}] = 0 \\ & [E_{23}, E_{24}] = 0 & \end{array}$$

To calculate $C^3\hat{g}$, we have only 2 distinct brackets:

$$[E_{12}, E_{14}] = 0 \qquad [E_{13}, E_{14}] = 0$$

which terminates the lower central series for $\hat{g}$. Thus we have for this Lie algebra:

$$\dim C^0\hat{g} = 6, \dim C^1\hat{g} = 3, \dim C^2\hat{g} = 1. \dim C^3\hat{g} = 0$$

We can also conclude that $\hat{g}$ is Lie nilpotent and has a center which is one-dimensional and is generated by the matrix E_{14}.

Now we return to illustrating the proof of Engel's Theorem. The first step in the proof was to show that there exists an ideal $\hat{h}$ in $\hat{g}$ of codimension one, and a vector Y in $\hat{g}$ not in $\hat{h}$ such that $\hat{g} = \hat{h} \oplus sp(Y)$. Since $\hat{h}$ is an ideal, we know for all Z in $\hat{h}$, $(ad(Z))(Y)$ is in $\hat{h}$. The maximal proper ideal $\hat{h}$ that we can identify is the 5-dimensional Lie subalgebra of $\hat{g}$ which has the basis $(E_{13}, E_{14}, E_{23}, E_{24}, E_{34})$. Thus each element of $\hat{h}$ has the form:

$$\begin{bmatrix} 0 & 0 & a_{13} & a_{14} \\ 0 & 0 & a_{23} & a_{24} \\ 0 & 0 & 0 & a_{34} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

By taking brackets we see that we do have a Lie subalgebra:

$$\begin{array}{ccc} [E_{13}, E_{14}] = 0 & [E_{13}, E_{23}] = 0 & [E_{13}, E_{24}] = 0 \\ [E_{13}, E_{34}] = E_{14} & [E_{14}, E_{23}] = 0 & [E_{14}, E_{24}] = 0 \\ [E_{14}, E_{34}] = 0 & [E_{23}, E_{24}] = 0 & [E_{23}, E_{34}] = E_{24} \\ & [E_{24}, E_{34}] = 0 & \end{array}$$

which is an ideal since:

$$\begin{array}{ccc} [E_{12}, E_{13}] = 0 & [E_{12}, E_{14}] = 0 & [E_{12}, E_{23}] = E_{13} \\ \multicolumn{3}{c}{[E_{12}, E_{24}] = E_{14} \qquad [E_{12}, E_{34}] = 0} \end{array}$$

Now we use our first application of induction in our proof of Engel''s Theorem. The Lie algebra of the Theorem is $ad(\hat{g})$. It is acting on the linear space $\hat{g}/\hat{h}$, which is a one-dimensional space. Thus essentially $ad(\hat{g})$ is acting on $\hat{g}$. To apply the theorem we now need to show that each element $ad(X)$, for X in $\hat{g}$, acting on $\hat{g}$ by the adjoint action is a nilpotent linear transformation in $End(\hat{g})$. We already know that $C^3\hat{g} = 0$. This means the $[\hat{g}, [\hat{g}, [\hat{g}, \hat{g}]]]$ is 0. But we can write this as $(ad(\hat{g})ad(\hat{g})ad(\hat{g}))(\hat{g}) = 0$. Thus for any X in $\hat{g}$, $(ad(X)ad(X)ad(X))(\hat{g}) = 0$, or $(ad(X))^3(\hat{g}) = 0$.

In the actual proof of the theorem we just calculated the brackets without any knowledge of the lower central series. We knew that after a finite number of repetitions of the bracket product we would reach the zero transformation since each X in $\hat{g}$ is a nilpotent linear transformation. In our case we have already seen above that after 3 repetitions of the associative multiplication in $\hat{g}$, we obtain the zero matrix. Now in computing the brackets in $\hat{g}$, we have

$$(ad(X))(A) = [X, A] = XA - AX$$

$$(ad(X)^2)(A) = [X, XA - AX] = XXA - XAX - XAX + AXX = XXA - 2XAX + AXX$$

$$(ad(X)^3)(A) = [X, XXA - 2XAX + AXX] = XXXA - XXAX - 2XXAX + 2XAXX + XAXX - AXXX = XXXA - 3XXAX + 3XAXX - AXXX$$

$$(ad(X)^4)(A) = [X, XXXA - 3XXAX + 3XAXX - AXXX] = XXXXA - XXXAX - 3XXXAX + 3XXAXX + 3XXAXX - 3XAXXX - XAXXX + AXXXX = XXXXA - 4XXXAX + 6XXAXX - 4XAXXX + AXXXX$$

Thus at this point the linear nilpotency kicks in. We have

$$XXXXA - 4XXXAX + 6XXAXX - 4XAXXX + AXXXX = -4XXXAX + 6XXAXX - 4XAXXX$$

Thus

$$(ad(X)^5)(A) = [X, -4XXXAX + 6XXAXX - 4XAXXX] = -4XXXXAX + 4XXXAXX + 6XXXAXX - 6XXAXXX - 4XXAXXX + 4XAXXXX = 10XXXAXX - 10XXAXXX$$

$$(ad(X)^6)(A) = [X, 10XXXAXX - 10XXAXXX] = 10XXXXAXX - 10XXXAXXX - 10XXXAXXX + 10XXAXXXX = -20XXXAXXX$$

$$(ad(X)^7)(A) = [X, -20XXXAXXX] = -20XXXXAXXX + 20XXXAXXXX = 0$$

which, of course, is the conclusion we have been seeking.

However, in the case we are analyzing, using our knowledge of the lower central series, which terminates at $C^3\hat{g}$, we know that the above calculation only needs to be carried to $((ad(X)^3)(A)$. For completeness we show these calculations.

$$X = \begin{bmatrix} 0 & x_{12} & x_{13} & x_{14} \\ 0 & 0 & x_{23} & x_{24} \\ 0 & 0 & 0 & x_{34} \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad A = \begin{bmatrix} 0 & a_{12} & a_{13} & a_{14} \\ 0 & 0 & a_{23} & a_{24} \\ 0 & 0 & 0 & a_{34} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$[X,A] = XA - AX =$$
$$\begin{bmatrix} 0 & 0 & x_{12}a_{23} - x_{23}a_{12} & x_{12}a_{24} + x_{13}a_{34} - x_{24}a_{12} - x_{34}a_{13} \\ 0 & 0 & 0 & x_{23}a_{34} - x_{34}a_{23} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$[X,[X,A]] = \begin{bmatrix} 0 & 0 & 0 & -(x_{12}a_{23} - x_{23}a_{12})x_{34} + x_{12}(x_{23}a_{34} - x_{34}a_{23}) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$[X,[X,[X,A]]] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Now we know that each element $ad(X)$, for X in $\hat{g}$, acting on $\hat{g}$ by the adjoint action, is a nilpotent linear transformation in $End(\hat{g})$. Thus we can use induction on the dimension of $\hat{g}$. Now $\hat{h}$ has dimension less than $\hat{g}$, and thus $ad(\hat{h})$ has dimension less than $ad(\hat{g})$. And we have shown in the proof that if $\hat{h}$ is an ideal, then $ad(\hat{g})$ acts on $\hat{g}/\hat{h}$, and thus $ad(\hat{h})$ acts on $\hat{g}/\hat{h}$. The conclusion of Engel's Theorem says that there exists a $\overline{Y}$ in $\hat{g}/\hat{h}$ which is not equal to zero but is an eigenvector with eigenvalue zero for all $ad(Z)$ in $ad(\hat{h})$. Unwinding this quotient, we assert that there exist a $Y \neq 0$ in $\hat{g}$ which is not in $\hat{h}$ such that for all Z in $\hat{h}$, $(ad(Z))(Y)$ is in $\hat{h}$. The matrix Y is given by the matrix E_{12}, which is not in $\hat{h}$, and the span of Y is represented by matrices of the form

$$\begin{bmatrix} 0 & a_{12} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Also for any Z in $\hat{h}$, $(ad(Z))(Y) = [Z,Y] = ZY - YZ$ is in $\hat{h}$:

$$\left[\begin{bmatrix} 0 & 0 & a_{13} & a_{14} \\ 0 & 0 & a_{23} & a_{24} \\ 0 & 0 & 0 & a_{34} \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right] = \begin{bmatrix} 0 & 0 & -a_{23} & -a_{24} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus we have shown that there exists an ideal $\hat{h}$ in $\hat{g}$ of codimension one, and a vector Y in $\hat{g}$ not in $\hat{h}$ such that $\hat{g} = \hat{h} \oplus sp(Y)$. This is the conclusion of the first part of the proof of Engel's Theorem.

The last part of the proof of Engel's Theorem used induction again. But now we focus on the ideal $\hat{h}$ identified above. This subalgebra has dimension one less than that of $\hat{g}$. Thus, using Engel's Theorem, with $\hat{h}$ acting on the linear space V, we identified a vector $u \neq 0$ in V such that for every Z in $\hat{h}$, $Z(u) = 0$. Since Z has the matrix form written on the basis (v_1, v_2, v_3, v_4) of V) of

$$\begin{bmatrix} 0 & 0 & a_{13} & a_{14} \\ 0 & 0 & a_{23} & a_{24} \\ 0 & 0 & 0 & a_{34} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

we see immediately that the subspace W of V such that $Z(W) = 0$ for all Z in $\hat{h}$ has the basis (v_1, v_2). But we also observe that the matrix Y identified above has the property of leaving invariant the space W since:

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Thus $Y(v_1) = 0$ and $Y(v_2) = v_1$. Since $\hat{g} = \hat{h} \oplus sp(Y)$, it is now easy to identify a vector $u \neq 0$ in V such that for all X in $\hat{g}$, $X(u) = 0$. Since $\hat{h}(W) = 0$, we need only examine $Y(v_1)$ and $Y(v_2)$. If we choose v_1, we see that $Y(v_1) = 0$ and we are finished. If we choose v_2, then $Y(v_2) = v_1$. But we know that $Y \in \hat{g}$ is linear nilpotent acting on V. In fact, since Y is equal to E_{12}, we have

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Then $Y^2(v_2) = 0 = Y(Y(v_2)) = Y(v_1)$. Thus once again we see that we can choose the vector $v_1 \neq 0$ in V which is a simultaneous eigenvector with eigenvalue 0 for all X in $\hat{g}$. And this is the conclusion of Engel's Theorem. Obviously we are dealing once again the matrices of the form

$$\begin{bmatrix} 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Changing focus, we now assume, if possible, that a 4-dimensional abstract nilpotent Lie algebra exists, called $\hat{n}$. Then given a basis of $\hat{n}$, the adjoint representation ad takes $\hat{n}$ into the 4x4 matrices $\widehat{gl}(\hat{n})$. We remark that if we let $\hat{n} = V$, then V is 4-dimensional and thus we are in the above context of $\widehat{gl}(V)$. From above we know that we have a nilpotent Lie subalgebra $\hat{g} = ad(\hat{n})$ of $\widehat{gl}(\hat{n})$, and with respect to the basis (v_1, v_2, v_3, v_4), the elements of this subalgebra are the upper triangular matrices with a zero diagonal. Now we would like to identify the images by ad of the four vectors (v_1, v_2, v_3, v_4) of $\hat{n}$ as matrices in $\hat{g}$. We therefore choose v_1 as a simultaneous eigenvector with eigenvalue 0 for all X in $\hat{g} = ad(\hat{n})$. Thus we have $[v_1, x] = 0$:

$$\begin{gathered}
(ad(v_1))(v_1) = [v_1, v_1] = 0 \\
(ad(v_1))(v_2) = [v_1, v_2] = -[v_2, v_1] = -(ad(v_2))(v_1) = 0 \\
(ad(v_1))(v_3) = [v_1, v_3] = -[v_3, v_1] = -(ad(v_3))(v_1) = 0 \\
(ad(v_1))(v_4) = [v_1, v_4] = -[v_4, v_1] = -(ad(v_4))(v_1) = 0
\end{gathered}$$

We see that $ad(v_1)$ is the zero matrix. This just says that v_1 is in the center of $\hat{n}$. We also know that the center of $\hat{n}$ is the kernel of the homomorphism ad. Continuing, we have

$$\begin{gathered}
(ad(v_2))(v_1) = [v_2, v_1] = 0 \\
(ad(v_2))(v_2) = [v_2, v_2] = 0 \\
(ad(v_2))(v_3) = [v_2, v_3] = a_{13}v_1 + a_{23}v_2 \\
(ad(v_2))(v_4) = [v_2, v_4] = a_{14}v_1 + a_{24}v_2 + a_{34}v_3
\end{gathered}$$

since we know that $ad(v_2)$ is an upper triangular matrix with zero diagonal. Continuing,

$$\begin{gathered}
(ad(v_3))(v_1) = [v_3, v_1] = 0 \\
(ad(v_3))(v_2) = [v_3, v_2] = b_{12}v_1 \\
(ad(v_3))(v_3) = [v_3, v_3] = 0 \\
(ad(v_3))(v_4) = [v_3, v_4] = b_{14}v_1 + b_{24}v_2 + b_{34}v_3
\end{gathered}$$

$$\begin{gathered}
(ad(v_4))(v_1) = [v_4, v_1] = 0 \\
(ad(v_4))(v_2) = [v_4, v_2] = c_{12}v_1 \\
(ad(v_4))(v_3) = [v_4, v_3] = c_{13}v_1 + c_{23}v_2 \\
(ad(v_4))(v_4) = [v_4, v_4] = 0
\end{gathered}$$

Now using the relation $[u, v] = -[v, u]$, we see

$$b_{12} = -a_{13} \qquad a_{23} = 0 \qquad c_{12} = -a_{14} \qquad a_{24} = 0 \qquad a_{34} = 0$$
$$c_{13} = -b_{14} \qquad c_{23} = -b_{24} \qquad b_{34} = 0$$

Thus we have

$$\begin{aligned}(ad(v_2))(v_1) &= [v_2, v_1] = 0\\(ad(v_2))(v_2) &= [v_2, v_2] = 0\\(ad(v_2))(v_3) &= [v_2, v_3] = a_{13}v_1\\(ad(v_2))(v_4) &= [v_2, v_4] = a_{14}v_1\end{aligned}$$

$$\begin{aligned}(ad(v_3))(v_1) &= [v_3, v_1] = 0\\(ad(v_3))(v_2) &= [v_3, v_2] = -a_{13}v_1\\(ad(v_3))(v_3) &= [v_3, v_3] = 0\\(ad(v_3))(v_4) &= [v_3, v_4] = b_{14}v_1 + b_{24}v_2\end{aligned}$$

$$\begin{aligned}(ad(v_4))(v_1) &= [v_4, v_1] = 0\\(ad(v_4))(v_2) &= [v_4, v_2] = -a_{14}v_1\\(ad(v_4))(v_3) &= [v_4, v_3] = -b_{14}v_1 - b_{24}v_2\\(ad(v_4))(v_4) &= [v_4, v_4] = 0\end{aligned}$$

Therefore the matrices are:

$$ad(v_1) = \begin{bmatrix} 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix},\ ad(v_2) = \begin{bmatrix} 0 & 0 & a_{13} & a_{14}\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix},$$

$$ad(v_3) = \begin{bmatrix} 0 & -a_{13} & 0 & b_{14}\\ 0 & 0 & 0 & b_{24}\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix},\ ad(v_4) = \begin{bmatrix} 0 & -a_{14} & -b_{14} & 0\\ 0 & 0 & -b_{24} & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}$$

We now want to check the homomorphism between $\hat{n}$ and this Lie subalgebra of $\hat{g}$. This means that we have to verify the following relation for all i and j:

$$ad[v_i, v_j] = [ad(v_i), ad(v_j)]$$

In the following calculations we first give the bracket in $\hat{n}$, and then map this element over to $ad(\hat{n})$. This is followed by mapping each factor of the bracket to $ad(\hat{n})$, and then calculating the bracket in $ad(\hat{n})$. The two calculations should give the same answer if we have a homomorphism.

$$[v_1, v_2] = 0 \longrightarrow ad([v_1, v_2]) = ad(0) = \begin{bmatrix} 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}$$

$$\left[\begin{bmatrix} 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}, \begin{bmatrix} 0 & 0 & a_{13} & a_{14}\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}\right] = \begin{bmatrix} 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}$$

$$[v_1, v_3] = 0 \longrightarrow ad([v_1, v_3]) = ad(0) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\left[\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -a_{13} & 0 & b_{14} \\ 0 & 0 & 0 & b_{24} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$[v_1, v_4] = 0 \longrightarrow ad([v_1, v_4]) = ad(0) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\left[\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -a_{14} & -b_{14} & 0 \\ 0 & 0 & -b_{24} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$[v_2, v_3] = a_{13}v_1 \longrightarrow ad([v_2, v_3]) = ad(a_{13}v_1) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\left[\begin{bmatrix} 0 & 0 & a_{13} & a_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -a_{13} & 0 & b_{14} \\ 0 & 0 & 0 & b_{24} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$[v_2, v_4] = a_{14}v_1 \longrightarrow ad([v_2, v_4]) = ad(a_{14}v_1) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\left[\begin{bmatrix} 0 & 0 & a_{13} & a_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -a_{14} & -b_{14} & 0 \\ 0 & 0 & -b_{24} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$[v_3, v_4] = b_{14}v_1 + b_{24}v_2 \longrightarrow ad([v_3, v_4]) = ad(b_{14}v_1 + b_{24}v_2) =$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + b_{24} \begin{bmatrix} 0 & 0 & a_{13} & a_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = b_{24} \begin{bmatrix} 0 & 0 & a_{13} & a_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\left[\begin{bmatrix} 0 & -a_{13} & 0 & b_{14} \\ 0 & 0 & 0 & b_{24} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -a_{14} & -b_{14} & 0 \\ 0 & 0 & -b_{24} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right] = \begin{bmatrix} 0 & 0 & a_{13}b_{24} & a_{14}b_{24} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} =$$

$$b_{24}\begin{bmatrix} 0 & 0 & a_{13} & a_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus brackets go into brackets by ad and the homomorphism is verified. And thus our assumption that we have a 4-dimensional nilpotent Lie algebra is justified.

We also observe that the three matrices

$$\begin{bmatrix} 0 & 0 & a_{13} & a_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -a_{13} & 0 & b_{14} \\ 0 & 0 & 0 & b_{24} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -a_{14} & -b_{14} & 0 \\ 0 & 0 & -b_{24} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

form a linearly independent set. Thus we have

$$\dim(\hat{n}) = \dim(\ker\, ad) + \dim(\text{image}\; ad)$$
$$4 = 1 + 3$$

and we observe that we have a 3-dimensional Lie subalgebra in the 6-dimensional Lie algebra $\hat{g}$ of all upper triangular matrices with diagonal zero in $\widehat{gl}(\hat{n})$. The kernel of the above map gives the center of the Lie algebra $\hat{n}$, which is $sp\{v_1\}$. Since we also know that $ad(\hat{n})$ is nilpotent in $\hat{g}$, we can affirm that $ad(\hat{n})$ also has a nontrivial center. From the above calculations of the bracket of $ad(\hat{n})$, we see that the matrix

$$\begin{bmatrix} 0 & 0 & a_{13} & a_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

and only this matrix has zero brackets with all other elements of $ad(\hat{n})$. Thus this matrix generates the one-dimensional center of $ad(\hat{n})$. We also see that

$$C^0(ad(\hat{n})) = ad(\hat{n});$$
$$C^1(ad(\hat{n})) = [ad(\hat{n}), ad(\hat{n})] = \{c(a_{13}E_{13} + a_{14}E_{14})\};$$
$$C^2(ad(\hat{n})) = 0$$

and thus the center of $ad(\hat{n}) = C^1(ad(\hat{n})) = sp\{E_{13}, E_{14}\}$, and

$$\dim C^0(ad(\hat{n})) = 3;\; \dim C^1(ad(\hat{n})) = 2;\; \dim C^2(ad(\hat{n})) = 0$$

2.9 Lie's Theorem

2.9.1 Some Remarks on Lie's Theorem. Recall that we took this excursion into nilpotent Lie algebras because we wanted to identify the automorphism A between the semisimple part of two Levi decompositions of a Lie algebra $\hat{g}$, where a Levi decomposition is a splitting at the level of linear spaces of an arbitrary Lie algebra $\hat{g}$ into its radical $\hat{r}$ and a semisimple Lie algebra $\hat{k}$.

$$\hat{g} = \hat{k} \oplus \hat{r}$$

Since the radical $\hat{r}$ is the maximal solvable ideal of the Lie algebra $\hat{g}$, let us take another excursion into solvable Lie algebras and prove a theorem comparable to Engel's Theorem for solvable Lie algebras. This is the famous Lie's Theorem. It reads:

> Let $\hat{s}$ be a solvable complex Lie subalgebra of $\widehat{gl}(V)$. Then there exists a nonzero vector $v \in V$ which is a simultaneous eigenvector for all X in $\hat{s}$ [with eigenvalue dependent on X].

Once again we are in the context of the 19th century and thus Lie's Theorem will be about matrices. Let V be a finite dimensional linear space and consider the Lie algebra $\widehat{gl}(V)$ of all endomorphisms of V. We are looking once again for simultaneous eigenvectors for some endomorphisms, but now the eigenvalue is not necessarily equal to zero. Thus immediately we must restrict ourselves to the algebraically closed field of scalars of characteristic 0 — in our case the complex numbers — in order to insure that the characteristic polynomial of any matrix can be factored linearly. Thus V must be a complex vector space and $\widehat{gl}(V)$ the complex endomorphisms of V. And the endomorphisms which give this property of having simultaneous eigenvectors will be a solvable complex Lie subalgebra of $\widehat{gl}(V)$.

We again make the following remarks, this time as we compare Lie's Theorem with Engel's Theorem. As we said above, the field of scalars for the linear space and the endomorphisms must be the complex numbers. Also the subalgebra in the Lie's Theorem is given in terms of Lie algebras, i.e., solvable, rather than as in the Engel's Theorem in terms of a property of linear transformations, i.e., every element of $\hat{g}$ is a nilpotent linear transformation. However in the context of the 19th century this was rather natural. We are now working essentially with matrices, and thus with $End(V)$. But $End(V)$ has a natural Lie algebra structure if the bracket is defined to be the commutator. In our notation this gives $End(V) = \widehat{gl}(V)$. Now in this context it is natural to define the derived series, and the situation in which the derived series of $\widehat{gl}(V)$ ends in zero should be a valuable property. Thus solvable Lie algebras become natural objects for consideration.

Now if we assume that Lie's Theorem is true, it is evident that for every solvable complex Lie algebra of linear transformations $\hat{s}$ we can find a basis for V such that every element in $\hat{s}$ can be represented on this basis by an upper triangular matrix. The proof is essentially the same as the proof given above that nilpotent Lie algebras of linear transformations can be represented by upper triangular matrices with a zero diagonal. We will not repeat the details. But one immediate conclusion is that every nilpotent Lie algebra is also solvable [which fact we already know]. But we also know that every solvable Lie algebra is not necessarily nilpotent. We give some calculations here so that we can get a feel for these statements. Let $\hat{s}$ be the solvable Lie subalgebra of upper triangular matrices in $\widehat{gl}(\mathbb{R}^4)$. We choose four matrices in $\hat{s}$: A, B, F. G:

$$A = \begin{bmatrix} 1 & 2 & -3 & 1 \\ 0 & -1 & 1 & -2 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & -2 \end{bmatrix} \quad B = \begin{bmatrix} -1 & 0 & -3 & 1 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & -2 & 3 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

$$F = \begin{bmatrix} 2 & 2 & 3 & 0 \\ 0 & -1 & 1 & -2 \\ 0 & 0 & -2 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad G = \begin{bmatrix} 2 & 2 & -3 & 0 \\ 0 & -2 & 1 & -2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Next we take the bracket $[A, B]$, which is in $D^1\hat{s} = C^1\hat{s}$:

$$[A, B] = \begin{bmatrix} 0 & 4 & 8 & 2 \\ 0 & 0 & -6 & -4 \\ 0 & 0 & 0 & 24 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

We see immediately that $[A, B]$ is an upper triangular matrix with a zero diagonal. We will show later that, in general, $D^1\hat{s}$ is in the nilpotent Lie subalgebra of upper triangular matrices with zero diagonal in $\widehat{gl}(\mathbb{R}^4)$. For now, however, we continue building up the lower central series for $\hat{s}$. Thus $[F, [A, B]]$ is in $C^2\hat{s}$ and:

$$[F, [A, B]] = \begin{bmatrix} 0 & 12 & 16 & 50 \\ 0 & 0 & -6 & 50 \\ 0 & 0 & 0 & -72 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

We observe that these matrices in $C^2\hat{s}$ have the same form as those in $C^1\hat{s}$, and we can show that this is true in general. Thus we can conclude that $C^k\hat{s} = 0$ for some k will not occur. Of course, this means that $\hat{s}$ is not a nilpotent Lie algebra. We now compute the brackets $[F, G]$ and $[[A, B], [F, G]]$:

$$[F,G] = \begin{bmatrix} 0 & -2 & -15 & 18 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 0 & -15 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad [[A,B],[F,G]] = \begin{bmatrix} 0 & 0 & 4 & 240 \\ 0 & 0 & 0 & -6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

As expected, $[F, G]$ is upper triangular with zero diagonal — an element in the nilpotent Lie subalgebra. Thus $[[A, B], [F, G]]$, which is in $D^2\hat{s}$, is a bracketing of two elements in a nilpotent Lie algebra, which action we know will push the bracket down in its lower central series. Continuing this process, we will reach the trivial Lie algebra 0. Thus we can conclude that for some k, $D^k\hat{s} = 0$ and we see that $\hat{s}$ is a solvable Lie algebra.

2.9.2 Proof of Lie's Theorem. We now give the proof of Lie's Theorem. In doing so, we will follow closely the proof of Engel's Theorem. [Once again we will be using induction on the dimension of $\hat{s}$.] When the dimension of $\hat{s}$ is one, we are dealing essentially with one nonzero linear transformation X in $\widehat{gl}(V)$. Now since we are in the context of the field of complex scalars, we know that X has an eigenvector v with eigenvalue λ in $\mathbb{C}$: $X(v) = \lambda \cdot v$. This fact proves the theorem in this case.

We now let $\hat{s}$ be a solvable complex Lie subalgebra of $\widehat{gl}(V)$ of dimension greater than one. First we want to find an ideal $\hat{h}$ of $\hat{s}$ of codimension one. Then this will give us a linear transformation Y in $\hat{s}$ whose span $sp(Y)$ is not in $\hat{h}$, giving $\hat{s} = \hat{h} \oplus sp(Y)$. In doing this we will also describe some interesting auxiliary structures of importance in mathematics. This time it is relatively easy to find $\hat{h}$. Since the condition on $\hat{s}$ is given in terms of Lie algebra structures, we take advantage of these structures. We choose the derived algebra of $\hat{s}$, the ideal $D^1\hat{s}$. We know $D^1\hat{s} \neq \hat{s}$, for if $D^1\hat{s} = \hat{s}$, then it would be impossible to find a k such that $D^k\hat{s} = 0$, and $\hat{s}$ would not be solvable. Now if $D^1\hat{s} = 0$, then $\hat{s}$ is abelian and thus diagonalizable and therefore satisfies the theorem. Thus we can assume that the dimension of $D^1\hat{s} \neq 0$ and is less than the dimension of $\hat{s}$. We now form the nonzero quotient $\hat{s}/D^1\hat{s}$, calling α the map from $\hat{s}$ to $\hat{s}/D^1\hat{s}$. But this process of moding out the subalgebra of commutators, which is what $D^1\hat{s}$ is, gives the abelianization of $\hat{s}$, i.e., $\hat{s}/D^1\hat{s}$ is an abelian Lie algebra. Here is the proof:

Taking the brackets of any two cosets $\overline{Y_1}$ and $\overline{Y_2}$, we have $[\overline{Y_1}, \overline{Y_2}]$. Writing this bracket in cosets language gives

$$[Y_1 + D^1\hat{s}, Y_2 + D^1\hat{s}] = [Y_1, Y_2] + [Y_1, D^1\hat{s}] + [D^1\hat{s}, Y_2] + [D^1\hat{s}, D^1\hat{s}]$$

Now $[Y_1, Y_2]$ is in $D^1\hat{s}$; and since $D^1\hat{s}$ is an ideal, the other three terms are also in $D^1\hat{s}$. We conclude that this bracket is in $D^1\hat{s}$, which is the zero coset. Thus $[\overline{Y_1}, \overline{Y_2}] = 0$. This gives us an abelian quotient Lie algebra.

Now we choose a codimension one subspace of $\hat{s}/D^1\hat{s}$. [We remark that if the dimension of $\hat{s}/D^1\hat{s}$ is one, then

$$\hat{s} = ker(\alpha) \oplus \alpha^{-1}(im(\alpha)) = D^1\hat{s} \oplus \hat{l}$$

where the dimension of $\hat{l}$ is one, giving $\hat{l} = sp(Y)$ where $Y \neq 0$ is in $\hat{s}$ but not in $D^1\hat{s}$. Since $D^1\hat{s}$ is an ideal, we have in this case effected the desired decomposition $\hat{s} = \hat{h} \oplus sp(Y)$.] Continuing, we have

$$\hat{s}/D^1\hat{s} = \hat{k}/D^1\hat{s} \oplus \hat{l}/D^1\hat{s}$$

where the dim $(\hat{s}/D^1\hat{s}) \geq 2$, dim $(\hat{k}/D^1\hat{s}) \geq 1$, and $\dim(\hat{l}/D^1\hat{s}) = 1$. This, of course, says that $\hat{k} + \hat{l} = \hat{s}$ and $\hat{k} \cap \hat{l} = D^1\hat{s}$. We now look at the dimensions. Let dim $\hat{s} = n$ and dim $D^1\hat{s} = m$. Then we have

$$\begin{aligned} n - m &= \dim \hat{k} - m + m + 1 - m \\ n &= \dim \hat{k} + 1 \end{aligned}$$

which says that $\hat{k}$ has codimension 1. Now $D^1\hat{s} \subset \hat{k}$. Thus we have the map

$$\hat{k} \xrightarrow{\alpha} \hat{k}/D^1\hat{s} \subset \hat{s}/D^1\hat{s}$$

Now since $\hat{s}/D^1\hat{s}$ is abelian, we have

$$[\hat{k}/D^1\hat{s}, \hat{s}/D^1\hat{s}] = 0 \subset \hat{k}/D^1\hat{s}$$

thus making $\hat{k}/D^1\hat{s}$ an ideal in $\hat{s}/D^1\hat{s}$. Now

$$\alpha[\hat{k}, \hat{s}] = [\alpha(\hat{k}), \alpha(\hat{s})] \subset \hat{k}/D^1\hat{s}$$

Thus

$$[\hat{k}, \hat{s}] \subset \alpha^{-1}(\hat{k}/D^1\hat{s}) = \hat{k}$$

which says that $\hat{k}$ is an ideal in $\hat{s}$. Thus we have our conclusion that $\hat{s}$ has a codimension 1 ideal $\hat{k}$. And thus we have $\hat{s} = \hat{h} \oplus sp(Y)$ with $Y \neq 0$ in $\hat{s}$ but with $sp(Y)$ not being in the ideal $\hat{h}$.

Again, following the proof of Engel's Theorem, we will do an induction on the dimension of $\hat{s}$. Since $\hat{s}$ is solvable, we know that the ideal $\hat{h}$ is also solvable and of dimension less than the dimension of $\hat{s}$. Thus, by induction, there exists a vector u of V which is a simultaneous eigenvector for all matrices Z in $\hat{h}$ with complex eigenvalue $\lambda(Z)$ dependent, of course, on Z, i.e., $Z(u) = \lambda(Z)(u)$. We observe that this determines an element λ in the dual $\hat{h}^*$ of $\hat{h}$.

Following the proof of Engel's Theorem we define a subspace W of V such that $w \in W$ means w is also a simultaneous eigenvector for all Z in $\hat{h}$ with complex eigenvalue $\lambda(Z)$, i.e., $Z(w) = \lambda(Z)(w)$. We remark that we have not changed the dual element in defining this subset W. Thus W is the eigenspace corresponding to the eigenvalue λ. Obviously u belongs to W.

We want to show also that Y stabilizes W, i.e., leaves W invariant. We proceed as follows. For all w in W, $Z(Y(w)) = \lambda(Z)(Y(w))$ for all Z in $\hat{h}$, where $\lambda(Z)$ is a complex eigenvalue for the simultaneous eigenvector $Y(w)$. Again let us take brackets: $Z(Y(w)) = Y(Z(w)) - [Y,Z](w)$. Since $\hat{h}$ is an ideal, we know that $[Y,Z]$ is in $\hat{h}$ and thus $[Y,Z](w) = \lambda([Y,Z])w$. Now $Y(Z(w)) = Y(\lambda(Z)w)$. Thus we have

$$Z(Y(w)) = Y(\lambda(Z)w) - \lambda([Y,Z])w = \lambda(Z)(Y(w)) - \lambda([Y,Z])w$$

This relation says that we must show $\lambda([Y,Z]) = 0$ in order for us to say that Y stabilizes W.

Thus we must find a way to evaluate the dual λ on $\hat{h}$. In doing so, we observe two facts. First Z operating on $Y(w)$ gives a linear combination in the subspace generated by w and $Y(w)$, and secondly the scalars are eigenvalues coming from the constant dual λ. Thus this suggests, as in the proof of Engel's Theorem, that we take iterates of Y acting on w: $(w, Y(w), Y^2(w), Y^3(w), \cdots, Y^k(w))$. [In Engel's Theorem, because of nilpotency, we knew that these iterates would arrive at zero, and this gave us the proof of Engel's Theorem. Now we lack the property of nilpotency but we can gain important information once again from these iterates. We know that $(w, Y(w), Y^2(w), Y^3(w), \cdots, Y^k(w))$ will be a maximal independent set of vectors in V for some k. We note that w is in W, and thus we are asking what happens when Y acts on W. Now either $Y(w)$ is in W or is not in W. If $Y(w)$ is in W, then Y stabilizes W and U, (the $Span(w, Y(w), Y^2(w), Y^3(w), \cdots, Y^k(w))$ is W itself. If not, then we ask whether $(w, Y(w))$ is stabilized by Y. But this says that we are now looking at the set $(w, Y(w), Y^2(w))$. If $Span(w, Y(w), Y^2(w)) = Span(w, Y(w))$, then $Y^2(w)$ is not independent, and we are finished. Thus we have the above conclusion that $(w, Y(w), Y^2(w), Y^3(w), \cdots, Y^k(w))$ will be a maximal independent set of vectors in V for some k. Thus they will span a subspace U of V of dimension $k+1 \le n$, where n is the dimension of V. It is evident that Y stabilizes U. And we remark that induction shows that $\hat{h}$ also stabilizes this subspace U. [We note that if $k = 0$, then the conclusion is trivial.] We assume that $Z((Y^{k-1})(w))$ is in U for all Z in $\hat{h}$. Then

$$Z((Y^k)(w)) = Z(Y((Y^{k-1})(w))) = Y(Z((Y^{k-1})(w))) - [Y, Z]((Y^{k-1})(w))$$

By induction $Z((Y^{k-1})(w))$ is in U and Y, operating on U, remains in U. Thus $Y(Z((Y^{k-1})(w)))$ is in U. Also $[Y, Z]$ is in $\hat{h}$ and thus $[Y, Z]((Y^{k-1})(w))$ is in U. We conclude that $\hat{h}$ stabilizes U. Indeed this means that $\hat{s}$ also stabilizes U.

We now let $l \leq k$. We observe that when $l = 1$, the above equality gives

$$Z(Y(w)) = Y(\lambda(Z)w) - \lambda([Y, Z])w = \lambda(Z)(Y(w)) - \lambda([Y, Z])w$$

which repeats the first equality given above. Continuing, we let $l = 2$. We have

$$Z((Y^2)(w)) = Z(Y(Y(w))) = Y(Z(Y(w))) - [Y, Z](Y(w)) =$$
$$\lambda(Z)(Y(Y(w))) - \lambda([Y, Z])(Y(w)) = \lambda(Z)(Y^2(w)) - \lambda([Y, Z])(Y(w))$$

We see the pattern developing. We are producing, for each Z in $\hat{h}$, with respect to the basis for U: $(w, Y(w), Y^2(w), Y^3(w), \cdot \cdot \cdot, Y^k(w))$, an upper triangular matrix for Z with constant diagonal entries equal to $\lambda(Z)$. This means that if we take the trace of this matrix, we obtain

$$trace(Z) = (k + 1) \cdot \lambda(Z)$$

We conclude that we have found a method of calculating the dual λ.

We return to the question of whether Y stabilizes the subspace W of all simultaneous eigenvectors under the action of $\hat{h}$. We know, when w is any element in W, that

$$Z(Y(w)) = \lambda(Z)(Y(w)) - \lambda([Y, Z])w$$

We now have a method of calculating $\lambda([Y, Z])$. Since $[Y, Z]$ is in $\hat{h}$, we have $trace([Y, Z]) = (k+1) \cdot \lambda([Y, Z])$. But the trace of any commutator $[Y, Z]$ in $\hat{h}$ is

$$trace([Y, Z]) = trace(YZ - ZY) = trace(YZ) - trace(ZY) = 0 =$$
$$(k + 1) \cdot \lambda([Y, Z])$$

since the $trace(YZ) = trace(ZY)$. This gives the conclusion that $\lambda([Y, Z]) = 0$, and finally we have $Z(Y(w)) = \lambda(Z)(Y(w))$, which says that $Y(w)$ is a simultaneous eigenvector for all Z in $\hat{h}$ with eigenvalue $\lambda(Z)$. Thus Y stabilizes W.

[It is interesting to remark that since $\lambda([Y, Z]) = 0$ whenever the bracket product is in $\hat{h}$, we can conclude that for each Z in $\hat{h}$, with respect to the basis for U: $(w, Y(w), Y^2(w), Y^3(w), \cdots, Y^k(w))$, the upper triangular matrix for Z with constant diagonal entries equal to $\lambda(Z)$ becomes a diagonal matrix with constant entries equal to $\lambda(Z)$.]

Now we know that Y is a linear transformation taking W into W, and thus has an eigenvector $v \neq 0$ in W with eigenvalue $\lambda(Y)$ — we are extending the dual λ now from $\hat{h}$ to $\hat{s}$ — since our field of scalars is $\mathbb{C}$. And since $v \in W$, it is also a simultaneous eigenvector for all Z in $\hat{h}$. Thus we have found a vector $v \in V$, $v \neq 0$, which is a simultaneous eigenvector for all X in $\hat{s}$. The proof is then complete.

2.9.3 Examples. An example again is enlightening. Let us use a V with dimension equal to 4, and give it a basis (v_1, v_2, v_3, v_4). Then $End(V) = \widehat{gl}(V)$ is the set of 4x4 matrices over $\mathbb{C}$. We look at the following set of upper triangular matrices $\hat{s}$ over $\mathbb{C}$:

$$X = \begin{bmatrix} a_{11} & a_{12} & a_{13} & 0 \\ 0 & 0 & a_{23} & 0 \\ 0 & 0 & a_{33} & 0 \\ 0 & 0 & 0 & a_{44} \end{bmatrix}$$

First let us show that $\hat{s}$ is a solvable Lie subalgebra. We see that $\hat{s}$ is 6-dimensional. We take the brackets in $\hat{s}$. Since the basis of $\hat{s}$ is:

$$(E_{11}, E_{12}, E_{13}, E_{23}, E_{33}, , E_{44})$$

we have 15 different such products:

$$\begin{array}{c} [E_{11}, E_{12}] = E_{12} \quad [E_{11}, E_{13}] = E_{13} \quad [E_{11}, E_{23}] = 0 \quad [E_{11}, E_{33}] = 0 \\ [E_{11}, E_{44}] = 0 \quad [E_{12}, E_{13}] = 0 \quad [E_{12}, E_{23}] = E_{13} \quad [E_{12}, E_{33}] = 0 \\ [E_{12}, E_{44}] = 0 \quad [E_{13}, E_{23}] = 0 \quad [E_{13}, E_{33}] = E_{13} \quad [E_{13}, E_{44}] = 0 \\ [E_{23}, E_{33}] = E_{23} \quad [E_{23}, E_{44}] = 0 \quad [E_{33}, E_{44}] = 0 \end{array}$$

Thus we see that the brackets close and we do have a Lie subalgebra. Also we have as a basis for $D^1\hat{s} = [\hat{s}, \hat{s}]$:

$$(E_{12}, E_{13}, E_{23})$$

We now take the brackets in $D^1\hat{s}$:

$$[E_{12}, E_{13}] = 0 \quad [E_{12}, E_{23}] = E_{13} \quad [E_{13}, E_{23}] = 0$$

Thus we have as a basis for $D^2\hat{s} = [D^1\hat{s}, D^1\hat{s}]$:

$$(E_{13})$$

Finally we have $D^3\hat{s} = [D^2\hat{s}, D^2\hat{s}] = 0$. And thus we see that $\hat{s}$ is solvable.

Moreover, we have

$$\dim\, \hat{s} = 6 \quad \dim\, D^1\hat{s} = 3 \quad \dim\, D^2\hat{s} = 1 \quad \dim\, D^3\hat{s} = 0$$

It is evident that given these matrices we see immediately that v_1 is a simultaneous eigenvector with eigenvalue $\lambda_1(X) = a_{11}$ [where, of course, a_{11} depends on X]; and also that v_4 is a simultaneous eigenvector with eigenvalue $\lambda_4(X) = a_{44}$ [where a_{44} depends on X]. Thus we are identifying duals λ_1 and λ_4 in $\hat{s}^*$ with this property. However what we would like to do is to follow the proof of the theorem in this case and see how the proof identifies these vectors.

The first step in the proof of Lie's Theorem in this case is to identify the ideal $\hat{h}$ of codimension one in $\hat{s}$. To do this we take the quotient algebra $\hat{s}/D^1\hat{s}$. This means we are taking matrices of the form

$$\begin{bmatrix} * & * & * & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & 0 & * \end{bmatrix}$$

and moding out by the matrices of the form

$$\begin{bmatrix} 0 & * & * & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

which gives us the 3-dimensional abelian Lie algebra isomorphic to the diagonal matrices

$$\begin{bmatrix} * & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & 0 & * \end{bmatrix}$$

From these computations we can say that the quotient of $\hat{s}$ by the commutator $D^1\hat{s}$ abelianizes $\hat{s}$.

We now want to choose a codimension one subspace of the 3-dimensional space $\hat{s}/D^1\hat{s}$. At this point we have three choices, and we think it would be quite informative to see how the proof of Lie's Theorem picks a simultaneous eigenvector in each of these choices.

First we choose the two-dimensional subspace with basis $(E_{33}+D^1\hat{s}, E_{44}+D^1\hat{s})$ in $\hat{s}/D^1\hat{s}$, which consists of matrices of the form

$$\begin{bmatrix} 0 & * & * & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & a_{33} & 0 \\ 0 & 0 & 0 & a_{44} \end{bmatrix}$$

The inverse image of this subspace by the quotient map gives an ideal $\hat{h}_1$ in $\hat{s}$ of dimension 5, thus of codimension one. $\hat{h}_1$ consists of all matrices of the form

$$\begin{bmatrix} 0 & * & * & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & 0 & * \end{bmatrix}$$

that is, it has the basis

$$(E_{12}, E_{13}, E_{23}, E_{33}, E_{44})$$

We check that it is an ideal in $\hat{s}$. From above we see that the brackets of $\hat{h}_1$ do close forming a Lie subalgebra. The only non-zero brackets are:

$$[E_{12}, E_{23}] = E_{13} \quad [E_{13}, E_{33}] = E_{13} \quad [E_{23}, E_{33}] = E_{23}$$

To verify ideal structure we need only bracket E_{11} with the basis of $\hat{h}_2$. From above we see that the only non-zero brackets are:

$$[E_{11}, E_{12}] = E_{12} \quad [E_{11}, E_{13}] = E_{13}$$

We see that we have closure again in $\hat{h}_1$, giving the ideal structure. We also see that $\hat{s} = \hat{h}_1 \oplus sp(Y_1)$, where Y_1 is the one-dimensional subspace of $\hat{s}$ generated by the matrix E_{11}.

First we check that $\hat{h}_1$ is solvable. From above we see the $D^1\hat{h}_1 = [\hat{h}_1, \hat{h}_1]$ has as a basis: (E_{13}, E_{23}). We see immediately that $D^2\hat{h}_1 = [D^1\hat{h}_1, D^1\hat{h}_1] = 0$ Thus $\hat{h}_1$ is solvable.

By induction, we know that there exists a vector u_1 in V which is a simultaneous eigenvector with eigenvalue $\lambda_1(Z)$ for each Z in $\hat{h}_1$:, i.e., $Z(u_1) = \lambda_1(Z)(u_1)$. We see that for the choices that we made this vector u_1 is either v_1 or v_4. For the choice of v_1, we see that $\lambda_1(Z) = 0$ for all Z in $\hat{h}_1$; for the choice of v_4, we see that $\lambda_1(Z) = a_{44}$ for all Z in $\hat{h}_1$, with, of course, a_{44} depending on Z.

Before we proceed with the analysis of the proof of Lie's Theorem for these choices, let us bring up at this point the other two possible choices for the codimension one ideal of $\hat{s}$. We now want to choose another codimension one subspace of the 3-dimensional space $\hat{s}/D^1\hat{s}$. We choose now the two-dimensional subspace with basis $(E_{11} + D^1\hat{s}, E_{44} + D^1\hat{s})$ in $\hat{s}/D^1\hat{s}$, which consists of matrices of the form

$$\begin{bmatrix} a_{11} & * & * & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{44} \end{bmatrix}$$

The inverse image of this subspace by the quotient map gives an ideal $\hat{h}_2$ in $\hat{s}$ of dimension 5, thus of codimension one. $\hat{h}_2$ consists of all matrices of the form

$$\begin{bmatrix} * & * & * & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & * \end{bmatrix}$$

that is, it has the basis

$$(E_{11}, E_{12}, E_{13}, E_{23}, E_{44})$$

We check that it is an ideal in $\hat{s}$. From above we see that the brackets of $\hat{h}_2$ do close forming a Lie subalgebra. Moreover, the only non-zero brackets are:

$$[E_{11}, E_{12}] = E_{12} \quad [E_{11}, E_{13}] = E_{13} \quad [E_{12}, E_{23}] = E_{13}$$

To verify the ideal structure we need only bracket E_{33} with the basis of $\hat{h}_1$. From above, we see that the only non-zero brackets are:

$$[E_{13}, E_{33}] = E_{13} \quad [E_{23}, E_{33}] = E_{23}$$

We see that we have closure again in $\hat{h}_2$, giving the ideal structure. We also see that $\hat{s} = \hat{h}_2 \oplus sp(Y_2)$, where Y_2 is the one-dimensional subspace of $\hat{s}$ generated by the matrix E_{33}.

First we check that $\hat{h}_2$ is solvable. From above we see the $D^1\hat{h}_2 = [\hat{h}_2, \hat{h}_2]$ has as a basis: (E_{12}, E_{13}). We see immediately that $D^2\hat{h}_2 = [D^1\hat{h}_2, D^1\hat{h}_2] = 0$ Thus $\hat{h}_2$ is solvable.

By induction, we know that there exists a vector u_2 in V which is a simultaneous eigenvector with eigenvalue $\lambda_2(Z)$ for each Z in $\hat{h}_2$, i.e., $Z(u_2) = \lambda_2(Z)(u_2)$. We see that for the choices that we made that this vector u_2 is again either v_1 or v_4. For the choice of v_1, we see that $\lambda_2(Z) = a_{11}$ for all Z in $\hat{h}_2$ with a_{11} depending on Z; for the choice of v_4, we see that $\lambda_2(Z) = a_{44}$ for all Z in $\hat{h}_2$, with a_{44} depending on Z.

We now want to choose the last possible codimension one subspace of the 3-dimensional space $\hat{s}/D^1\hat{s}$. We choose the two-dimensional subspace with basis $(E_{11} + D^1\hat{s}, E_{33} + D^1\hat{s})$ in $\hat{s}/D^1\hat{s}$, which consists of matrices of the form

$$\begin{bmatrix} a_{11} & * & * & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & a_{33} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The inverse image of this subspace by the quotient map gives an ideal $\hat{h}_3$ in $\hat{s}$ of dimension 5, thus of codimension one. $\hat{h}_3$ consists of all matrices of the form

$$\begin{bmatrix} * & * & * & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

that is, it has the basis

$$(E_{11}, E_{12}, E_{13}, E_{23}, E_{33})$$

We check that it is an ideal in $\hat{s}$. From above we see that the brackets of $\hat{h}_3$ do close forming a Lie subalgebra. The only non-zero brackets are:

$$[E_{11}, E_{12}] = E_{12} \quad [E_{11}, E_{13}] = E_{13} \quad [E_{12}, E_{23}] = E_{13}$$
$$[E_{13}, E_{33}] = E_{13} \quad [E_{23}, E_{33}] = E_{23}$$

To verify ideal structure we need only bracket E_{44} with the basis of $\hat{h}_3$. From above we see that there are no non-zero brackets. We also see that we have closure in $\hat{h}_3$, giving the ideal structure. Moreover, $\hat{s} = \hat{h}_3 \oplus sp(Y_3)$, where Y_3 is the one-dimensional subspace of $\hat{s}$ generated by the matrix E_{44}.

First we check that $\hat{h}_3$ is solvable. From above, we see the $D^1\hat{h}_3 = [\hat{h}_3, \hat{h}_3]$ has as a basis: (E_{12}, E_{13}, E_{23}). We see $D^2\hat{h}_3 = [D^1\hat{h}_3, D^1\hat{h}_3]$ has a basis (E_{13}). Thus $D^3\hat{h}_3 = [D^2\hat{h}_3, D^2\hat{h}_3] = 0$ and $\hat{h}_3$ is solvable.

By induction, we know that there exists a vector u_3 in V which is a simultaneous eigenvector with eigenvalue $\lambda_3(Z)$ for each Z in $\hat{h}_3$, i.e., $Z(u_3) = \lambda_3(Z)(u_3)$. We see that, for the choices that we made, this vector u_3 is again either v_1 or v_4. For the choice of v_1, we see that $\lambda_3(Z) = a_{11}$ for all Z in $\hat{h}_3$ with a_{11} depending on Z. For the choice of v_4, we see that $\lambda_2(Z) = 0$ for all Z in $\hat{h}_3$.

Summarizing at this point, we have found three decompositions of $\hat{s}$:

$$\hat{s} = \hat{h}_1 \oplus sp(Y_1) \qquad \hat{s} = \hat{h}_2 \oplus sp(Y_2) \qquad \hat{s} = \hat{h}_3 \oplus sp(Y_3)$$

where

$$\hat{h}_1 = \left\{ \begin{bmatrix} 0 & * & * & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & 0 & * \end{bmatrix} \right\} \quad \hat{h}_2 = \left\{ \begin{bmatrix} * & * & * & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & * \end{bmatrix} \right\} \quad \hat{h}_3 = \left\{ \begin{bmatrix} * & * & * & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right\}$$

and where the Y_i are:

$$Y_1 = E_{11} \qquad Y_2 = E_{33} \qquad Y_3 = E_{44}$$

giving

$$\hat{s} = \hat{h}_1 \oplus sp(E_{11}) \qquad \hat{s} = \hat{h}_2 \oplus sp(E_{33}) \qquad \hat{s} = \hat{h}_3 \oplus sp(E_{44})$$

The possible simultaneous eigenvectors u_i for each $\hat{h}_i$ with their respective eigenvalues λ_i are:

$$\{v_1, v_4\} \text{ for } \hat{h}_1 \text{ with } Z(v_1) = \lambda_1(Z)(v_1) = 0;\ Z(v_4) = \lambda_1(Z)(v_4) = a_{44}v_4$$
$$\{v_1, v_4\} \text{ for } \hat{h}_2 \text{ with } Z(v_1) = \lambda_2(Z)(v_1) = a_{11}v_4;\ Z(v_4) = \lambda_2(Z)(v_4) = a_{44}v_4$$
$$\{v_1, v_4\} \text{ for } \hat{h}_3 \text{ with } Z(v_1) = \lambda_3(Z)(v_1) = a_{11}v_4;\ Z(v_4) = \lambda_3(Z)(v_4) = 0$$

The amazing conclusion is that the same two eigenvectors are chosen no matter what the ideal $\hat{h}_i$ is. To conclude the proof of Lie's Theorem all we have to do is show that each Y_i also acts on v_1 or v_4 as an eigenvector, that is, $Y_i(v_j) = \lambda_i(Y_i)(v_j)$. We have

$$E_{11}(v_1) = v_1; E_{11}(v_4) = 0$$
$$E_{33}(v_1) = 0; E_{33}(v_4) = 0$$
$$E_{44}(v_1) = 0; E_{44}(v_4) = v_4$$

and thus we can conclude that either v_1 or v_4 satisfies the conclusion of the theorem.

What we would like to do now in order to conclude the analysis of Lie's Theorem is to see how the reasoning in general was able to come to this same conclusion. At this point we identify the subspace W of V to be the subspace defined by all elements of V such that for each w in W, w is a simultaneous eigenvector for all Z in the ideal $\hat{h}$, i.e., $Z(w) = \lambda(Z)(w)$. We recall that W is just the eigenspace corresponding to each eigenvalue λ. In all six of the cases above, two for each $\hat{h}_i$, this W is one-dimensional, i.e., $W = sp(v_1)$ or $W = sp(v_4)$. Then we affirm that Y stabilizes this subspace, that is, $Y(w) \in W$, since., $Z(Y(w)) = \lambda(Z)(Y(w))$ for all Z in $\hat{h}$.

In our example we have:

$$Z(E_{11}(v_1)) = Z(v_1) = \lambda(Z)(v_1) = 0 \qquad \lambda(Z)(E_{11}(v_1)) = \lambda(Z)(v_1) = 0$$
$$Z(E_{11}(v_4)) = Z(0) = 0 \qquad \lambda(Z)(E_{11}(v_4)) = \lambda(Z)(0) = 0$$

$$Z(E_{33}(v_1)) = Z(0) = 0 \qquad \lambda(Z)(E_{33}(v_1)) = \lambda(Z)(0) = 0$$
$$Z(E_{33}(v_4)) = Z(0) = 0 \qquad \lambda(Z)(E_{33}(v_4)) = \lambda(Z)(0) = 0$$
$$Z(E_{44}(v_1)) = Z(0) = 0 \qquad \lambda(Z)(E_{44}(v_1)) = \lambda(Z)(0) = 0$$
$$Z(E_{44}(v_4)) = Z(v_4) = \lambda(Z)(v_4) = 0 \qquad \lambda(Z)(E_{44}(v_4)) = \lambda(Z)(v_4)) = 0$$

It is interesting to remark that the relation $Z(Y(w)) = \lambda(Z)(Y(w))$ gave the zero vector in W in all cases.

However, in the proof we had to work with the bracket product;

$$Z(Y(w)) = Y(\lambda(Z)w) - \lambda([Y,Z])w = \lambda(Z)(Y(w)) - \lambda([Y,Z])w$$

and we had to prove that $\lambda([Y,Z]) = 0$. Thus in the proof we were faced with the task of evaluating the dual λ on $\hat{h}$. This evaluation we did by taking iterates of w by Y, forming a linear subspace U of V which has the basis $(w, Y(w), Y^2(w), Y^3(w), \cdots, Y^k(w))$, with $k+1 \leq n$ and where n is the dimension of V. From this calculation we obtain

$$trace(Z) = (k+1) \cdot \lambda(Z)$$

which says that

$$\lambda(Z) = \frac{trace(Z)}{k+1}$$

Thus we have succeeded in evaluating the dual λ on $\hat{h}$. In our example we we see that $(w = Y_i^0(w))$ is a basis for U for all $\hat{h}_i$ for any $w = v_j$ in W, that is, the dimension of all the W's is one. In general our calculation is:

$$\begin{aligned} w &= v_j \\ Y(w) &= Y(v_j) \\ Y^2(w) &= Y^2(v_j) = Y(Y(v_j)) \end{aligned}$$

For $\hat{h}_1$, this gives

$$\begin{array}{rr} w = v_1 & w = v_4 \\ E_{11}(v_1) = v_1 & E_{11}(v_4) = 0 \end{array}$$

which show that after one iteration we begin to repeat. Continuing, we have for $\hat{h}_2$,

$$\begin{array}{rr} w = v_1 & w = v_4 \\ E_{33}(v_1) = 0 & E_{33}(v_4) = 0 \end{array}$$

which show that after one iteration we begin to repeat. For $\hat{h}_3$, this gives

$$\begin{array}{rr} w = v_1 & w = v_4 \\ E_{44}(v_1) = 0 & E_{44}(v_4) = v_4 \end{array}$$

which show that after one iteration we begin to repeat.

Now in order to write each Z in $\hat{h}_i$ in the basis (w) we need the following information:

$$\begin{array}{cc} Z(v_1) = \lambda_1(Z)(v_1) = 0 & Z(v_4) = \lambda_1(Z)(v_4) = a_{44}v_4 \\ Z(v_1) = \lambda_2(Z)(v_1) = a_{11}v_1 & Z(v_4) = \lambda_2(Z)(v_4) = a_{44}v_4 \\ Z(v_1) = \lambda_3(Z)(v_1) = a_{11}v_1 & Z(v_4) = \lambda_3(Z)(v_4) = 0 \end{array}$$

For $\hat{h}_i$, the basis for U is (v_1) or (v_4). On these bases we form the 1x1 matrices for Z.

$$\begin{array}{cccccc} \text{For } \hat{h}_1, & Z(v_1) = 0 & Z(v_4) = a_{44}v_4 & \text{giving matrices} & [0] & [a_{44}] \\ \text{For } \hat{h}_2, & Z(v_1) = a_{11}v_1 & Z(v_4) = a_{44}v_4 & \text{giving matrices} & [a_{11}] & [a_{44}] \\ \text{For } \hat{h}_3, & Z(v_1) = a_{11}v_1 & Z(v_4) = 0 & \text{giving matrices} & [a_{11}] & [0] \end{array}$$

Finally, we can calculate the traces of these matrices, which are, of course, trivial calculations. For $\hat{h}_1$ we have

$$Z(v_1) = \lambda(Z)(v_1) = 0 \qquad \lambda(Z) = \tfrac{trace(Z)}{1} = trace[0] = 0$$
$$Z(v_4) = \lambda(Z)(v_4) = a_{44}v_4 \qquad \lambda(Z) = \tfrac{trace(Z)}{1} = trace[a_{44}] = a_{44}$$

For $\hat{h}_2$ we have

$$Z(v_1) = \lambda(Z)(v_1) = a_{11}v_1 \qquad \lambda(Z) = \tfrac{trace(Z)}{1} = trace[a_{11}] = a_{11}$$
$$Z(v_4) = \lambda(Z)(v_4) = a_{44}v_4 \qquad \lambda(Z) = \tfrac{trace(Z)}{1} = trace[a_{44}] = a_{44}$$

For $\hat{h}_3$ we have

$$Z(v_1) = \lambda(Z)(v_1) = a_{11}v_1 \qquad \lambda(Z) = \tfrac{trace(Z)}{1} = trace[a_{11}] = a_{11}$$
$$Z(v_4) = \lambda(Z)(v_4) = 0 \qquad \lambda(Z) = \tfrac{trace(Z)}{1} = trace[0] = 0$$

The previous part was just an interlude in the proof of Lie's Theorem where we showed how to calculate the dual λ. To complete the proof of the theorem, we have to show how we can conclude to a simultaneous eigenvector for all X in $\hat{s}$. We now have $\hat{s}$ written in three ways:

$$\hat{s} = \hat{h}_1 \oplus sp(E_{11}) \qquad \hat{s} = \hat{h}_2 \oplus sp(E_{33}) \qquad \hat{s} = \hat{h}_3 \oplus sp(E_{44})$$

Thus

$$X = Z + a_{11}E_{11} \text{ with } Z \text{ in } \hat{h}_1 \qquad X = Z + a_{33}E_{33} \text{ with } Z \text{ in } \hat{h}_2$$
$$X = Z + a_{44}E_{44} \text{ with } Z \text{ in } \hat{h}_3$$

$$w = v_1 \qquad w = v_4$$
$$E_{44}(v_1) = 0 \qquad E_{44}(v_4) = v_4$$

which show that after one iteration we begin to repeat.

We know

$$\{v_1, v_4\} \text{ for } \hat{h}_1 \text{ with } Z(v_1) = \lambda_1(Z)(v_1) = 0;\ Z(v_4) = \lambda_1(Z)(v_4) = a_{44}v_4$$
$$\{v_1, v_4\} \text{ for } \hat{h}_2 \text{ with } Z(v_1) = \lambda_2(Z)(v_1) = a_{11}v_1;\ Z(v_4) = \lambda_2(Z)(v_4) = a_{44}v_4$$
$$\{v_1, v_4\} \text{ for } \hat{h}_3 \text{ with } Z(v_1) = \lambda_3(Z)(v_1) = a_{11}v_1;\ Z(v_4) = \lambda_3(Z)(v_4) = 0$$

and

$$E_{11}(v_1) = v_1; E_{11}(v_4) = 0$$
$$E_{33}(v_1) = 0; E_{33}(v_4) = 0$$
$$E_{44}(v_1) = 0; E_{44}(v_4) = v_4$$

We conclude

$$X(v_1) = Z(v_1) + a_{11}E_{11}(v_1) \text{ with } Z \text{ in } \hat{h}_1$$
$$X(v_4) = Z(v_4) + a_{11}E_{11}(v_4) \text{ with } Z \text{ in } \hat{h}_1$$

$$X(v_1) = Z(v_1) + a_{33}E_{33}(v_1) \text{ with } Z \text{ in } \hat{h}_2$$
$$X(v_4) = Z(v_4) + a_{33}E_{33}(v_4) \text{ with } Z \text{ in } \hat{h}_2$$

$$X(v_1) = Z(v_1) + a_{44}E_{44}(v_1) \text{ with } Z \text{ in } \hat{h}_3$$
$$X(v_4) = Z(v_4) + a_{44}E_{44}(v_4) \text{ with } Z \text{ in } \hat{h}_3$$

$$X(v_1) = \lambda_1(Z)(v_1) + a_{11}v_1 = 0 + a_{11}v_1 = a_{11}v_1 \text{ with } Z \text{ in } \hat{h}_1$$
$$\text{giving } \lambda_1(X) = a_{11} \text{ for } v_1$$
$$X(v_4) = \lambda_1(Z)(v_4) + a_{11}0 = a_{44}v_4 + 0 = a_{44}v_4 \text{ with } Z \text{ in } \hat{h}_1$$
$$\text{giving } \lambda_1(X) = a_{44} \text{ for } v_4$$

$$X(v_1) = \lambda_2(Z)(v_1) + a_{33}0 = a_{11}v_1 + 0 = a_{11}v_1 \text{ with } Z \text{ in } \hat{h}_2$$
$$\text{giving } \lambda_2(X) = a_{11} \text{ for } v_1$$
$$X(v_4) = \lambda_2(Z)(v_4) + a_{33}0 = a_{44}v_4 + 0 = a_{44}v_4 \text{ with } Z \text{ in } \hat{h}_2$$
$$\text{giving } \lambda_2(X) = a_{44} \text{ for } v_4$$

$$X(v_1) = \lambda_3(Z)(v_1) + a_{44}0 = a_{11}v_1 + 0 = a_{11}v_1 \text{ with } Z \text{ in } \hat{h}_3$$
$$\text{giving } \lambda_3(X) = a_{11} \text{ for } v_1$$
$$X(v_4) = \lambda_3(Z)(v_4) + a_{44}v_4 = 0 + a_{44}v_4 = a_{44}v_4 \text{ with } Z \text{ in } \hat{h}_3$$
$$\text{giving } \lambda_3(X) = a_{44} \text{ for } v_4$$

We observe that no matter how we choose X we obtain the same eigenvalues and indeed these eigenvalues agree with the matrix X

$$X = \begin{bmatrix} a_{11} & a_{12} & a_{13} & 0 \\ 0 & 0 & a_{23} & 0 \\ 0 & 0 & a_{33} & 0 \\ 0 & 0 & 0 & a_{44} \end{bmatrix}$$

We believe that this example gives a wonderful concrete understanding of Lie's Theorem.

But in order to get a better feel for the solvable case we repeat what we did above for the nilpotent case.

We assume a 4-dimensional abstract solvable Lie algebra $\hat{s}$ exists. Then given a basis in $\hat{s}$, the adjoint representation ad takes $\hat{s}$ into the 4x4 matrices $\widehat{gl}(\hat{s})$. We remark again that if we let $\hat{s} = V$, then V is 4-dimensional and thus we are in the above context of $\widehat{gl}(V)$. We thus know that we have a 10-dimensional solvable Lie subalgebra $\hat{g}$ of $\widehat{gl}(\hat{s})$, and on the basis (v_1, v_2, v_3, v_4), the elements of this subalgebra are the upper triangular matrices. Now the image of ad is a solvable Lie subalgebra of $\widehat{gl}(\hat{s})$. What we would like to show is that this subalgebra is a subalgebra of $\hat{g}$ in $\widehat{gl}(\hat{s})$. Thus we would like to identify the images by ad of the four vectors (v_1, v_2, v_3, v_4) as matrices in $\hat{g}$. Now we know that v_1 is a simultaneous eigenvector with eigenvalue $\lambda_1(X)$ for all X in $\hat{g}$. But $ad(v_1)$ is in $\hat{g}$. Thus we have

$$\begin{aligned}
(ad(v_1))(v_1) &= [v_1, v_1] = \lambda_1(ad(v_1))(v_1) = 0 \\
(ad(v_1))(v_2) &= [v_1, v_2] = -[v_2, v_1] = -(ad(v_2))(v_1) = -\lambda_1(ad(v_2))(v_1) \\
(ad(v_1))(v_3) &= [v_1, v_3] = -[v_3, v_1] = -(ad(v_3))(v_1) = -\lambda_1(ad(v_3))(v_1) \\
(ad(v_1))(v_4) &= [v_1, v_4] = -[v_4, v_1] = -(ad(v_4))(v_1) = -\lambda_1(ad(v_4))(v_1)
\end{aligned}$$

$$\begin{aligned}
(ad(v_2))(v_1) &= [v_2, v_1] = \lambda_1(ad(v_2))(v_1) \\
(ad(v_2))(v_2) &= [v_2, v_2] = 0 \\
(ad(v_2))(v_3) &= [v_2, v_3] = a_{13}v_1 + a_{23}v_2 + a_{33}v_3 \\
(ad(v_2))(v_4) &= [v_2, v_4] = a_{14}v_1 + a_{24}v_2 + a_{34}v_3 + a_{44}v_4
\end{aligned}$$

since we know that $ad(v_2)$ is an upper triangular matrix. Continuing,

$$\begin{aligned}
(ad(v_3))(v_1) &= [v_3, v_1] = \lambda_1(ad(v_3))(v_1) \\
(ad(v_3))(v_2) &= [v_3, v_2] = b_{12}v_1 + b_{22}v_2 \\
(ad(v_3))(v_3) &= [v_3, v_3] = 0 \\
(ad(v_3))(v_4) &= [v_3, v_4] = b_{14}v_1 + b_{24}v_2 + b_{34}v_3 + b_{44}v_4
\end{aligned}$$

$$\begin{aligned}
(ad(v_4))(v_1) &= [v_4, v_1] = \lambda_1(ad(v_4))(v_1) \\
(ad(v_4))(v_2) &= [v_4, v_2] = c_{12}v_1 + c_{22}v_2 \\
(ad(v_4))(v_3) &= [v_4, v_3] = c_{13}v_1 + c_{23}v_2 + c_{33}v_3 \\
(ad(v_4))(v_4) &= [v_4, v_4] = 0
\end{aligned}$$

Now using the relation $[u, v] = -[v, u]$, we see

$$\begin{gathered}
b_{12} = -a_{13} \qquad b_{22} = -a_{23} \qquad a_{33} = 0 \qquad c_{12} = -a_{14} \\
c_{22} = -a_{24} \qquad a_{34} = 0 \qquad a_{44} = 0 \qquad c_{13} = -b_{14} \\
c_{23} = -b_{24} \qquad c_{33} = -b_{34} \qquad b_{44} = 0
\end{gathered}$$

Thus we have

$$\begin{gathered}
(ad(v_2))(v_1) = [v_2, v_1] = \lambda_1(ad(v_2))(v_1) \\
(ad(v_2))(v_2) = [v_2, v_2] = 0 \\
(ad(v_2))(v_3) = [v_2, v_3] = a_{13}v_1 + a_{23}v_2 \\
(ad(v_2))(v_4) = [v_2, v_4] = a_{14}v_1 + a_{24}v_2
\end{gathered}$$

$$\begin{gathered}
(ad(v_3))(v_1) = [v_3, v_1] = \lambda_1(ad(v_3))(v_1) \\
(ad(v_3))(v_2) = [v_3, v_2] = -a_{13}v_1 - a_{23}v_2 \\
(ad(v_3))(v_3) = [v_3, v_3] = 0 \\
(ad(v_3))(v_4) = [v_3, v_4] = b_{14}v_1 + b_{24}v_2 + b_{34}v_3
\end{gathered}$$

$$(ad(v_4))(v_1) = [v_4, v_1] = \lambda_1(ad(v_4))(v_1)$$
$$(ad(v_4))(v_2) = [v_4, v_2] = -a_{14}v_1 - a_{24}v_2$$
$$(ad(v_4))(v_3) = [v_4, v_3] = -b_{14}v_1 - b_{24}v_2 - b_{34}v_3$$
$$(ad(v_4))(v_4) = [v_4, v_4] = 0$$

Thus the matrices are:

$$ad(v_1) = \begin{bmatrix} 0 & -\lambda_1(ad(v_2)) & -\lambda_1(ad(v_3)) & -\lambda_1(ad(v_4)) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$
$$ad(v_2) = \begin{bmatrix} \lambda_1(ad(v_2)) & 0 & a_{13} & a_{14} \\ 0 & 0 & a_{23} & a_{24} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$
$$ad(v_3) = \begin{bmatrix} \lambda_1(ad(v_3)) & -a_{13} & 0 & b_{14} \\ 0 & -a_{23} & 0 & b_{24} \\ 0 & 0 & 0 & b_{34} \\ 0 & 0 & 0 & 0 \end{bmatrix},$$
$$ad(v_4) = \begin{bmatrix} \lambda_1(ad(v_4)) & -a_{14} & -b_{14} & 0 \\ 0 & -a_{24} & -b_{24} & 0 \\ 0 & 0 & -b_{34} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

We now want to check the homomorphism between $\hat{s}$ and this Lie subalgebra $ad(\hat{s})$ of $\hat{g}$ by doing the following computations:

$$ad[v_i, v_j] = [ad(v_i), ad(v_j)]$$

$$[v_1, v_2] = -\lambda_1(ad(v_2))(v_1) \longrightarrow ad([v_1, v_2]) = ad(-\lambda_1(ad(v_2))(v_1)) =$$
$$-\lambda_1(ad(v_2))ad(v_1) =$$
$$-\lambda_1(ad(v_2)) \begin{bmatrix} 0 & -\lambda_1(ad(v_2)) & -\lambda_1(ad(v_3)) & -\lambda_1(ad(v_4)) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} =$$
$$\begin{bmatrix} 0 & (\lambda_1(ad(v_2)))^2 & (\lambda_1(ad(v_2)))(\lambda_1(ad(v_3))) & (\lambda_1(ad(v_2)))(\lambda_1(ad(v_4))) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix};$$

$$[ad(v_1), ad(v_2)] =$$
$$\begin{bmatrix} 0 & -\lambda_1(ad(v_2)) & -\lambda_1(ad(v_3)) & -\lambda_1(ad(v_4)) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1(ad(v_2)) & 0 & a_{13} & a_{14} \\ 0 & 0 & a_{23} & a_{24} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} -$$

$$\begin{bmatrix} \lambda_1(ad(v_2)) & 0 & a_{13} & a_{14} \\ 0 & 0 & a_{23} & a_{24} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & -\lambda_1(ad(v_2)) & -\lambda_1(ad(v_3)) & -\lambda_1(ad(v_4)) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & 0 & -(\lambda_1(ad(v_2)))a_{23} & -(\lambda_1(ad(v_2)))a_{24} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} -$$

$$\begin{bmatrix} 0 & -(\lambda_1(ad(v_2)))^2 & -\lambda_1(ad(v_2))\lambda_1(ad(v_3)) & -\lambda_1(ad(v_2))\lambda_1(ad(v_4)) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} =$$

[Because of the length of some entries in these matrices, it is now necessary to write this matrix in two parts. The $*_{ij}$ indicates where these entries occur.]

$$\begin{bmatrix} 0 & (\lambda_1(ad(v_2)))^2 & -(\lambda_1(ad(v_2)))a_{23} + (\lambda_1(ad(v_2)))(\lambda_1(ad(v_3))) & *_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & *_{12} & *_{13} & -(\lambda_1(ad(v_2)))a_{24} + (\lambda_1(ad(v_2)))(\lambda_1(ad(v_4))) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

We observe that we can effect a homomorphism if $a_{23} = 0$ and $a_{24} = 0$. This gives the matrix

$$\begin{bmatrix} 0 & (\lambda_1(ad(v_2)))^2 & (\lambda_1(ad(v_2)))(\lambda_1(ad(v_3))) & (\lambda_1(ad(v_2)))(\lambda_1(ad(v_4))) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Continuing,

$$[v_1, v_3] = -\lambda_1(ad(v_3))(v_1) \longrightarrow ad([v_1, v_3]) = ad(-\lambda_1(ad(v_3))(v_1)) =$$
$$-\lambda_1(ad(v_3))ad(v_1) =$$
$$-\lambda_1(ad(v_3)) \begin{bmatrix} 0 & -\lambda_1(ad(v_2)) & -\lambda_1(ad(v_3)) & -\lambda_1(ad(v_4)) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & (\lambda_1(ad(v_3)))(\lambda_1(ad(v_2))) & \lambda_1(ad(v_3))^2 & (\lambda_1(ad(v_3)))(\lambda_1(ad(v_4))) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$[ad(v_1), ad(v_3)] =$$

$$\begin{bmatrix} 0 & -\lambda_1(ad(v_2)) & -\lambda_1(ad(v_3)) & -\lambda_1(ad(v_4)) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1(ad(v_3)) & -a_{13} & 0 & b_{14} \\ 0 & -a_{23} & 0 & b_{24} \\ 0 & 0 & 0 & b_{34} \\ 0 & 0 & 0 & 0 \end{bmatrix} -$$

$$\begin{bmatrix} \lambda_1(ad(v_3)) & -a_{13} & 0 & b_{14} \\ 0 & -a_{23} & 0 & b_{24} \\ 0 & 0 & 0 & b_{34} \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & -\lambda_1(ad(v_2)) & -\lambda_1(ad(v_3)) & -\lambda_1(ad(v_4)) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & (\lambda_1(ad(v_2)))a_{23} & 0 & -(\lambda_1(ad(v_2)))b_{24} - (\lambda_1(ad(v_3)))b_{34} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} -$$

$$\begin{bmatrix} 0 & -\lambda_1(ad(v_3))\lambda_1(ad(v_2)) & -(\lambda_1(ad(v_3)))^2 & -\lambda_1(ad(v_3))\lambda_1(ad(v_4)) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} =$$

since we know that $a_{23} = 0$,

[Again, because of the length of some entries in these matrices, it is now necessary to write this matrix in two parts. And again the $*_{ij}$ indicates where these entries occur.]

$$\begin{bmatrix} 0 & (\lambda_1(ad(v_3)))(\lambda_1(ad(v_2))) & \lambda_1(ad(v_3))^2 & *_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & *_{12} & *_{13} & -(\lambda_1(ad(v_2)))b_{24} - (\lambda_1(ad(v_3)))b_{34} + (\lambda_1(ad(v_3)))(\lambda_1(ad(v_4))) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

We observe that we can effect a homomorphism if $b_{24} = 0$ and $b_{34} = 0$:

$$\begin{bmatrix} 0 & (\lambda_1(ad(v_2)))(\lambda_1(ad(v_3))) & \lambda_1(ad(v_3))^2 & (\lambda_1(ad(v_3)))(\lambda_1(ad(v_4))) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Continuing,

$$[v_1, v_4] = -\lambda_1(ad(v_4))(v_1) \longrightarrow ad([v_1, v_4]) = ad(-\lambda_1(ad(v_4))(v_1)) =$$
$$-\lambda_1(ad(v_4))ad(v_1) =$$
$$-\lambda_1(ad(v_4)) \begin{bmatrix} 0 & -\lambda_1(ad(v_2)) & -\lambda_1(ad(v_3)) & -\lambda_1(ad(v_4)) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} =$$
$$\begin{bmatrix} 0 & (\lambda_1(ad(v_4)))(\lambda_1(ad(v_2))) & (\lambda_1(ad(v_4)))(\lambda_1(ad(v_3))) & \lambda_1(ad(v_4))^2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$[ad(v_1), ad(v_4)] =$$
$$\begin{bmatrix} 0 & -\lambda_1(ad(v_2)) & -\lambda_1(ad(v_3)) & -\lambda_1(ad(v_4)) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1(ad(v_4)) & -a_{14} & -b_{14} & 0 \\ 0 & -a_{24} & -b_{24} & 0 \\ 0 & 0 & -b_{34} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} -$$
$$\begin{bmatrix} \lambda_1(ad(v_4)) & -a_{14} & -b_{14} & 0 \\ 0 & -a_{24} & -b_{24} & 0 \\ 0 & 0 & -b_{34} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & -\lambda_1(ad(v_2)) & -\lambda_1(ad(v_3)) & -\lambda_1(ad(v_4)) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & (\lambda_1(ad(v_2)))a_{24} & (\lambda_1(ad(v_2)))b_{24} + (\lambda_1(ad(v_3)))b_{34} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} -$$
$$\begin{bmatrix} 0 & -\lambda_1(ad(v_4))\lambda_1(ad(v_2)) & -\lambda_1(ad(v_4))\lambda_1(ad(v_3)) & -(\lambda_1(ad(v_4)))^2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} =$$

since we know that $a_{24} = 0; b_{24} = 0; b_{34} = 0,$

$$\begin{bmatrix} 0 & \lambda_1(ad(v_4))\lambda_1(ad(v_2)) & \lambda_1(ad(v_4))\lambda_1(ad(v_3)) & (\lambda_1(ad(v_4)))^2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

We observe that we effect the homomorphism immediately in this calculation.

Continuing, since we know that $a_{23} = 0$, we have

$$[v_2, v_3] = a_{13}v_1 + a_{23}v_2 = a_{13}v_1 \longrightarrow ad([v_2, v_3]) = ad(a_{13}v_1) = a_{13}ad(v_1) =$$

$$a_{13}\begin{bmatrix} 0 & -\lambda_1(ad(v_2)) & -\lambda_1(ad(v_3)) & -\lambda_1(ad(v_4)) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$[ad(v_2), ad(v_3)] =$$

since we know that $a_{23} = 0$; $a_{24} = 0$; $b_{24} = 0$; $b_{34} = 0$, and

$$\left(\begin{bmatrix} \lambda_1(ad(v_2)) & 0 & a_{13} & a_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1(ad(v_3)) & -a_{13} & 0 & b_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right) -$$
$$\left(\begin{bmatrix} \lambda_1(ad(v_3)) & -a_{13} & 0 & b_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1(ad(v_2)) & 0 & a_{13} & a_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right) =$$

$$\begin{bmatrix} (\lambda_1(ad(v_2))(\lambda_1(ad(v_3))) & (\lambda_1(ad(v_2)))(-a_{13}) & 0 & (\lambda_1(ad(v_2)))b_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} -$$
$$\begin{bmatrix} \lambda_1(ad(v_3))(\lambda_1(ad(v_2))) & 0 & \lambda_1(ad(v_3))(a_{13}) & \lambda_1(ad(v_3))(a_{14}) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & -(\lambda_1(ad(v_2)))a_{13} & -(\lambda_1(ad(v_3)))a_{13} & \lambda_1(ad(v_2))b_{14} - \lambda_1(ad(v_3))a_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

We observe that we can effect a homomorphism if

$$-a_{13}(\lambda_1(ad(v_4))) = \lambda_1(ad(v_2))b_{14} - \lambda_1(ad(v_3))a_{14}$$

Continuing, since we know that $a_{24} = 0$, we have

$$[v_2, v_4] = a_{14}v_1 \longrightarrow ad([v_2, v_4]) = ad(a_{14}(v_1)) = a_{14}ad(v_1) =$$
$$a_{14}\begin{bmatrix} 0 & -\lambda_1(ad(v_2)) & -\lambda_1(ad(v_3)) & -\lambda_1(ad(v_4)) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$[ad(v_2), ad(v_4)] =$$

since we know that $a_{23} = 0$; $a_{24} = 0$; $b_{24} = 0$; $b_{34} = 0$,

$$\left(\begin{bmatrix} \lambda_1(ad(v_2)) & 0 & a_{13} & a_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1(ad(v_4)) & -a_{14} & -b_{14} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right) -$$
$$\left(\begin{bmatrix} \lambda_1(ad(v_4)) & -a_{14} & -b_{14} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1(ad(v_2)) & 0 & a_{13} & a_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right) =$$

$$\begin{bmatrix} (\lambda_1(ad(v_2))(\lambda_1(ad(v_4))) & (\lambda_1(ad(v_2)))(-a_{14}) & (\lambda_1(ad(v_2)))(-b_{14}) & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} -$$
$$\begin{bmatrix} (\lambda_1(ad(v_4))(\lambda_1(ad(v_2))) & 0 & \lambda_1(ad(v_4))(a_{13}) & \lambda_1(ad(v_4))(a_{14}) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & -(\lambda_1(ad(v_2)))a_{14} & -(\lambda_1(ad(v_2)))b_{14} - \lambda_1(ad(v_4))a_{13} & -\lambda_1(ad(v_4))a_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

We observe that we can effect a homomorphism if

$$-a_{14}\lambda_1(ad(v_3)) = -\lambda_1(ad(v_2))b_{14} - \lambda_1(ad(v_4))a_{13}.$$

Continuing, since we know that $b_{24} = 0$ and $b_{34} = 0$, we have

$$[v_3, v_4] = b_{14}v_1 \longrightarrow ad([v_3, v_4]) = ad(b_{14}(v_1)) = b_{14}ad(v_1) =$$
$$b_{14}\begin{bmatrix} 0 & -\lambda_1(ad(v_2)) & -\lambda_1(ad(v_3)) & -\lambda_1(ad(v_4)) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$[ad(v_3), ad(v_4)] =$$

since we know that $a_{23} = 0$; $a_{24} = 0$; $b_{24} = 0$; $b_{34} = 0$,

$$\left(\begin{bmatrix} \lambda_1(ad(v_3)) & -a_{13} & 0 & b_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1(ad(v_4)) & -a_{14} & -b_{14} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right) -$$

$$\left(\begin{bmatrix} \lambda_1(ad(v_4)) & -a_{14} & -b_{14} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1(ad(v_3)) & -a_{13} & 0 & b_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right) =$$

$$\begin{bmatrix} (\lambda_1(ad(v_3))(\lambda_1(ad(v_4))) & (\lambda_1(ad(v_3)))(-a_{14}) & (\lambda_1(ad(v_3)))(-b_{14}) & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} -$$

$$\begin{bmatrix} \lambda_1(ad(v_4))(\lambda_1(ad(v_3))) & \lambda_1(ad(v_4))(-a_{13}) & 0 & \lambda_1(ad(v_4))(b_{14}) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & -\lambda_1(ad(v_3))a_{14} + \lambda_1(ad(v_4))a_{13} & -(\lambda_1(ad(v_3)))b_{14} & -(\lambda_1(ad(v_4)))b_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

We observe that we can effect a homomorphism if

$$-b_{14}\lambda_1(ad(v_2)) = -\lambda_1(ad(v_3))a_{14} + \lambda_1(ad(v_4))a_{13}.$$

Besides $a_{23} = 0; a_{24} = 0; b_{24} = 0; b_{34} = 0$, we have found the following three relations should hold:

$$\begin{aligned} -a_{13}\lambda_1(ad(v_4)) &= \lambda_1(ad(v_2))b_{14} - \lambda_1(ad(v_3))a_{14} \\ -a_{14}\lambda_1(ad(v_3)) &= -\lambda_1(ad(v_2))b_{14} - \lambda_1(ad(v_4))a_{13} \\ -b_{14}\lambda_1(ad(v_2)) &= -\lambda_1(ad(v_3))a_{14} + \lambda_1(ad(v_4))a_{13}. \end{aligned}$$

But on examining these three relations, we see that they all give the same equality:

$$b_{14}\lambda_1(ad(v_2)) - a_{14}\lambda_1(ad(v_3)) + a_{13}\lambda_1(ad(v_4)) = 0.$$

It is evident now that we can rescale the vectors v_2, v_3, v_4 so that $a_{13} = a_{14} = b_{14} = 1$, giving the relation between the eigenvalues for the simultaneous eigenvector v_1:

$$\lambda_1(ad(v_2)) - \lambda_1(ad(v_3)) + \lambda_1(ad(v_4)) = 0.$$

Substituting then this one relation, we obtain

$$[ad(v_2), ad(v_3)] = \begin{bmatrix} 0 & -\lambda_1(ad(v_2)) & -\lambda_1(ad(v_3)) & -\lambda_1(ad(v_4)) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$[ad(v_2), ad(v_4)] = \begin{bmatrix} 0 & -\lambda_1(ad(v_2)) & -\lambda_1(ad(v_3)) & -\lambda_1(ad(v_4)) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$[ad(v_3), ad(v_4)] = \begin{bmatrix} 0 & -\lambda_1(ad(v_2)) & -\lambda_1(ad(v_3)) & -\lambda_1(ad(v_4)) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus brackets go into brackets by ad and the homomorphism is verified.

The matrices which are the images of ad under this homomorphism are:

$$ad(v_1) = \begin{bmatrix} 0 & -\lambda_1(ad(v_2)) & -\lambda_1(ad(v_3)) & -\lambda_1(ad(v_4)) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$ad(v_2) = \begin{bmatrix} \lambda_1(ad(v_2)) & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \; ad(v_3) = \begin{bmatrix} \lambda_1(ad(v_3)) & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$ad(v_4) = \begin{bmatrix} \lambda_1(ad(v_4)) & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

In addition the eigenvalues and the arbitrary constants are tied together by the relation:

$$\lambda_1(ad(v_2)) - \lambda_1(ad(v_3)) + \lambda_1(ad(v_4)) = 0.$$

We remark that the image of $\hat{s}$ under ad lives in $\hat{g}$, the Lie algebra of upper diagonal matrices, which is a Lie subalgebra of $\widehat{gl}(\hat{s})$, all written with respect to the rescaled vectors v_1, v_2, v_3, v_4. Indeed the image is the subspace generated by the matrices $E_{11}, E_{12}, E_{13}, E_{14}$. Now these four matrices produce six independent brackets, of which the only non-zero ones are:

$$[E_{11}, E_{12}] = E_{12} \qquad [E_{11}, E_{13}] = E_{13} \qquad [E_{11}, E_{14}] = E_{14}$$

We observe that these brackets close in the subspace, and thus the subspace is a Lie subalgebra of $ad(\hat{s})$. Also from the brackets that we have found, we see that $[ad(\hat{s}), ad(\hat{s})] \neq 0$, and thus $D^1 ad(\hat{s}) \neq 0$. But we see from the above brackets of the basis vectors $E_{11}, E_{12}, E_{13}, E_{14}$ that $D^2 ad(\hat{s}) = 0$, and thus $ad(\hat{s})$ is a solvable Lie subalgebra of $\hat{g}$.

The question to ask, however, is the following. Is ad an isomorphism or a homomorphism with a nontrivial kernel? This means we are asking the question: are the images under ad of the four basis vectors of $\hat{s}$, v_1, v_2, v_3, v_4, linearly independent? Doing the linear algebra, we see that the matrix which determines this calculation in the four-dimensional subspace $sp(E_{11}, E_{12}, E_{13}, E_{14})$ is

$$\begin{bmatrix} 0 & \lambda_1(ad(v_2)) & \lambda_1(ad(v_3)) & \lambda_1(ad(v_4)) \\ -\lambda_1(ad(v_2)) & 0 & -1 & -1 \\ -\lambda_1(ad(v_3)) & 1 & 0 & -1 \\ -\lambda_1(ad(v_4)) & 1 & 1 & 0 \end{bmatrix}$$

If we calculate the determinant of this matrix, we obtain

$$(\lambda_1(ad(v_2)))^2 - 2\lambda_1(ad(v_2))\lambda_1(ad(v_3)) + (\lambda_1(ad(v_3)))^2 +$$
$$2\lambda_1(ad(v_2))\lambda_1(ad(v_4)) - 2\lambda_1(ad(v_3))\lambda_1(ad(v_4)) + (\lambda_1(ad(v_4)))^2$$

However the above determinant can be written as

$$(\lambda_1(ad(v_2)) - \lambda_1(ad(v_3)) + \lambda_1(ad(v_4)))^2$$

We recognize that this is the square of the expression

$$\lambda_1(ad(v_2)) - \lambda_1(ad(v_3)) + \lambda_1(ad(v_4))$$

and thus $= 0$. This guarantees that we have a homomorphism but not an isomorphism between $\hat{s}$ and $ad(\hat{s})$. Thus we know that the above matrix is singular and that ad has a kernel. On the assumption that $\lambda_1(ad(v_2)) \neq 0$, we row-reduce this matrix and obtain

$$\begin{bmatrix} 1 & 0 & \frac{1}{\lambda_1(ad(v_2))} & \frac{1}{\lambda_1(ad(v_2))} \\ 0 & 1 & \frac{\lambda_1(ad(v_3))}{\lambda_1(ad(v_2))} & \frac{\lambda_1(ad(v_4))}{\lambda_1(ad(v_2))} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus we know that the vectors

$$u_3 = -(\frac{1}{\lambda_1(ad(v_2))})v_1 - (\frac{\lambda_1(ad(v_3))}{\lambda_1(ad(v_2))})v_2 + v_3$$

and

$$u_4 = -(\frac{1}{\lambda_1(ad(v_2))})v_1 - (\frac{\lambda_1(ad(v_4))}{\lambda_1(ad(v_2))})v_2 + v_4$$

are in the kernel of ad, and thus are a basis for the center of $\hat{s}$. We confirm these statements below:

$$ad(u_3) = ad(-(\frac{1}{\lambda_1(ad(v_2))})v_1 - (\frac{\lambda_1(ad(v_3))}{\lambda_1(ad(v_2))})v_2 + v_3) =$$

$$-(\frac{1}{\lambda_1(ad(v_2))})ad(v_1) - (\frac{\lambda_1(ad(v_3))}{\lambda_1(ad(v_2))})ad(v_2) + ad(v_3) =$$

$$-(\frac{1}{\lambda_1(ad(v_2))})\begin{bmatrix} 0 & -\lambda_1(ad(v_2)) & -\lambda_1(ad(v_3)) & -\lambda_1(ad(v_4)) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} -$$

$$(\frac{\lambda_1(ad(v_3))}{\lambda_1(ad(v_2))})\begin{bmatrix} \lambda_1(ad(v_2)) & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} \lambda_1(ad(v_3)) & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & 1 & \frac{\lambda_1(ad(v_3))}{\lambda_1(ad(v_2))} & \frac{\lambda_1(ad(v_4))}{\lambda_1(ad(v_2))} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} \lambda_1(ad(v_3)) & 0 & \frac{\lambda_1(ad(v_3))}{\lambda_1(ad(v_2))}) & \frac{\lambda_1(ad(v_3))}{\lambda_1(ad(v_2))}) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} +$$

$$\begin{bmatrix} \lambda_1(ad(v_3)) & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & 0 & 0 & \frac{\lambda_1(ad(v_4))}{\lambda_1(ad(v_2))} - \frac{\lambda_1(ad(v_3))}{\lambda_1(ad(v_2))}) + 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

if we apply our condition

$$\lambda_1(ad(v_2)) - \lambda_1(ad(v_3)) + \lambda_1(ad(v_4)) = 0.$$

Continuing

$$ad(u_4) = ad(-(\tfrac{1}{\lambda_1(ad(v_2))})v_1 - (\tfrac{\lambda_1(ad(v_4))}{\lambda_1(ad(v_2))})v_2 + v_4) =$$

$$-(\tfrac{1}{\lambda_1(ad(v_2))})ad(v_1) - (\tfrac{\lambda_1(ad(v_4))}{\lambda_1(ad(v_2))})ad(v_2) + ad(v_4) =$$

$$-(\tfrac{1}{\lambda_1(ad(v_2))})\begin{bmatrix} 0 & -\lambda_1(ad(v_2)) & -\lambda_1(ad(v_3)) & -\lambda_1(ad(v_4)) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} -$$

$$(\tfrac{\lambda_1(ad(v_4))}{\lambda_1(ad(v_2))})\begin{bmatrix} \lambda_1(ad(v_2)) & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} \lambda_1(ad(v_4)) & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & 1 & \frac{\lambda_1(ad(v_3))}{\lambda_1(ad(v_2))} & \frac{\lambda_1(ad(v_4))}{\lambda_1(ad(v_2))} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} \lambda_1(ad(v_4)) & 0 & \frac{\lambda_1(ad(v_4))}{\lambda_1(ad(v_2))}) & \frac{\lambda_1(ad(v_4))}{\lambda_1(ad(v_2))}) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} +$$

$$\begin{bmatrix} \lambda_1(ad(v_4)) & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & 0 & \frac{\lambda_1(ad(v_3))}{\lambda_1(ad(v_2))} - \frac{\lambda_1(ad(v_4))}{\lambda_1(ad(v_2))}) - 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

if we again apply our condition

$$\lambda_1(ad(v_2)) - \lambda_1(ad(v_3)) + \lambda_1(ad(v_4)) = 0.$$

On the assumption that $\lambda_1(ad(v_2)) \neq 0$, we now know that $\hat{s}$ is a solvable Lie algebra with a two-dimensional center, and that $ad(\hat{s})$ is a two dimensional solvable Lie subalgebra of the solvable Lie subalgebra $\hat{g}$ of the Lie algebra $\widehat{gl}(\hat{s})$.

We see immediately that the center of $\hat{s}$ is generated by the two basis vectors u_3 and u_4, while the image of $\hat{s}$ in $\hat{g}$ by ad is generated by the basis vectors E_{11} and E_{12}. Since $[E_{11}, E_{12}] = E_{12}$, we see that $D^1 ad(\hat{s}) \neq 0$ and that $D^2 ad(\hat{s}) = 0$, confirming our above results. We remark that Lie's Theorem says that $ad(\hat{s})$ now has a simultaneous eigenvector, which is either $u_1 = v_1$ or $u_2 = v_2$, with corresponding eigenvalue $\lambda_1(ad(u_i))$, $i = 1$ or 2, giving 0 or $\lambda_1(ad(v_2) \neq 0$.

If $\lambda_1(ad(v_2)) = 0$, then we see that in order to satisfy the condition

$$\lambda_1(ad(v_2)) - \lambda_1(ad(v_3)) + \lambda_1(ad(v_4)) = 0.$$

either $\lambda_1(ad(v_3)) = \lambda_1(ad(v_4))$ if $\lambda_1(ad(v_3)) \neq 0$, or $\lambda_1(ad(v_3)) = \lambda_1(ad(v_4)) = 0$. [If $\lambda_1(ad(v_2)) = \lambda_1(ad(v_3)) = \lambda_1(ad(v_4)) = 0$, then $ad(\hat{s})$ is a nilpotent Lie algebra, and $\hat{s}$ becomes a nilpotent Lie algebra, and we return to the example of 2.8.4 with $b_{24} = 0$.]

Thus we are now in the case when $\lambda_1(ad(v_2)) = 0$ and $\lambda_1(ad(v_3)) = \lambda_1(ad(v_4))$. And again we ask whether ad is an isomorphism or a homomorphism with a nontrivial kernel? This means again that we are asking the question: are the images under ad of the four basis vectors of $\hat{s}$, v_1, v_2, v_3, v_4, linearly independent? Doing the linear algebra, we see that the matrix which determines this calculation in the four-dimensional subspace $sp(E_{11}, E_{12}, E_{13}, E_{14})$ is

$$\begin{bmatrix} 0 & 0 & \lambda_1(ad(v_3)) & \lambda_1(ad(v_3)) \\ 0 & 0 & -1 & -1 \\ -\lambda_1(ad(v_3)) & 1 & 0 & -1 \\ -\lambda_1(ad(v_3)) & 1 & 1 & 0 \end{bmatrix}$$

It is straightforward that the determinant of this matrix is equal to 0. Row reducing this matrix we obtain

$$\begin{bmatrix} 1 & \frac{-1}{\lambda_1(ad(v_3))} & 0 & \frac{1}{\lambda_1(ad(v_3))} \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus we know that the vectors

$$u_2 = \tfrac{1}{\lambda_1(ad(v_3))}(v_1) + v_2$$

and

$$u_4 = -\tfrac{1}{\lambda_1(ad(v_3))}(v_1) - v_3 + v_4$$

are in the kernel of ad, and thus are a basis for the center of $\hat{s}$. We confirm these statements in what follows:

$$ad(u_2) = ad(\tfrac{1}{\lambda_1(ad(v_3))}v_1 + v_2) = \tfrac{1}{\lambda_1(ad(v_3))}ad(v_1) + ad(v_2) = 0 + ad(v_2) =$$
$$\begin{bmatrix} 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Continuing

$$ad(u_4) = ad(-\tfrac{1}{\lambda_1(ad(v_3)}(v_1) - v_3 + v_4) = -\tfrac{1}{\lambda_1(ad(v_3)}(ad(v_1) - ad(v_3) + ad(v_4) =$$

$$-\tfrac{1}{\lambda_1(ad(v_3)}\begin{bmatrix} 0 & 0 & -\lambda_1(ad(v_3) & -\lambda_1(ad(v_3) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} \lambda_1(ad(v_3) & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} +$$

$$\begin{bmatrix} \lambda_1(ad(v_3) & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

In this case we have shown that the vectors

$$u_2 = \tfrac{1}{\lambda_1(ad(v_3))}v_1 + v_2 \qquad \text{and} \qquad u_4 = -\tfrac{1}{\lambda_1(ad(v_3))}v_1 - v_3 + v_4$$

are in the center of $\hat{s}$. We see immediately that the center of $\hat{s}$ is generated by the two basis vectors u_2 and u_4, while the image of $\hat{s}$ in $\hat{g}$ by ad is generated by the basis vectors E_{11} and E_{13}. Since $[E_{11}, E_{13}] = E_{13}$, we see that $D^1 ad(\hat{s}) \neq 0$ and that $D^2 ad(\hat{s}) = 0$, confirming our above results. We remark that Lie's Theorem says that $ad(\hat{s})$ now has a simultaneous eigenvector, which is either $u_1 = v_1$ or $u_3 = v_3$, with corresponding eigenvalue $\lambda_1(ad(u_i))$, $i = 1$ or 3, giving 0 or $\lambda_1(ad(v_3) \neq 0$.

The final case has $\lambda_1(ad(v_2)) = \lambda_1(ad(v_3)) = \lambda_1(ad(v_4)) = 0$. In this case the vector v_1 belongs to the center of $\hat{s}$ since it is a simultaneous eigenvector with eigenvalue 0. We also have in the center

$$u_4 = v_2 - v_3 + v_4$$

We confirm this latter statement below:

$$ad(u_4) = ad(v_2) - ad(v_3) + ad(v_4) =$$

$$\begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Continuing

$$ad(u_4) = ad(-\tfrac{1}{\lambda_1(ad(v_3)}(v_1) - v_3 + v_4) = -\tfrac{1}{\lambda_1(ad(v_3)}(ad(v_1) - ad(v_3) + ad(v_4) =$$

$$-\frac{1}{\lambda_1(ad(v_3)}\begin{bmatrix} 0 & 0 & -\lambda_1(ad(v_3) & -\lambda_1(ad(v_3) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} \lambda_1(ad(v_3) & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} +$$

$$\begin{bmatrix} \lambda_1(ad(v_3) & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

We see immediately that the center of $\hat{s}$ is generated by the two basis vectors v_1 and u_4, while the image of $\hat{s}$ in $\hat{g}$ by ad is generated by the basis vectors E_{12} and E_{13}. Since $[E_{12}, E_{13}] = 0$, we see that $D^1 ad(\hat{s}) = 0$. and thus $ad(\hat{s})$ is abelian. We remark that Lie's Theorem says that $ad(\hat{s})$ now has two simultaneous eigenvectors, v_1 and u_4 with corresponding eigenvalues 0 and 0. We also observe that $ad(\hat{s})$ is nilpotent.

2.10 Some Remarks on Semisimple Lie Algebras (2)

Recall that we took this excursion into nilpotent and solvable Lie algebras because we wanted to identify the automorphism A between the semisimple parts of two Levi decompositions of a Lie algebra $\hat{g}$, where a Levi decomposition is a splitting at the level of linear spaces of an arbitrary Lie algebra $\hat{g}$ into its radical $\hat{r}$ and a semisimple Lie algebra $\hat{k}$.

$$\hat{g} = \hat{k} \oplus \hat{r}$$

Thus we now want to examine the other piece of this direct sum decomposition, that is, we want to examine more in detail the structure of a semisimple Lie algebra. We now change notation and call the semisimple Lie algebra $\hat{g}$.

2.10.1 The Homomorphic Image of Semisimple Lie Algebra is Semisimple. First we assert the homomorphic image of a semisimple Lie algebra is also semisimple. Above we proved that the quotient of a Lie algebra by its radical is a semisimple Lie algebra. If we let the homomorphism be represented by $\phi : \hat{g} \longrightarrow \phi(\hat{g}) = \hat{h}$, we see that the proof essentially said that if the Lie algebra $\hat{h}$ had a solvable ideal $\hat{s}$, its pre-image $\phi^{-1}(\hat{s})$ would also be solvable. But since $\hat{g}$ is semisimple, then this solvable ideal must be the zero ideal, and thus its image must also be the zero ideal. We conclude that $\hat{h}$ is semisimple.

2.10.2 $\hat{\mathbf{g}}$ Semisimple Implies $\mathbf{D^1\hat{g} = \hat{g}}$. Next we want to show that if $\hat{g}$ is semisimple, then $D^1\hat{g} = [\hat{g}, \hat{g}] = \hat{g}$. Since $D^1\hat{g}$ is an ideal, we again take the quotient algebra $\hat{g}/D^1\hat{g}$. [We know that $D^1\hat{g} \neq 0$, for if it were 0, then $\hat{g}$ would be solvable.] Now suppose that $D^1\hat{g} \neq \hat{g}$. We have already remarked that this quotient is the abelianization of the Lie algebra $\hat{g}$. Thus this quotient is a nonzero abelian Lie algebra. However we know that the quotient of a semisimple Lie algebra $\hat{g}$ is a semisimple Lie algebra. But an nonzero abelian Lie algebra is solvable and thus its radical is not zero. We conclude that the quotient is the zero Lie algebra, and thus $D^1\hat{g} = \hat{g}$.

2.10.3 Simple Lie Algebras. We now give the another class of Lie algebras which form the building blocks of the other Lie algebras — the simple Lie algebras. We define a *simple Lie algebra* $\hat{a}$ to be a Lie algebra which is not abelian, and whose only ideals are the improper ideals 0 and $\hat{a}$. We can show that every simple Lie algebra is also semisimple. For let $\hat{r}$ be the radical of $\hat{a}$ and $\neq 0$. Since $\hat{r}$ is an ideal, it can only be 0 or $\hat{a}$. If the radical is $\hat{a}$, then we know that $\hat{a}$ has a nonzero abelian ideal. But this says that $\hat{a}$ is this nonzero abelian ideal. But by definition no simple Lie algebra can be abelian. Thus we conclude that the radical of $\hat{a} = 0$, and thus $\hat{a}$ is semisimple.

The final fact that we want to establish is that every semisimple Lie algebra can be decomposed into a direct sum of ideals, each of which is a simple Lie algebra. And thus we see that the building blocks of the Lie algebras are the simple Lie algebras, and with the Levi decomposition theorem, the radical of the Lie algebra.

2.11 The Killing Form (1)

2.11.1 Structure of the Killing Form and its Fundamental Theorem. In order to obtain the above conclusion about semisimple Lie algebras, we introduce a powerful tool in the study of Lie algebras — the *Killing form*. First we define the Killing form $\hat{B}$, and then we specialize it to the adjoint representation.

The term "form" here carries with it the usual meaning: given a linear space over a scalar field $\mathbb{F}$ of characteristic 0 – see below – [recall that because we are considering only the fields of the real numbers or the complex numbers, we meet the condition that the scalar field have characteristic 0], a form $\hat{B}$ is a bilinear function on the linear space with values in the field of scalars. But the linear space involved in this definition is a very special one linked to the concept of a Lie algebra. Thus we begin with a linear space V and take its Lie algebra of the endomorphisms of V, $\widehat{gl}(V)$, and then we take a Lie subalgebra $\hat{g}$ of this Lie algebra. Given a basis for V, we then define $\hat{B}$ on this linear space $\hat{g}$ to be

$$\hat{B} : \hat{g} \times \hat{g} \longrightarrow \mathbb{F}$$

$$(X,Y) \longmapsto \hat{B}(X,Y) := trace(X \circ Y)$$

where $X \circ Y$ is just the product of two matrices. [We might remark that in a matrix representation of $End(V)$, since the trace is just the sum of the elements on the diagonal of the matrix, and since we do not want a sum to be zero because of a finite characteristic of the field of scalars, we have restricted the field of scalars to have characteristic 0.] Since, as we shall see, the trace of a linear transformation is independent of the choice of a basis, we have a valid definition. What is amazing is that we are reducing the information of a product of two matrices X and Y to information about its product's diagonal only, and then reducing once more this information to just the sum of the elements on this product's diagonal. Yet as in other uses of this concept in mathematics this information is so sufficiently rich that it gives information about the structure of the original matrices and ultimately also about the structure of the Lie algebra $\hat{g}$.

Indeed this definition gives us a bilinear form.

$$\hat{B}(X_1 + X_2, Y) = trace((X_1 + X_2) \circ Y) = trace(X_1 \circ Y + X_2 \circ Y) = \\ trace(X_1 \circ Y) + trace(X_2 \circ Y) = \hat{B}(X_1, Y) + \hat{B}(X_2, Y)$$

Likewise we have $\hat{B}(X, Y_1 + Y_2) = \hat{B}(X, Y_1) + \hat{B}(X, Y_2)$. Also for c in $\mathbb{F}$

$$\hat{B}(cX, Y) = trace((cX) \circ Y) = trace(c(X \circ Y)) = c(trace(X \circ Y)) = \\ c(\hat{B}(X,Y))$$

Likewise we have $\hat{B}(X, cY) = c(\hat{B}(X,Y))$.

Our first observation is that $\hat{B}$ is symmetric since the trace function is symmetric:

$$\hat{B}(X,Y) = trace(X \circ Y) = trace(Y \circ X) = \hat{B}(Y,X)$$

At this point we have used just the properties of a linear algebra. However when we bring in the structure of a Lie algebra, we obtain another property of the trace function called its associative structure. We use the structure of the Lie bracket to show that

$$\hat{B}([X,Y],Z) = \hat{B}(X,[Y,Z])$$

We have

$$\begin{gathered}\hat{B}([X,Y],Z) = trace([X,Y]\circ Z) = \\ trace((X\circ Y - Y\circ X)\circ Z) = \\ trace(X\circ Y\circ Z - Y\circ X\circ Z) = \\ trace(X\circ Y\circ Z) - trace(Y\circ X\circ Z) = \\ trace(X\circ Y\circ Z) - trace(X\circ Z\circ Y) = \\ trace(X\circ Y\circ Z - X\circ Z\circ Y) = \\ trace(X\circ (Y\circ Z - Z\circ Y)) = \\ trace(X\circ [Y,Z]) = \hat{B}(X,[Y,Z])\end{gathered}$$

We observe in this calculation that we used the definition and properties of the Lie bracket, and thus we would expect the form $\hat{B}$ to reflect in some manner the structure of the Lie algebra $\hat{g}$.

We will be working with the condition that $\hat{B}(X,X) = 0$ for all X in $\hat{g}$. We want to remark now that this condition is equivalent to $\hat{B}(X,Y) = 0$ for all X, Y in $\hat{g}$. Obviously if $\hat{B}(X,Y) = 0$ for all X, Y in $\hat{g}$, then $\hat{B}(X,X) = 0$ for all X in $\hat{g}$. Now suppose that $\hat{B}(X,X) = 0$ for all X in $\hat{g}$. Then we know that $\hat{B}(X+Y, X+Y) = 0$ for all X, Y in $\hat{g}$. We have $0 = \hat{B}(X+Y, X+Y) = \hat{B}(X,X) + \hat{B}(X,Y) + \hat{B}(Y,X) + \hat{B}(Y,Y) = 0 + \hat{B}(X,Y) + \hat{B}(Y,X) + 0 = 2\hat{B}(X,Y)$, by symmetry, giving $0 = 2\hat{B}(X,Y)$. But since our field of scalars is not *mod* 2, we can conclude that $\hat{B}(X,Y) = 0$ for X, Y in $\hat{g}$.

We note that we have assumed that the Lie algebra $\hat{g}$ is a Lie subalgebra of $\widehat{gl}(V)$, where V is a finite dimensional, linear space over the field $\mathbb{C}$. Thus $\hat{g}$ is a *matrix* algebra. The fundamental and difficult theorem about $\hat{B}$ that we wish to prove is the following. (We will call this theorem *Theorem* $\hat{B}$).

> Let $\hat{g}$ be a subalgebra of $\widehat{gl}(V)$. Let the form $\hat{B}$ satisfy the condition that for all X in $D^1\hat{g}$ and for all Y in $\hat{g}$, $\hat{B}(X,Y) = 0$. Then such an X is a nilpotent linear transformation.

On first glance this seems like a formidable task since there appears to be no easy connection between these two statements. And indeed the proof is not conceptually easy. We return to linear algebra and use the result that if a linear transformation X in $\widehat{gl}(V)$ has all of its eigenvalues equal to zero, then it is linearly nilpotent, which is the conclusion we are seeking.

Since $\hat{g}$ is a subalgebra of the matrix algebra $\widehat{gl}(V)$ on $End(V)$, we can represent X in $\hat{g}$ as a matrix. Being a linear transformation, X has a characteristic polynomial. Since our scalar field is algebraically closed and the dimension of V is n, we know that the characteristic polynomial for X factors into n linear factors, which give the eigenvalues $\lambda_1, \cdots, \lambda_n$, where, of course, some eigenvalues will be repeated according to their multiplicities. Now we must show that all these eigenvalues are zero, under the supposition that for all X in $D^1\hat{g}$ and for all Y in $\hat{g}$, $\hat{B}(X, Y) = 0$. For then we can conclude that X is a nilpotent linear transformation.

In our proof we will use the Jordan Decomposition Theorem, which says that X, now written in its canonical Jordan form, can be expressed uniquely as a sum of two linear transformations, $X = S_X + N_X$, where S_X is diagonal and N_X is nilpotent, with $S_X N_X = N_X S_X$, and in which both S_X and N_X can be written as polynomials in X [with zero constant term]. Since S_X is a diagonal matrix, it has the n eigenvalues $\lambda_1, \cdots, \lambda_n$, where, of course, some eigenvalues will be repeated according to their multiplicities.

We remark the following. Suppose we use the hypothesis that for all X in $\hat{g}$, $\hat{B}(X, X) = 0$. This would give us $trace(X \circ X) = 0$. Since the trace is independent of the basis chosen to calculate it, we know from the form of the Jordan canonical form that $trace(X \circ X) = trace(S_X \circ S_X) = \lambda_1^2 + \cdots + \lambda_n^2 = 0$. But this is not sufficient to give us the conclusion that every $\lambda_i = 0$ since we are not working over the real numbers but over the complex numbers.

However, if we take the conjugate matrix $\overline{S_X}$, then we know that $trace(S_X \circ \overline{S_X}) = \lambda_1 \overline{\lambda}_1 + \cdots + \lambda_n \overline{\lambda}_n = |\lambda_1|^2 + \cdots + |\lambda_n|^2$. Of course, if we show that this sum is zero, then we can conclude that $|\lambda_i|^2 = 0$, giving $\lambda_i = 0$ for all i, which indeed is the conclusion we are seeking.

Thus the assumption that $\hat{B}(X, X) = 0$ for X in $\hat{g}$ is not sufficient. But if we assume that $\hat{B}(X, Y) = 0$ for X in $D^1\hat{g}$ and Y in $\hat{g}$, we will be able to reach our desired conclusion after a long and rather involved analysis of $\hat{g}$. But this just reveals how vital the nature of $D^1\hat{g}$ is in the structure of a Lie algebra $\hat{g}$.

We proceed as follows. What we need to do is to take advantage of the fact that the Killing form is associative in $\widehat{gl}(V)$. We do this in the following manner. Since X is in $[\hat{g}, \hat{g}]$, we can say that with respect to the basis chosen above, X is a sum of commutators $[Y_r, Z_r]$ with Y_r and Z_r in $\hat{g}$. We examine the term $trace(X \circ \overline{S_X})$. Since the trace function is linear, we need look at only $trace([Y_r, Z_r] \circ \overline{S_X})$. Now $[Y_r, Z_r]$ is indeed in $D^1\hat{g}$, but we cannot say that $\overline{S_X}$ is in $\hat{g}$, and thus we cannot use our hypothesis that $\hat{B}(X, Y) = 0$ for X in $D^1\hat{g}$ and Y in $\hat{g}$. But by using the associativity of the Killing Form, we have $trace([Y_r, Z_r] \circ \overline{S_X}) = trace(Y_r \circ [Z_r, \overline{S_X}])$, and if we can show that $[Z_r, \overline{S_X}]$ is in $D^1\hat{g}$, then our hypothesis says that $\hat{B}(Y_r, [Z_r, \overline{S_X}]) = trace(Y_r \circ [Z_r \overline{S_X}]) = 0$, which means [by associativity and linearity of the trace function] that $trace(X \circ \overline{S_X}) = 0$. This is equivalent to $trace(S_X \circ \overline{S_X}) = 0$ and this fact gives us our conclusion that all the eigenvalues of X are zero and this, in turn, says that X is a nilpotent linear transformation. Thus we need to begin with the hypothesis that X is in $D^1\hat{g}$ and not just in $\hat{g}$.

All this implies that we are now reduced to proving that $[Z_r, \overline{S_X}]$ is an element in $D^1\hat{g}$, knowing that Z_r is in $\hat{g}$.

But first, we need to back off from $\overline{S_X}$ and return to S_X, and then to X, which we know is an element in $D^1\hat{g}$. We now use some facts from linear algebra. It can be proven that ad respects the Jordan decomposition. This means the following. We know that X is in $\widehat{gl}(V) = End(V)$. Thus X has the Jordan decomposition $X = S_X + N_X$. Now $ad(X)$ is in $\widehat{gl}(\widehat{gl}(V)) = End(\widehat{gl}(V))$, and thus is again a linear transformation. It therefore has a Jordan decomposition $ad(X) = S_{ad(X)} + N_{ad(X)}$. Now when we say that ad respects the Jordan decomposition, we mean $ad(X) = ad(S_X + N_X) = ad(S_X) + ad(N_X) = S_{ad(X)} + N_{ad(X)}$, with $ad(S_X) = S_{ad(X)}$ and $ad(N_X) = N_{ad(X)}$. Now we use the fact that in the Jordan decomposition $S_{ad(X)}$ is a polynomial in $ad(X)$ without constant term, i.e., we can express $S_{ad(X)}$ as:

$$S_{ad(X)} = c_1 ad(X) + c_2(ad(X))^2 + \cdots + c_s(ad(X))^s$$

Now we know that

$$\begin{gathered}[S_X, Z_r] = ad(S_X)Z_r = S_{ad(X)}Z_r = \\ (c_1 ad(X) + c_2(ad(X))^2 + \cdots + c_s(ad(X))^s)Z_r = \\ c_1 ad(X)Z_r + c_2((ad(X))^2)Z_r + \cdots + c_s((ad(X))^s)Z_r = \\ c_1[X, Z_r] + c_2[X, [X, Z_r]] + \cdots + c_s[X, [\cdots[X, Z_r]\cdots]]\end{gathered}$$

But we also know that X is in $D^1\hat{g}$ and Z_r is in $\hat{g}$. Thus $[X, Z_r] = ad(X)Z_r$ is in $D^1\hat{g}$. We conclude that $[S_X, Z_r]$ is in $D^1\hat{g}$.

Now we can show that $[\overline{S_X}, Z_r]$ is also in $D^1\hat{g}$. We recall that we have so chosen a basis in V that X is in its Jordan canonical form. This makes S_X a diagonal matrix. We now take advantage of the fact from linear algebra that, knowing that S_X is in the form of a diagonal matrix, then $ad(S_X) = S_{ad(X)}$ is also in the form of a diagonal matrix. Thus since S_X is a diagonal matrix, then the conjugate of S_X, $\overline{S_X}$, is a diagonal matrix, and $ad(\overline{S_X}) = \overline{ad(S_X)} = \overline{S_{ad(X)}}$. But now we know that $\overline{S_{ad(X)}}$ can be written as a polynomial in $S_{ad(X)}$ [without constant term]:

$$\overline{S_{ad(X)}} = d_1 ad(S_X) + d_2(ad(S_X))^2 + \cdots + d_t(ad(S_X))^t$$

Thus, in our chosen basis, we have

$$\begin{gathered}ad(\overline{S_X}) \cdot Z_r = (d_1 ad(S_X) + d_2(ad(S_X))^2 + \cdots + d_t(ad(S_X))^t) \cdot Z_r = \\ d_1 ad(S_X)Z_r + d_2(ad(S_X))^2)Z_r + \cdots + d_t(ad(S_X))^t)Z_r\end{gathered}$$

And we know that $ad(S_X) \cdot Z_r$ is in $D^1\hat{g}$. We conclude that $ad(\overline{S_X}) \cdot Z_r$ is in $D^1\hat{g}$, and we have finally reached the conclusion that we have been seeking. Thus *Theorem* $\hat{B}$ has been proven.

We think that this proof is an exciting and remarkable proof. At so many junctures we seemed to have been foiled, but there was always another fact that we had not yet used, Using them, one by one, we eventually won over the day. Now before we continue we want to comment more on the three steps above which we merely quoted: 1) the fact that the adjoint representation respects the Jordan decomposition; 2) if S_X is a diagonal matrix, then $ad(S_X) = S_{ad(X)}$ is also a diagonal matrix. 3) the fact that the matrix $\overline{S_{ad(X)}} = \overline{ad(S_X)} = ad(\overline{S_X})$ is diagonal, and that $\overline{S_{ad(X)}}$ can be written as a polynomial in $S_{ad(X)}$ without constant term. We will not prove these statements, since the proofs are straightforward but tedious, or they pertain to other parts of mathematics than Lie algebras. But we do want to construct an example involving a space of high enough dimension to show how one would construct a general proof in these cases.

First, we examine the fact that the adjoint representation respects the Jordan decomposition. Thus, for a linear transformation X in $\widehat{gl}(V)$, we have its Jordan Decomposition $X = S_X + N_X$, where S_X is diagonalizable and N_X is a nilpotent linear transformation. Now the adjoint representation of $\widehat{gl}(V)$ puts X into $ad(X)$ in $\widehat{gl}(\widehat{gl}(V))$, which is again a set of linear transformations. Thus $ad(X)$ has a Jordan decomposition $ad(X) = S_{ad(X)} + N_{ad(X)}$. We want to assert that $ad(S_X) = S_{ad(X)}$ and $ad(N_X) = N_{ad(X)}$.

We now write X in a basis which gives X the Jordan canonical form. This means that S_X is realized as a diagonal matrix with the eigenvalues of X, including repetitions, down the diagonal; and N_X is realized as a nilpotent linear matrix, i.e., an upper triangular matrix with a zero diagonal. Since we are working on an example here in dimension three, we need to use the canonical basis of the 3x3 matrices, E_{ij}. Thus

$$S_X = \lambda_1 E_{11} + \lambda_2 E_{22} + \lambda_3 E_{33}$$

where the λ_i are the three eigenvalues of X. Now

$$ad(S_X) = \lambda_1 ad(E_{11}) + \lambda_2 ad(E_{22}) + \lambda_3 ad(E_{33})$$

Thus we need to evaluate $ad(E_{ii})$ on the basis E_{ij}:

$$ad(E_{kk}) \cdot E_{ij} = [E_{kk}, E_{ij}]$$
$$= E_{kj}, \text{ if } i = k; \qquad = -E_{ik}, \text{ if } j = k; \qquad \text{and } = 0 \text{ otherwise.}$$

Our job is to find eigenvalues that work. Thus, we form the 9x9 matrices, where we have chosen the nine basis vectors in the following order:

$$(E_{11}, E_{21}, E_{31}, E_{12}, E_{22}, E_{32}, E_{13}, E_{23}, E_{33})$$

We begin with $\lambda_1 ad(E_{11})$.

$$\lambda_1 ad(E_{11}) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Next we calculate $\lambda_2 ad(E_{22})$.

$$\lambda_2 ad(E_{22}) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\lambda_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\lambda_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Finally we form $\lambda_3 ad(E_{33})$.

$$\lambda_3 ad(E_{33}) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\lambda_3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\lambda_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

We now have

$$ad(S_X) =$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_2-\lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_3-\lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_1-\lambda_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_3-\lambda_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_1-\lambda_3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_2-\lambda_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus we have the beautiful conclusion that $ad(S_X)$ is a diagonal matrix, and thus is a good candidate for S_{adX}.

We now calculate $ad(N_X)$. Recall that we have so chosen the basis for $\hat{g}$ so that $X = S_X + N_X$ is in its Jordan canonical form. Thus S_X is diagonal and N_X is an upper triangular matrix with zero diagonal and with the only non-zero terms being ones just above the diagonal. Since we are in the case of three dimensions, let us suppose the N_X matrix has the form

$$N_X = \begin{bmatrix} 0 & a_1 & 0 \\ 0 & 0 & a_2 \\ 0 & 0 & 0 \end{bmatrix}$$

where a_1 and a_2 are either zero or one. Thus we suppose that we start with a matrix X with Jordan canonical form

$$X = \begin{bmatrix} \lambda_1 & a_1 & 0 \\ 0 & \lambda_2 & a_2 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

which has the three eigenvalues λ_1,λ_2,λ_3. Thus we can write

$$N_X = a_1 E_{12} + a_2 E_{23}$$

$$ad(N_X) = a_1 ad(E_{12}) + a_2 ad(E_{23})$$

Proceeding, we need to evaluate $ad(E_{12})$ and $ad(E_{23})$ on the basis E_{ij}. We have

$$ad(E_{12}) \cdot E_{ij} = [E_{12}, E_{ij}]$$
$$= E_{1j}, \text{ if } i = 2; \quad = -E_{i2}, \text{ if } j = 1; \quad \text{and } = 0 \text{ otherwise}$$
$$ad(E_{23}) \cdot E_{ij} = [E_{23}, E_{ij}]$$
$$= E_{2j}, \text{ if } i = 3; \quad = -E_{i3}, \text{ if } j = 2; \quad \text{and } = 0 \text{ otherwise}$$

Thus we have:

$$a_1 ad(E_{12}) = \begin{bmatrix} 0 & a_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -a_1 & 0 & 0 & 0 & a_1 & 0 & 0 & 0 & 0 \\ 0 & -a_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -a_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$a_2 ad(E_{23}) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -a_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -a_2 & 0 & 0 & 0 & a_2 \\ 0 & 0 & 0 & 0 & 0 & -a_2 & 0 & 0 & 0 \end{bmatrix}$$

and this gives

$$ad(N_X) = \begin{bmatrix} 0 & a_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -a_1 & 0 & 0 & 0 & a_1 & 0 & 0 & 0 & 0 \\ 0 & -a_1 & 0 & 0 & 0 & a_2 & 0 & 0 & 0 \\ 0 & 0 & -a_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -a_2 & 0 & 0 & 0 & a_1 & 0 \\ 0 & 0 & 0 & 0 & -a_2 & 0 & 0 & 0 & a_2 \\ 0 & 0 & 0 & 0 & 0 & -a_2 & 0 & 0 & 0 \end{bmatrix}$$

We observe that $ad(N_X)$ is not in an obvious linear nilpotent form. since it is not in upper [or lower] triangular form with a zero diagonal. However it does have a zero diagonal and if we take powers of $ad(N_X)$:

$$(ad(N_X))^2 = \begin{bmatrix} 0 & 0 & a_1a_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -2a_1^2 & 0 & 0 & 0 & a_1a_2 & 0 & 0 & 0 \\ 0 & 0 & -2a_1a_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_1a_2 & 0 & 0 & 0 & -2a_1a_2 & 0 & 0 & 0 & a_1a_2 \\ 0 & a_1a_2 & 0 & 0 & 0 & -2a_2^2 & 0 & 0 & 0 \\ 0 & 0 & a_1a_2 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$(ad(N_X))^3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -3a_1^2a_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3a_1^2a_2 & 0 & 0 & 0 & -3a_1a_2^2 & 0 & 0 & 0 \\ 0 & 0 & 3a_1a_2^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

we find that $(ad(N_X))^3$ does have a lower triangular form with zero diagonal, and thus $ad(N_X)$ is a nilpotent linear transformation, since $(ad(N_X))^4 \neq 0$, but $(ad(N_X))^5 = 0$.

$$(ad(N_X))^4 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6a_1^2a_2^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$(ad(N_X))^5 = 0$$

We can therefore affirm that $ad(N_X)$ is a candidate for $N_{ad(X)}$. But we need to check the commutativity of $ad(S_X)$ and $ad(N_X)$. We know that $S_XN_X = N_XS_X$. Now we have

$$S_X = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \qquad N_X = \begin{bmatrix} 0 & a_1 & 0 \\ 0 & 0 & a_2 \\ 0 & 0 & 0 \end{bmatrix}$$

where the λ_i's can have repetitions and a_1 and a_2 are either one or zero. We calculate

$$S_X N_X - N_X S_X = \begin{bmatrix} 0 & \lambda_1 a_1 - \lambda_2 a_1 & 0 \\ 0 & 0 & \lambda_2 a_2 - \lambda_3 a_2 \\ 0 & 0 & 0 \end{bmatrix} =$$
$$\begin{bmatrix} 0 & (\lambda_1 - \lambda_2) a_1 & 0 \\ 0 & 0 & (\lambda_2 - \lambda_3) a_2 \\ 0 & 0 & 0 \end{bmatrix}$$

which we know must be the zero matrix. Thus we now need to find what values of the λ_i's and the a_j's are possible in order for $0 = S_X N_X - N_X S_X$ to be true. We have three possibilities. The first is to have λ_1, λ_2, and λ_3 all distinct. Then we know that S_X is diagonal with no repetitions and thus $\lambda_1 - \lambda_2 \neq 0$ and $\lambda_2 - \lambda_3 \neq 0$. Thus the only way $S_X N_X - N_X S_X$ can be 0 is if $a_1 = 0$ and $a_2 = 0$. In this case we must choose N_X to be the zero matrix and we have $S_X N_X - N_X S_X = 0$. The second possibility is to have one repetition of eigenvalues, say: $\lambda_1 = \lambda_2$ with $\lambda_3 \neq \lambda_1$. In this case $a_1 = 1$ and $a_2 = 0$, giving again $S_X N_X - N_X S_X = 0$. The third possibility is to have all three eigenvalues identical and thus $a_1 = a_2 = 1$ can be the choice giving $S_X N_X - N_X S_X = 0$. Thus in all three cases we can find values so that $S_X N_X = N_X S_X$. Given these values we now need to calculate $ad(S_X)ad(N_X) - ad(N_X)ad(S_X)$.

$$ad(S_X) =$$
$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\lambda_1 + \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 & \\ 0 & 0 & -\lambda_1 + \lambda_3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_1 - \lambda_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\lambda_2 + \lambda_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_1 - \lambda_3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_2 - \lambda_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$ad(N_X) = \begin{bmatrix} 0 & a_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -a_1 & 0 & 0 & 0 & a_2 & 0 & 0 & 0 & 0 \\ 0 & -a_1 & 0 & 0 & 0 & a_2 & 0 & 0 & 0 \\ 0 & 0 & -a_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -a_2 & 0 & 0 & 0 & a_1 & 0 \\ 0 & 0 & 0 & 0 & a_2 & 0 & 0 & 0 & a_2 \\ 0 & 0 & 0 & 0 & 0 & -a_2 & 0 & 0 & 0 \end{bmatrix}$$

and

$$ad(S_X)ad(N_X) - ad(N_X)ad(N_X) =$$

$$\begin{bmatrix} 0 & -a_1(-\lambda_1+\lambda_2) & 0 & 0 & 0 & \cdot\cdot\cdot\cdot \\ 0 & 0 & a_2(\lambda_2-\lambda_3) & 0 & 0 & \cdot\cdot\cdot\cdot \\ 0 & 0 & 0 & 0 & 0 & \cdot\cdot\cdot\cdot \\ a_1(-\lambda_1+\lambda_2) & 0 & 0 & 0 & a_1(\lambda_1-\lambda_2) & \cdot\cdot\cdot\cdot \\ 0 & a_1(-\lambda_1+\lambda_2) & 0 & 0 & 0 & \cdot\cdot\cdot\cdot \\ 0 & 0 & a_1(-\lambda_1+\lambda_2) & 0 & 0 & \cdot\cdot\cdot\cdot \\ 0 & 0 & 0 & a_2(-\lambda_2+\lambda_3) & 0 & \cdot\cdot\cdot\cdot \\ 0 & 0 & 0 & 0 & -a_2(\lambda_2-\lambda_3) & \cdot\cdot\cdot\cdot \\ 0 & 0 & 0 & 0 & 0 & \cdot\cdot\cdot\cdot \end{bmatrix}$$

—-

$$\begin{bmatrix} \cdot\cdot\cdot\cdot\cdot & 0 & 0 & 0 & 0 \\ \cdot\cdot\cdot\cdot\cdot & 0 & 0 & 0 & 0 \\ \cdot\cdot\cdot\cdot\cdot & 0 & 0 & 0 & 0 \\ \cdot\cdot\cdot\cdot\cdot & 0 & 0 & 0 & 0 \\ \cdot\cdot\cdot\cdot\cdot & -a_2(-\lambda_2+\lambda_3) & 0 & 0 & 0 \\ \cdot\cdot\cdot\cdot\cdot & 0 & 0 & 0 & 0 \\ \cdot\cdot\cdot\cdot\cdot & 0 & 0 & a_1(\lambda_1-\lambda_2) & 0 \\ \cdot\cdot\cdot\cdot\cdot & 0 & 0 & 0 & a_2(\lambda_2-\lambda_3) \\ \cdot\cdot\cdot\cdot\cdot & a_2(-\lambda_2+\lambda_3) & 0 & 0 & 0 \end{bmatrix}$$

We now examine the following three cases that basically take care of all possible cases:

1) $\lambda_1 \neq \lambda_2 \neq \lambda_3$ with $a_1 = a_2 = 0$
2) $\lambda_1 = \lambda_2 \neq \lambda_3$ with $a_1 \neq 0$ and $a_2 = 0$
3) $\lambda_1 = \lambda_2 = \lambda_3$ with $a_1, a_2 \neq 0$

And we observe that in each of the three cases we have $ad(S_X)ad(N_X) - ad(N_X)ad(S_X) = 0$, giving us desired the commutativity, i.e.:

$$ad(S_X)ad(N_X) = ad(N_X)ad(S_X)$$

Thus, with the right choices of eigenvalues and off diagonals, we can find an X in a 3-dimensional $\hat{g}$ where ad does preserve the Jordan decomposition of X and $ad(X)$:

$$X = S_X + N_X \longrightarrow ad(X) = ad(S_X + N_X) = ad(S_X) + ad(N_X) = S_{ad(X)} + N_{ad(X)}$$

with, of course, $ad(S_X) = S_{ad(X)}$ and $ad(N_X) = N_{ad(X)}$. [We remark that we have not examined the other two properties of the Jordan decomposition, i.e., 1) the uniqueness of the decomposition; 2) the fact that the diagonalizable part and the linear nilpotent part are each polynomials (without constant term) in the linear transformation X.]

The example given above shows how we could come to this same conclusion in general. One observation that can be made from this example is that ad does not preserve the Jordan canonical form. It does take the diagonal matrix S_X over to the diagonal matrix $S_{ad(X)}$, but the nilpotent matrix $N_{ad(X)}$ does not have the correct form. This should be expected since in forming N_X, we had to choose a very special basis in $\hat{g}$, but in forming $N_{ad(X)}$, we just used the given canonical basis (E_{ij}) in $ad(\hat{g})$, and it would be surprising if that produced the special basis needed in $ad(\hat{g})$ to form its Jordan canonical form.

However we note that the commutativity relationship can easily be proven in general. We start with the fact that $S_X N_X = N_X S_X$ and we want to show that

$$S_{ad(X)}N_{ad(X)} = N_{ad(X)}S_{ad(X)}$$

We operate these expressions on an arbitrary element Z in $\widehat{gl}(V)$.

$$\begin{gathered} S_{ad(X)}N_{ad(X)} \cdot Z = ad(S_X)ad(N_X) \cdot Z = [S_X, [N_X, Z]] = [S_X, N_X Z - Z N_X] = \\ [S_X, N_X Z] - [S_X, Z N_X] = S_X N_X Z - N_X Z S_X - S_X Z N_X + Z N_X S_X = \\ N_X S_X Z - N_X Z S_X - S_X Z N_X + Z S_X N_X = \\ N_X(S_X Z - Z S_X) + (-S_X Z + Z S_X)N_X = \\ N_X([S_X, Z]) - ([S_X, Z])N_X = [N_X, [S_X, Z]] = ad(N_X)ad(S_X) \cdot Z = \\ N_{ad(X)}S_{ad(X)} \cdot Z \end{gathered}$$

from which we can conclude that $S_{ad(X)}N_{ad(X)} = N_{ad(X)}S_{ad(X)}$.

Finally there is one more observation we should make. We used in the proof the important step that $ad(\overline{S_X}) = \overline{ad(S_X)}$. But just examining the following expressions

$$\begin{aligned} S_X &= \lambda_1 E_{11} + \lambda_2 E_{22} + \lambda_3 E_{33} \\ ad(S_X) &= \lambda_1 ad(E_{11}) + \lambda_2 ad(E_{22}) + \lambda_3 ad(E_{33}) \end{aligned}$$

and

$$\begin{aligned} \overline{S_X} &= \overline{\lambda_1} E_{11} + \overline{\lambda_2} E_{22} + \overline{\lambda_3} E_{33} \\ ad(\overline{S_X}) &= \overline{\lambda_1} ad(E_{11}) + \overline{\lambda_2} ad(E_{22}) + \overline{\lambda_3} ad(E_{33}) \end{aligned}$$

we immediately see that we have the desired conclusion $ad(\overline{S_X}) = \overline{ad(S_X)}$.

We still have two more comments to make before we can close down this proof. The second comment stated that if S_X is a diagonal matrix, then $ad(S_X) = S_{ad(X)}$ is also a diagonal matrix. And in the three-dimensional case above we have shown the validity of this relationship. Finally, the third comment used the fact that for the diagonal matrix $\overline{S_{ad(X)}}$ we have $\overline{S_{ad(X)}} = \overline{ad(S_X)}$ and that it can be written as a polynomial in $S_{ad(X)}$ without constant term. This latter is just the expression that the conjugate of any n-dimensional complex diagonal matrix can be written as a linear combination of that n-dimensional complex diagonal matrix without constant term. We give a few examples of this phenomenon, which, though they do not give a general proof, will suffice for our exposition.

When we are working in $\mathbb{C}$, it is nothing but the relationship $\overline{c}c = |c|^2$.

$$\overline{c} = \tfrac{|c|^2}{c} = (\tfrac{\overline{c}}{c})c$$

Using real notation, we have

$$c_1 - c_2 i = \tfrac{c_1^2 + c_2^2}{c_1 + c_2 i} = (\tfrac{c_1 - c_2 i}{c_1 + c_2 i})(c_1 + c_2 i)$$

This says that

$$c_1 - c_2 i = (a_1 + a_2 i)(c_1 + c_2 i) = (a_1 c_1 - a_2 c_2) + i(a_1 c_2 + a_2 c_1)$$

and gives the linear equations

$$\begin{aligned} c_1 &= c_1 a_1 - c_2 a_2 \\ -c_2 &= c_2 a_1 + c_1 a_2 \end{aligned}$$

whose unique solution is $a_1 = \frac{c_1^2 - c_2^2}{c_1^2 + c_2^2}$; $a_2 = \frac{-2c_1c_2}{c_1^2 + c_2^2}$.

We observe that

$$a_1 + a_2 i = \frac{c_1^2 - c_2^2}{c_1^2 + c_2^2} + i(\frac{-2c_1c_2}{c_1^2 + c_2^2}) = \frac{(c_1^2 - c_2^2) + i(-2c_1c_2)}{c_1^2 + c_2^2} = \frac{(c_1 - c_2 i)^2}{c_1^2 + c_2^2} = \frac{(c_1 - c_2 i)^2}{(c_1 + c_2 i)(c_1 - c_2 i)} = \frac{c_1 - c_2 i}{c_1 + c_2 i}$$

which is exactly the coefficient of $c_1 + c_2 i$ that gives $c_1 - c_2 i$.

When we are working in $\mathbb{C}^2$, we see how we can generalize the above procedure. We seek $(a_1 + a_2 i)$ and $(b_1 + b_2 i)$ such that the following matrix equation is satisfied:

$$\begin{bmatrix} c_1 - c_2 i & 0 \\ 0 & d_1 - d_2 i \end{bmatrix} = (a_1 + a_2 i)\begin{bmatrix} c_1 + c_2 i & 0 \\ 0 & d_1 + d_2 i \end{bmatrix} + (b_1 + b_2 i)\begin{bmatrix} c_1 + c_2 i & 0 \\ 0 & d_1 + d_2 i \end{bmatrix}^2$$

However, instead of working in general we will show the procedure for a specific numerical example. This will be sufficient to guide us to the generalization. Our example is

$$\begin{bmatrix} 1 - 2i & 0 \\ 0 & 3 - 4i \end{bmatrix} = (a_1 + a_2 i)\begin{bmatrix} 1 + 2i & 0 \\ 0 & 3 + 4i \end{bmatrix} + (b_1 + b_2 i)\begin{bmatrix} 1 + 2i & 0 \\ 0 & 3 + 4i \end{bmatrix}^2 = (a_1 + a_2 i)\begin{bmatrix} 1 + 2i & 0 \\ 0 & 3 + 4i \end{bmatrix} + (b_1 + b_2 i)\begin{bmatrix} -3 + 4i & 0 \\ 0 & -7 + 24i \end{bmatrix}$$

which gives the following equations:

$$\begin{aligned} 1 &= 1a_1 - 2a_2 - 3b_1 - 4b_2 \\ -2 &= 2a_1 + 1a_2 + 4b_1 - 3b_2 \\ 3 &= 3a_1 - 4a_2 - 7b_1 - 24b_2 \\ -4 &= 4a_1 + 3a_2 + 24b_1 - 7b_2 \end{aligned}$$

which have the unique solution

$$a_1 = \tfrac{-22}{25} \qquad a_2 = \tfrac{-19}{25} \qquad b_1 = \tfrac{1}{25} \qquad b_2 = \tfrac{-3}{25}$$

Thus we have

$$\begin{bmatrix} 1-2i & 0 \\ 0 & 3-4i \end{bmatrix} =$$

$$(-\tfrac{22}{25} - \tfrac{19}{25}i)\begin{bmatrix} 1+2i & 0 \\ 0 & 3+4i \end{bmatrix} + (\tfrac{1}{25} - \tfrac{3}{25}i)\begin{bmatrix} 1+2i & 0 \\ 0 & 3+4i \end{bmatrix}^2$$

which expansion shows how the conjugate diagonal matrix can be expressed as a polynomial [without constant term] in the diagonal matrix.

And thus for a Lie algebra $\hat{g}$, a Lie subalgebra of $\widehat{gl}(V)$, where V is a linear space over the field $\mathbb{C}$, we can affirm the fundamental theorem *Theorem* $\hat{B}$:

> Let the form $\hat{B}$ satisfy the condition that for all X in $D^1\hat{g}$ and all Y in $\hat{g}$, $\hat{B}(X,Y) = 0$. Then X is a nilpotent linear transformation.

2.11.2 Two Theorems for Solvable and Semisimple Lie Algebras.

The first modification of $\hat{B}$ is to define the form that has traditionally been called the *Killing Form*, and to give it the symbol B. It is said to be the Killing form B of a finite dimensional Lie algebra $\hat{g}$ over a field $\mathbb{F}$ of characteristic 0. Thus it differs from $\hat{B}$ by the Lie algebra over which it is defined. $\hat{B}$ was defined over a subalgebra $\hat{g}$ of the Lie algebra $\widehat{gl}(V)$ of brackets in $End(V)$ of a linear space V over a field $\mathbb{F}$ of characteristic 0. Now B starts with an arbitrary Lie algebra $\hat{g}$ over a field $\mathbb{F}$ of characteristic 0, and is defined as a bilinear form over $\hat{g}$ by using the adjoint map ad from $\hat{g}$ into the Lie subalgebra $ad(\hat{g})$ of the Lie algebra of brackets $\widehat{gl}(\hat{g})$. Thus, after choosing a basis for $\hat{g}$, we have

$$\begin{aligned} B : \hat{g} \times \hat{g} &\longrightarrow \mathbb{F} \\ (x,y) &\longmapsto B(x,y) := trace(ad(x) \circ ad(y)) \end{aligned}$$

This says that $\hat{B}$ is defined on the Lie subalgebra $ad(\hat{g})$ of the Lie algebra $\widehat{gl}(\hat{g})$, that is, $B(x,y) = \hat{B}(ad(x), ad(y))$. Now since ad is linear on $\hat{g}$, i.e.,

$$ad(x+y) = ad(x) + ad(y) \qquad ad(cx) = cad(x)$$

we see immediately from the bilinearity of $\hat{B}$ that B is also bilinear on $\hat{g}$. Finally since the adjoint ad is a homomorphism of Lie algebras, i.e.,

$$ad[x,y] = [ad(x), ad(y)]$$

we also see that the associative property of $\hat{B}$ is preserved, i.e.,

$$B([x,y],z) = B(x,[y,z])$$

We remark that since this homomorphism is just another way of writing the Jacobi identity of $\hat{g}$, we would expect that the Killing form B would reflect the structure of the Lie algebra $\hat{g}$.

With this new tool of the Killing form we can prove two remarkable theorems:

> A Lie algebra $\hat{g}$ over $\mathbb{R}$ or $\mathbb{C}$ is solvable if and only if the Killing form $B(x, x) = 0$ for all x in $D^1\hat{g}$.

and

> A Lie algebra $\hat{g}$ over $\mathbb{R}$ or $\mathbb{C}$ is semisimple if and only if its Killing form B is nondegenerate.

These will be proved below but their proofs require much preparation.

2.11.3 Solvable Lie Algebra $\hat{\mathbf{s}}$ over $\mathbb{C}$ Implies $\mathbf{D^1\hat{s}}$ is a Nilpotent Lie Algebra. Before we get to these proofs, we pause a moment to prove something that we will use in the following development. First we show that for $\hat{s}$, a solvable Lie algebra over $\mathbb{C}$, $ad(D^1\hat{s})$ is a set of nilpotent linear transformations in $\widehat{gl}(\hat{s})$. We have seen this phenomenon many times in our examples above and now we would like to confirm it in general. By definition $ad(D^1\hat{s}) = ad([\hat{s}, \hat{s}]) \subset [ad(\hat{s}), ad(\hat{s})]$. Since $\hat{s}$ is solvable and ad is a homomorphism of Lie algebras, we know that $ad(\hat{s})$ is a solvable Lie algebra of linear transformations in $\widehat{gl}(\hat{s})$. Thus using Lie's Theorem, we know that we can find a basis for $\hat{s}$ such that all the linear transformations in $ad(\hat{s})$ are represented as upper triangular matrices. We now take the brackets of these matrices. Let (a_{ij}) be an i-th row of the matrix A in $ad(\hat{s})$ and let (b_{ji}) be the i-th column of the matrix B in $ad(\hat{s})$. We calculate the diagonal element $c_{ii} = \sum_j (a_{ij}) \cdot (b_{ji})$ of the product matrix $A \cdot B$. Since we are dealing with upper triangular matrices, the know the 0 elements of the row (a_{ij}) are $(a_{i,1}, a_{i,2}, \cdots, a_{i,i-1})$. And also we know that the 0 elements of the column (b_{ji}) are $(b_{i+1,i}, b_{i+2,i}, \cdots, b_{n,i})$ [where n is the dimension of $\hat{s}$]. Thus $c_{ii} = \sum_j (a_{ij}) \cdot (b_{ji}) = a_{ii} \cdot b_{ii}$. Now we do the same for the product $B \cdot A$. Let (b_{ij}) be a i-th row of the matrix B and (a_{ji}) be the i-th column of the matrix A. We calculate the diagonal element $d_{ii} = \sum_j (b_{ij}) \cdot (a_{ji})$ of the product matrix $B \cdot A$. It is obvious that $d_{ii} = b_{ii} \cdot a_{ii}$. Thus we can conclude that all the entries on the diagonal of the bracket $[A, B] = AB - BA$ are equal to zero, giving us an upper triangular matrix with a zero diagonal. We conclude that $[A, B] \in [ad(\hat{s}), ad(\hat{s})]$ is a nilpotent linear transformation in $\widehat{gl}(\hat{s})$. Thus by 2.8.2 we can conclude that $D^1\hat{s}$ is a nilpotent Lie algebra. [Since we used Lie's Theorem in the proof, this conclusion now applies only to Lie algebras over $\mathbb{C}$.]

2.12 Changing the Scalar Field

We focus now on the proof of

> A Lie algebra $\hat{g}$ over $\mathbb{R}$ or $\mathbb{C}$ is solvable if and only if the Killing form $B(x, x) = 0$ for all x in $D^1\hat{g}$

We see that we are affirming that this theorem is true if the scalar field for the Lie algebra is either $\mathbb{R}$ or $\mathbb{C}$. However in proving this theorem we will need to use Lie's Theorem which says

> Let $\hat{s}$ be a solvable complex Lie subalgebra of $\widehat{gl}(V)$. Then there exists a nonzero vector $v \in V$ which is a simultaneous eigenvector for all X in $\hat{s}$ [with eigenvalue dependent on X].

and therefore we need $\mathbb{C}$, an algebraically closed field, in that proof. Thus we must separate the proof for a Lie algebra over $\mathbb{C}$ from the proof for a Lie algebra over $\mathbb{R}$. But obviously the two scalar fields are connected to one another and thus it will be natural to ask when one begins with one field what information carries over to the other.

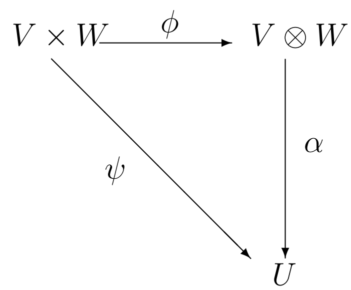
2.12.1 From $\mathbb{R}$ to $\mathbb{C}$: the Linear Space Structure. We first move from a $\mathbb{R}$-linear space to a $\mathbb{C}$-linear space. Thus, given a real linear space, we ask what kind of complex linear structures does it determine? In this part of our exposition we will assume many details from linear algebra. Our initial task is to set up notation. First of all, we need a knowledge of the tensor product in linear algebra. Thus, we let V and W be two finite dimensional linear spaces over a field $\mathbb{F}$ whose characteristic is 0, with dim $V = m$ and dim $W = n$. We then form the free linear space $Free(V \times W)$ on the Cartesian product $V \times W$, that is, we take all pairs (v, w) in $V \times W$ as a basis over the scalars $\mathbb{F}$. This forms an infinite dimensional linear space $Free(V \times W)$ over $\mathbb{F}$. We then place relations on this space by defining a subspace N generated by elements of the form:

$$(v_1 + v_2, w) - (v_1, w) - (v_2, w) \qquad (v, w_1 + w_2) - (v, w_1) - (v, w_2)$$
$$(cv, w) - c(v, w) \qquad (v, cw) - c(v, w)$$

with $v, v_1, v_2 \in V$, $w, w_1, w_2 \in W$, and $c \in \mathbb{F}$. Now we define a <u>bilinear</u> space over $\mathbb{F}$: $V \otimes W := Free(V \times W)/N$. The image of (v, w) in $Free(V \times W)$ by the projection of $Free(V \times W)$ on $V \otimes W$ is given the symbol $v \otimes w$, which is read "v tensor w". We note that we have immediately the following <u>bilinear</u> relations on $V \otimes W$:

$$(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w \qquad v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2$$
$$(cv) \otimes w = c(v \otimes w) = v \otimes (cw)$$

This process essentially is the multilinearization of linear spaces. In our case we take a typical bilinear map ψ of $V \times W$ to a linear space U and express it by defining a bilinear map ϕ of $V \times W$ to a newly defined bilinear space $V \otimes W$ and a newly defined bilinear map α of $V \otimes W$ to the linear space U. This map is such that $\psi = \alpha \circ \phi$.



We can easily show that if V has a basis $(v_1, \cdots, v_m)$ and W has a basis $(w_1, \cdots, w_n)$, then $V \otimes W$ has a basis $(v_1 \otimes w_1, \cdots, v_i \otimes w_j, \cdots, v_m \otimes w_n)$, where i runs from 1 to m and j runs from 1 to n. Thus the dimension of $V \otimes W$ is mn, the product of the dimensions of V and of W. (We can, in this vein, also use for any scalar field $\mathbb{F}$ and any linear space V over $\mathbb{F}$ the identification $\mathbb{F} \otimes V \cong V$ by $r \otimes v \longmapsto rv$.)

Recall that bilinearity means that the map is linear in both factors:

$$\begin{gathered}
(x_1 + x_2, y) \longmapsto \phi(x_1 + x_2, y) = (x_1 + x_2) \otimes y = x_1 \otimes y + x_2 \otimes y \\
(ax, y) \longmapsto \phi(ax, y) = ax \otimes y = a(x \otimes y) \\
(x, y_1 + y_2) \longmapsto \phi(x, y_1 + y_2) = x \otimes (y_1 + y_2) = x \otimes y_1 + x \otimes y_2 \\
(x, by) \longmapsto \phi(x, by) = (x \otimes by) = b(x \otimes y)
\end{gathered}$$

and

$$\begin{gathered}
(x_1 + x_2, y) \longmapsto \psi(x_1 + x_2, y) = \psi(x_1, y) + \psi(x_2, y) \\
(ax, y) \longmapsto \psi(ax, y) = a(\psi(x, y)) \\
(x, y_1 + y_2) \longmapsto \psi(x, y_1 + y_2) = \psi(x, y_1) + \psi(x, y_2) \\
(x, by) \longmapsto \psi(x, by) = b(\psi(x, y))
\end{gathered}$$

and thus we have

$$\begin{gathered}
\alpha((x_1 + x_2) \otimes y) = \psi(x_1, y) + \psi(x_2, y); \alpha(ax \otimes y)) = a(\psi(x, y)) \\
\alpha(x \otimes (y_1 + y_2)) = \psi(x, y_1) + \psi(x, y_2); \alpha((x \otimes by)) = b(\psi(x, y))
\end{gathered}$$

We begin with the field $\mathbb{R}$ of characteristic 0 whose algebraic closure is the field$\mathbb{C}$]. Suppose now we have a linear space V over $\mathbb{R}$. We want to define a linear space V^c over $\mathbb{C}$ [called the complexification of V] by using tensor products. Now to make sense of a tensor product, we need both factors of the tensor product to be linear spaces over the same field. Since $\mathbb{C}$ is a linear space over $\mathbb{C}$, and since $\mathbb{R}$ is a subfield of $\mathbb{C}$, we know that $\mathbb{C}$ is also a linear space over $\mathbb{R}$. We give this real linear space the symbol $\mathbb{C}^r$ to distinguish it from the same set with its complex linear space structure, that is, from $\mathbb{C}$. Now we can define a linear space V^c over $\mathbb{C}$ by using tensor products.

This means that V^c is the same set of elements as $\mathbb{C}^r \otimes_{\mathbb{R}} V$. Also since $\mathbb{C}^r \otimes_{\mathbb{R}} V$ has an additive structure, V^c repeats this additive structure. What is different is the scalar multiplication structure. V^c is to be a $\mathbb{C}$-linear space, while $\mathbb{C}^r \otimes_{\mathbb{R}} V$ is a $\mathbb{R}$-linear space. Thus we have to make $\mathbb{C}^r \otimes_{\mathbb{R}} V$ into a $\mathbb{C}$-linear space.

We now have an $\mathbb{R}$-linear space $\mathbb{C}^r \otimes_{\mathbb{R}} V$. [We might remark that this notation is frequently shortened to $\mathbb{C} \otimes_{\mathbb{R}} V$, where by the ambiguous $\mathbb{C}$ we mean the $\mathbb{R}$-linear space $\mathbb{C}^r \otimes_{\mathbb{R}} V$. Since the full symbol is $\mathbb{C}^r \otimes_{\mathbb{R}} V$, the fact that we are tensoring over $\mathbb{R}$ implies that we must be considering $\mathbb{C}$ as a real linear space.]

With these constructs in mind, We can now define V^c to be a $\mathbb{C}$-linear space.

By the construction of the tensor product we already have an addition on V^c. We have

$$V^c \times V^c \xrightarrow{+} V^c$$

given by

$$\begin{aligned}(\mathbb{C} \otimes_{\mathbb{R}} V) \times (\mathbb{C} \otimes_{\mathbb{R}} V) &\longrightarrow \mathbb{C} \otimes_{\mathbb{R}} V\\ (c_1 \otimes u, c_2 \otimes v) &\longmapsto (c_1 \otimes u) + (c_2 \otimes v)\end{aligned}$$

and this defines an addition in V^c which associates, commutes, has an identity element 0 and has inverses. Next, we need to define a $\mathbb{C}$-scalar multiplication in V^c. The following definition seems to be the obvious one.

$$\begin{aligned}\mathbb{C} \times V^c &\longrightarrow V^c\\ \mathbb{C} \times (\mathbb{C} \otimes_{\mathbb{R}} V) &\longrightarrow \mathbb{C} \otimes_{\mathbb{R}} V\\ (c_1, c \otimes v) &\longmapsto c_1(c \otimes v) := (c_1 c) \otimes v\end{aligned}$$

We check that this is indeed a scalar multiplication. The following properties are straightforward.

$$(c_1 + c_2)(c \otimes v) = ((c_1 + c_2)(c)) \otimes v = (c_1 c + c_2 c) \otimes v = (c_1 c) \otimes v + (c_2 c) \otimes v$$

and

$$(c_1c_2)(c \otimes v) = ((c_1c_2)c) \otimes v = (c_1(c_2c)) \otimes v = c_1((c_2c) \otimes v))$$
$$1(c \otimes v) = (1 \cdot c) \otimes v = c \otimes v$$

Thus we only need to verify that this definition is linear over V^c, i.e.,

$$c(c_1 \otimes u + c_2 \otimes v) = c(c_1 \otimes u) + c(c_2 \otimes v)$$

But there is nothing in our definition that allows us to do this distribution. This is a whole new relationship. However, the righthand side of this expression can be handled by our definition. We have a real basis for the real linear space $\mathbb{C} \otimes_{\mathbf{R}} V$:

$$\{1 \otimes v_1, \cdots, 1 \otimes v_n, i \otimes v_1, \cdots, i \otimes v_n\}$$

and we know that an $\mathbb{R}$-basis for $\mathbb{C}$ considered as a real linear space is $\{1, i\}$. Thus everything on the right hand side of

$$c(c_1 \otimes u + c_2 \otimes v) = c(c_1 \otimes u) + c(c_2 \otimes v)$$

can be expressed as being in a real linear space. Thus we need to verify

$$c(c_1 \otimes u + c_2 \otimes v) = c(c_1 \otimes u) + c(c_2 \otimes v)$$

We begin by examining the expression on the right: $c(c_1 \otimes u) + c(c_2 \otimes v)$. At this point we use the $\mathbb{R}$-basis for $\mathbb{C} \otimes_{\mathbf{R}} V$. We know that an $\mathbb{R}$-basis for $\mathbb{C}$ considered as a real linear space is $\{1, i\}$. We choose a $\mathbb{R}$-basis for V to be $\{v_1, \cdots, v_n\}$. Then a $\mathbb{R}$-basis for $\mathbb{C} \otimes_{\mathbf{R}} V$ is

$$\{1 \otimes v_1, \cdots, 1 \otimes v_n, i \otimes v_1, \cdots, i \otimes v_n\}$$

Using the definition of scalar multiplication, we have

$$c(c_1 \otimes u) + c(c_2 \otimes v) = (cc_1) \otimes u + (cc_2) \otimes v =$$
$$((a+bi)(a_1+b_1i)) \otimes (\textstyle\sum_{k=1}^n(r_kv_k)) + ((a+bi)(a_2+b_2i)) \otimes (\textstyle\sum_{k=1}^n(s_kv_k)) =$$
$$((aa_1-bb_1)+(ab_1+ba_1)i) \otimes (\textstyle\sum_{k=1}^n(r_kv_k))+$$
$$((aa_2-bb_2)+(ab_2+ba_2)i) \otimes (\textstyle\sum_{k=1}^n(s_kv_k)) =$$
$$((aa_1-bb_1) \otimes (\textstyle\sum_{k=1}^n(r_kv_k)) + ((ab_1+ba_1)i) \otimes (\textstyle\sum_{k=1}^n(r_kv_k))+$$
$$((aa_2-bb_2) \otimes (\textstyle\sum_{k=1}^n(s_kv_k)) + ((ab_2+ba_2)i) \otimes (\textstyle\sum_{k=1}^n(s_kv_k)) =$$
$$\textstyle\sum_{k=1}^n((aa_1-bb_1) \otimes (r_kv_k)) + \sum_{k=1}^n(((ab_1+ba_1)i) \otimes (r_kv_k))+$$
$$\textstyle\sum_{k=1}^n((aa_2-bb_2) \otimes (s_kv_k)) + \sum_{k=1}^n(((ab_2+ba_2)i) \otimes (s_kv_k)) =$$
$$\textstyle\sum_{k=1}^n(r_k((aa_1-bb_1)(1 \otimes v_k))) + \sum_{k=1}^n(r_k((ab_1+ba_1)(i \otimes v_k)))+$$
$$\textstyle\sum_{k=1}^n(s_k((aa_2-bb_2)(1 \otimes v_k))) + \sum_{k=1}^n(s_k((ab_2+ba_2)(i \otimes v_k))) =$$
$$\textstyle\sum_{k=1}^n((r_k(aa_1-bb_1)+s_k(aa_2-bb_2))(1 \otimes v_k)+$$
$$(r_k(ab_1+ba_1)+s_k(ab_2+ba_2))(i \otimes v_k))$$

Using the definition of scalar multiplication again, we can write

$$(r_k(ab_1+ba_1)+s_k(ab_2+ba_2))(i\otimes v_k)=$$
$$((r_k(ab_1+ba_1)+s_k(ab_2+ba_2)))(i\cdot 1\otimes v_k)=$$
$$((r_k(ab_1+ba_1)+s_k(ab_2+ba_2)))i)(1\otimes v_k)$$

giving

$$\textstyle\sum_{k=1}^n((r_k(aa_1-bb_1)+s_k(aa_2-bb_2))(1\otimes v_k)+$$
$$(r_k(ab_1+ba_1)+s_k(ab_2+ba_2))(i\otimes v_k))=$$
$$\textstyle\sum_{k=1}^n((r_k(aa_1-bb_1)+s_k(aa_2-bb_2))(1\otimes v_k)+$$
$$((r_k(ab_1+ba_1)+s_k(ab_2+ba_2))i)(1\otimes v_k)=$$
$$\textstyle\sum_{k=1}^n((r_k(aa_1-bb_1)+s_k(aa_2-bb_2))+((r_k(ab_1+ba_1)+s_k(ab_2+ba_2))i)(1\otimes v_k)$$

Changing notation from $\mathbb{C}\otimes_{\mathbb{R}} V$ to V^c, we have

$$(a+bi)((a_1+b_1i)(\textstyle\sum_{k=1}^n(r_kv_k))+(a+bi)((a_2+b_2i)(\sum_{k=1}^n(s_kv_k))=$$
$$\textstyle\sum_{k=1}^n(((r_k(aa_1-bb_1)+s_k(aa_2-bb_2))+(r_k(ab_1+ba_1)+s_k(ab_2+ba_2))i)v_k$$

Returning to the expression

$$c(c_1\otimes u+c_2\otimes v)=c(c_1\otimes u)+c(c_2\otimes v)$$

we now examine the expression of the left. We write $c_1\otimes u+c_2\otimes v$ in terms of the above basis.

$$c_1\otimes u+c_2\otimes v=(a_1+b_1i)\otimes(\textstyle\sum_{k=1}^n(r_kv_k))+(a_2+b_2i)\otimes(\sum_{k=1}^n(s_kv_k))=$$
$$a_1\otimes(\textstyle\sum_{k=1}^n(r_kv_k))+b_1i\otimes(\sum_{k=1}^n(r_kv_k))+$$
$$a_2\otimes(\textstyle\sum_{k=1}^n(s_kv_k))+b_2i\otimes(\sum_{k=1}^n(s_kv_k))=$$
$$\textstyle\sum_{k=1}^n(a_1\otimes(r_kv_k))+\sum_{k=1}^n(b_1i\otimes(r_kv_k))+$$
$$\textstyle\sum_{k=1}^n(a_2\otimes(s_kv_k))+\sum_{k=1}^n(b_2i\otimes(s_kv_k))=$$
$$\textstyle\sum_{k=1}^n r_k(a_1\otimes v_k)+\sum_{k=1}^n r_k(b_1i\otimes v_k)+$$
$$\textstyle\sum_{k=1}^n s_k(a_2\otimes v_k)+\sum_{k=1}^n s_k(b_2i\otimes v_k)=$$
$$\textstyle\sum_{k=1}^n(r_ka_1(1\otimes v_k)+s_ka_2(1\otimes v_k)+r_kb_1(i\otimes v_k)+s_kb_2(i\otimes v_k))=$$
$$\textstyle\sum_{k=1}^n((r_ka_1+s_ka_2)(1\otimes v_k)+(r_kb_1+s_kb_2)(i\otimes v_k))$$

Using the definition of scalar multiplication, we write

$$(r_kb_1+s_kb_2)(i\otimes v_k)=(r_kb_1+s_kb_2)(i\cdot 1\otimes v_k)=((r_kb_1+s_kb_2)i)(1\otimes v_k)$$

Changing notation from $\mathbb{C}\otimes_{\mathbb{R}} V$ to V^c, we now have

$$(a_1+b_1i)(\textstyle\sum_{k=1}^n(r_kv_k))+(a_2+b_2i)(\sum_{k=1}^n(s_kv_k))=$$
$$\textstyle\sum_{k=1}^n((r_ka_1+s_ka_2)+((r_kb_1+s_kb_2)i))v_k$$

What we would like to establish is

$$c(c_1 \otimes u + c_2 \otimes v) = c(c_1 \otimes u) + c(c_2 \otimes v)$$

The right side of this identity we have already calculated directly.

$$c(c_1 \otimes u) + c(c_2 \otimes v) =$$
$$\sum_{k=1}^{n}(((r_k(aa_1 - bb_1) + s_k(aa_2 - bb_2)) + (r_k(ab_1 + ba_1) + s_k(ab_2 + ba_2))i)(v_k)$$

We also have calculated $c_1 \otimes u + c_2 \otimes v$ of the left side. We would like to multiply this by c to get the expression

$$c(c_1 \otimes u + c_2 \otimes v) =$$
$$c(\sum_{k=1}^{n}((r_k a_1 + s_k a_2) + ((r_k b_1 + s_k b_2)i))v_k)$$

However, we have nothing that allows us to move c over the summation sign. What we need is $\mathbb{C}$-linearity with respect to addition in V^c. Let us for the moment suppose that we do have this. Returning to the tensor notation $\mathbb{C} \otimes_{\mathbb{R}} V$, we would have

$$c(\sum_{k=1}^{n}((r_k a_1 + s_k a_2) + ((r_k b_1 + s_k b_2)i))(1 \otimes v_k)) =$$
$$(a + bi)(\sum_{k=1}^{n}((r_k a_1 + s_k a_2) + ((r_k b_1 + s_k b_2)i))(1 \otimes v_k)) =$$
$$\sum_{k=1}^{n}((a + bi)((r_k a_1 + s_k a_2) + ((r_k b_1 + s_k b_2)i)))(1 \otimes v_k)) =$$
$$\sum_{k=1}^{n}((a(r_k a_1 + s_k a_2) - b(r_k b_1 + s_k b_2))+$$
$$(a(r_k b_1 + s_k b_2) + b(r_k a_1 + s_k a_2))i)(1 \otimes v_k)$$

Returning to V^c notation gives

$$c(\sum_{k=1}^{n}((r_k a_1 + s_k a_2) + ((r_k b_1 + s_k b_2)i))v_k) =$$
$$\sum_{k=1}^{n}((a(r_k a_1 + s_k a_2) - b(r_k b_1 + s_k b_2)) + (a(r_k b_1 + s_k b_2) + b(r_k a_1 + s_k a_2))i)v_k$$

We compare the expressions on the left side and the right side and we see that they are equal. Thus we can conclude in V^c that it is "natural" to define

$$c(c_1 \otimes u + c_2 \otimes v) := (cc_1) \otimes u + (cc_2) \otimes v$$

and thus, *with this natural definition*, we make V^c into a $\mathbb{C}$-linear space.

In order to make more intuitive the above construction, let us repeat it for the case where $V = \mathbb{R}$. This means that we want the complexification of $\mathbb{R}$, i.e., $\mathbb{R}^c$, and we show that naturally $\mathbb{R}^c = \mathbb{C} \otimes_{\mathbb{R}} \mathbb{R}$ can be made into a complex linear space $\mathbb{C}$. Fortunately, we can visualize this situation by taking any line in a plane to represent $\mathbb{R}$, and showing that its complexification $\mathbb{C}$ is the plane. Now any line is a one-dimensional real linear space, and the plane is a two- dimensional real linear space, but also it is a one-dimensional complex linear space.

Thus we begin with real linear space $\mathbb{C} \otimes_{\mathbb{R}} \mathbb{R}$. Every element in this space is written as

$$c \otimes k = (a + bi) \otimes k = a \otimes k + bi \otimes k = a(1 \otimes k) + b(i \otimes k)$$

We recognize $(1 \otimes k, i \otimes k)$ as a real basis for $\mathbb{C} \otimes_{\mathbb{R}} \mathbb{R}$, and thus we have $c \otimes k$ written in a 2-dimensional real basis. Interpreting this geometrically on the plane, we see that this says that if we take a plane with a distinguished point O [called the origin of the plane], and any line through O, and any point k on this line, which is represented by $\mathbb{R}$, and multiply this by a, we obtain the point $a \otimes k = ak$ on this line. Then if we rotate this line counterclockwise by 90 degrees around the origin [multiplying by i] and we obtain a point $i \otimes k = ik$ on the line perpendicular to the original line through the origin. Multiplying this by b, we obtain another point $bi \otimes k = bik$ on this line perpendicular to the original line. Adding these two points [vector addition in the plane] gives us a point $c \otimes k$ in the plane.

We remark that all of the above can be interpreted in the notation of V^r. This means that we are interpreting $(\mathbb{R}^c)^r$. If we write $c \otimes k$ as ck, that is, a complex scalar c times a 1-dimensional real vector k, we have

$$ck = (a + ib)k = ak + b(ik)$$

which shows a 1-dimensional complex vector written as a 2-dimensional real vector using the special basis $(1k, ik)$.

Now we have defined a complex scalar multiplication in $\mathbb{C} \otimes_{\mathbb{R}} \mathbb{R}$ as

$$c_1(c \otimes k) := (c_1 c) \otimes k$$

Using this definition, $c \otimes k$ becomes

$$c \otimes k = (a + bi) \otimes k = a(1 \otimes k) + b(i \otimes k) = a(1 \otimes k) + bi(1 \otimes k) = (a + bi)(1 \otimes k) = ck$$

and this says that the point in the plane can also be obtained by multiplying the basis element k by the complex scalar c. Thus we have changed the 2-dimensional real space $\mathbb{C} \otimes_{\mathbb{R}} \mathbb{R}$ into the one-dimensional complex vector space $\mathbb{R}^c$, which is the complexification of $\mathbb{R}$. And, of course, $\mathbb{R}^c = \mathbb{C}$.

We remark that if we take advantage of the natural isomorphism $\mathbb{C} \otimes_{\mathbb{C}} V^c \cong V^c$ so that $c \otimes v$, which maps to cv by scalar multiplication, will be considered just an element u in V^c. Then, in this context there is no need to write a scalar in front of the vector. Also we have this very useful identity where we use the definition of complex scalar multiplication in $\mathbb{C} \otimes_{\mathbb{R}} V$.

$$
\begin{gathered}
V^c = \mathbb{C} \otimes_{\mathbb{R}} V = (\mathbb{R} \oplus \mathbb{R}i) \otimes_{\mathbb{R}} V = \\
(\mathbb{R} \otimes_{\mathbb{R}} V) \oplus ((\mathbb{R}i) \otimes_{\mathbb{R}} V) = \\
V \oplus (i \otimes V) = V \oplus iV
\end{gathered}
$$

which again shows that $V \subset V^c$. This also shows that $\dim_{\mathbb{C}} V^c = \dim_{\mathbb{R}} V$.

We also conclude that we have two notations for the same set: $\mathbb{C} \otimes_{\mathbb{R}} V$ and V^c. The first is a $2n$-dimensional real linear space, while the second is an n-dimensional complex linear space.

2.12.2 From $\mathbb{R}$ to $\mathbb{C}$: The Lie Algebra Structure. Now we let V be a real Lie algebra, e.g., $V = \hat{g}$, where $\hat{g}$ is a real Lie algebra. We form $\hat{g}^c$ and we wish to give it the structure of a complex Lie algebra. Thus, we need to define a bracket in $\hat{g}^c$. We choose:

$$
\begin{gathered}
\hat{g}^c \times \hat{g}^c = (\mathbb{C} \otimes_{\mathbb{R}} \hat{g}) \times (\mathbb{C} \otimes_{\mathbb{R}} \hat{g}) \longrightarrow \mathbb{C} \otimes_{\mathbb{R}} \hat{g} = \hat{g}^c \\
(c_1 \otimes u, c_2 \otimes v) \longmapsto [c_1 \otimes u, c_2 \otimes v] := c_1 c_2 \otimes [u, v]
\end{gathered}
$$

where, of course, $[u, v]$ is the bracket in $\hat{g}$.

With this definition of bracket in $\hat{g}^c$ we show that $\hat{g}^c$ is a Lie algebra over $\mathbb{C}$. First, we note that

$$
\begin{gathered}
[c_1 \otimes u, c_2 \otimes v] = [(a_1 + ib_1) \otimes u, (a_2 + ib_2) \otimes v] = \\
[a_1 \otimes u + (ib_1) \otimes u, a_2 \otimes v + (ib_2) \otimes v] = [a_1 \otimes u + i(b_1 \otimes u), a_2 \otimes v + i(b_2 \otimes v)]
\end{gathered}
$$

Now we let $a_1 \otimes u = x_1$; $b_1 \otimes u = y_1$; $a_2 \otimes v = x_2$; $b_2 \otimes v = y_2$. This gives

$$
[(x_1 + iy_1), (x_2 + iy_2)] = [x_1, x_2] - [y_1, y_2] + i([x_1, y_2] + [y_1, x_2])
$$

Now

$$
\begin{gathered}
c_1 c_2 \otimes [u, v] = (a_1 + ib_1)(a_2 + ib_2) \otimes [u, v] = \\
((a_1 a_2 - b_1 b_2) + i(a_1 b_2 + b_1 a_2)) \otimes [u, v] = \\
(a_1 a_2 - b_1 b_2) \otimes [u, v] + i(a_1 b_2 + b_1 a_2)) \otimes [u, v]
\end{gathered}
$$

Comparing these two expressions, we have

$$
\begin{gathered}
(a_1 a_2) \otimes [u, v] = [a_1 u, a_2 v] = [x_1, x_2] \\
-(b_1 b_2) \otimes [u, v] = -[b_1 u, b_2 v] = -[y_1, y_2] \\
(a_1 b_2) \otimes [u, v] = [a_1 u, b_2 v] = [x_1, y_2] \\
(b_1 a_2) \otimes [u, v] = [b_1 u, a_2 v] = [y_1, x_2]
\end{gathered}
$$

Thus we see that by just using $\mathbb{C}$-linearity in $\hat{g}^c$ we have a natural identity of a Lie bracket in $\hat{g}^c$.

It is straightforward that scalar multiplication is bilinear with respect to this multiplication. For c in $\mathbb{C}$ and $c_1 \otimes u$ and $c_2 \otimes v$ in $(\mathbb{C} \otimes_{\mathbb{R}} \hat{g})$, we have

$$c[c_1 \otimes u, c_2 \otimes v] = c(c_1c_2 \otimes [u,v]) = (c(c_1c_2)) \otimes [u,v] =$$
$$((cc_1)c_2) \otimes [u,v] = [(cc_1) \otimes u, c_2 \otimes v] = [c(c_1 \otimes u), c_2 \otimes v]$$

and

$$c[c_1 \otimes u, c_2 \otimes v] = c(c_1c_2 \otimes [u,v]) = (c(c_1c_2)) \otimes [u,v] =$$
$$(c_1(cc_2)) \otimes [u,v] = [c_1 \otimes u, (cc_2) \otimes v] = [c_1 \otimes u, c(c_2 \otimes v)]$$

We also need to show that this multiplication distributes on the right and on the left, that is, it is bilinear with respect to addition.

$$[c_1 \otimes u, (c_2 \otimes v + c_3 \otimes w)] = [c_1 \otimes u, c_2 \otimes v] + [c_1 \otimes u), c_3 \otimes w]$$
$$[c_1 \otimes u + c_2 \otimes v, c_3 \otimes w] = [c_1 \otimes u), c_3 \otimes w] + [c_2 \otimes v, c_3 \otimes w]$$

For left distribution we want to show

$$[c_1 \otimes u, (c_2 \otimes v + c_3 \otimes w)] = [c_1 \otimes u, c_2 \otimes v] + [c_1 \otimes u, c_3 \otimes w]$$

We reduce first the righthand side of this equation. We work with a basis $(v_1, ..., v_n)$ in $\hat{g}$.

$$[c_1 \otimes u, c_2 \otimes v] + [c_1 \otimes u, c_3 \otimes w] = c_1c_2 \otimes [u,v] + c_1c_3 \otimes [u,w] =$$
$$(a_1 + b_1i)(a_2 + b_2i) \otimes [\textstyle\sum_{i=1}^n (r_iv_i), \sum_{j=1}^n (s_jv_j)] +$$
$$(a_1 + b_1i)(a_3 + b_3i) \otimes [\textstyle\sum_{i=1}^n (r_iv_i), \sum_{k=1}^n (t_kv_k)] =$$
$$((a_1a_2 - b_1b_2) + (a_1b_2 + b_1a_2)i) \otimes \textstyle\sum_{i=1}^n \sum_{j=1}^n r_is_j[v_i, v_j] +$$
$$((a_1a_3 - b_1b_3) + (a_1b_3 + b_1a_3)i) \otimes \textstyle\sum_{i=1}^n \sum_{k=1}^n r_it_k[v_i, v_k] =$$
$$\textstyle\sum_{i=1}^n \sum_{j=1}^n ((a_1a_2 - b_1b_2) + (a_1b_2 + b_1a_2)i) \otimes r_is_j[v_i, v_j] +$$
$$\textstyle\sum_{i=1}^n \sum_{k=1}^n ((a_1a_3 - b_1b_3) + (a_1b_3 + b_1a_3)i) \otimes r_it_k[v_i, v_k]$$

Handling the lefthand side of the equation is more delicate. We work first with $c_2 \otimes v + c_3 \otimes w$.

$$c_2 \otimes v + c_3 \otimes w = (a_2 + b_2i) \otimes \textstyle\sum_{j=1}^n (s_jv_j) + (a_3 + b_3i) \otimes \sum_{k=1}^n (t_kv_k) =$$
$$\textstyle\sum_{j=1}^n ((a_2 + b_2i) \otimes (s_jv_j)) + \sum_{k=1}^n ((a_3 + b_3i) \otimes (t_kv_k))$$

Now we have

$$[c_1 \otimes u, (c_2 \otimes v + c_3 \otimes w)] =$$
$$[c_1 \otimes u, \textstyle\sum_{j=1}^n ((a_2 + b_2i) \otimes (s_jv_j)) + \sum_{k=1}^n ((a_3 + b_3i) \otimes (t_kv_k))]$$

Once again we see that we have nothing that makes legitimate moving brackets across addition in $\hat{g}^c = (\mathbb{C} \otimes_{\mathbb{R}} \hat{g})$. Again, let us assume, for the moment, that we can. This gives

$$[c_1 \otimes u, \sum_{j=1}^n ((a_2 + b_2 i) \otimes (s_j v_j)) + \sum_{k=1}^n ((a_3 + b_3 i) \otimes (t_k v_k))] =$$
$$[c_1 \otimes u, \sum_{j=1}^n ((a_2 + b_2 i) \otimes (s_j v_j))] + [c_1 \otimes u, \sum_{k=1}^n ((a_3 + b_3 i) \otimes (t_k v_k))] =$$
$$\sum_{j=1}^n [c_1 \otimes u, ((a_2 + b_2 i) \otimes (s_j v_j))] + \sum_{k=1}^n [c_1 \otimes u, ((a_3 + b_3 i) \otimes (t_k v_k))] =$$

Now we can apply the definition of the Lie bracket in $\hat{g}^c = (\mathbb{C} \otimes_{\mathbb{R}} \hat{g})$.

$$\sum_{j=1}^n [c_1 \otimes u, ((a_2 + b_2 i) \otimes (s_j v_j))] + \sum_{k=1}^n [c_1 \otimes u, ((a_3 + b_3 i) \otimes (t_k v_k))] =$$
$$\sum_{j=1}^n c_1(a_2 + b_2 i) \otimes [u, s_j v_j] + \sum_{k=1}^n c_1(a_3 + b_3 i) \otimes [u, t_k v_k]$$

We now expand c_1 and u.

$$\sum_{j=1}^n c_1(a_2 + b_2 i) \otimes [u, s_j v_j] + \sum_{k=1}^n c_1(a_3 + b_3 i) \otimes [u, t_k v_k] =$$
$$\sum_{j=1}^n (a_1 + b_1 i)(a_2 + b_2 i) \otimes [\sum_{i=1}^n (r_i v_i), s_j v_j] +$$
$$\sum_{k=1}^n (a_1 + b_1 i)(a_3 + b_3 i) \otimes [\sum_{i=1}^n (r_i v_i), t_k v_k]$$

Now since we know that brackets in $\hat{g}$ are bilinear with respect to addition and real scalars, we have

$$\sum_{j=1}^n (a_1 + b_1 i)(a_2 + b_2 i) \otimes [\sum_{i=1}^n (r_i v_i), s_j v_j] +$$
$$\sum_{k=1}^n (a_1 + b_1 i)(a_3 + b_3 i) \otimes [\sum_{i=1}^n (r_i v_i), t_k v_k] =$$
$$\sum_{i=1}^n \sum_{j=1}^n (a_1 a_2 - b_1 b_2 + (a_1 b_2 + b_1 a_2) i) \otimes r_i s_j [v_i, v_j] +$$
$$\sum_{i=1}^n \sum_{k=1}^n (a_1 a_3 - b_1 b_3 + (a_1 b_3 + b_1 a_3) i) \otimes r_i t_k [v_i, v_k]$$

Now we wanted to show that

$$[c_1 \otimes u, (c_2 \otimes v + c_3 \otimes w)] = [c_1 \otimes u, c_2 \otimes v] + [c_1 \otimes u, c_3 \otimes w]$$

We calculated

$$[c_1 \otimes u, c_2 \otimes v] + [c_1 \otimes u, c_3 \otimes w] =$$
$$\sum_{i=1}^n \sum_{j=1}^n ((a_1 a_2 - b_1 b_2) + (a_1 b_2 + b_1 a_2) i) \otimes r_i s_j [v_i, v_j] +$$
$$\sum_{i=1}^n \sum_{k=1}^n ((a_1 a_3 - b_1 b_3) + (a_1 b_3 + b_1 a_3) i) \otimes r_i t_k [v_i, v_k]$$

Assuming that we can move brackets across addition in $\hat{g}^c = (\mathbb{C} \otimes_{\mathbb{R}} \hat{g})$, we obtained

$$[c_1 \otimes u, (c_2 \otimes v + c_3 \otimes w)] =$$
$$\sum_{j=1}^n (a_1 + b_1 i)(a_2 + b_2 i) \otimes [\sum_{i=1}^n (r_i v_i), s_j v_j] +$$
$$\sum_{k=1}^n (a_1 + b_1 i)(a_3 + b_3 i) \otimes [\sum_{i=1}^n (r_i v_i), t_k v_k] =$$
$$\sum_{i=1}^n \sum_{j=1}^n (a_1 a_2 - b_1 b_2 + (a_1 b_2 + b_1 a_2) i) \otimes r_i s_j [v_i, v_j] +$$
$$\sum_{i=1}^n \sum_{k=1}^n (a_1 a_3 - b_1 b_3 + (a_1 b_3 + b_1 a_3) i) \otimes r_i t_k [v_i, v_k]$$

We see that these two expressions are identical. Thus it is reasonable *to define* that brackets are linear with respect to addition in $\hat{g}^c = (\mathbb{C} \otimes_{\mathbb{R}} \hat{g})$. And obviously the same conclusion is true for right distribution.

We also need to show the anticommutativity of the bracket product in $\hat{g}^c$. It comes from this property holding in $\hat{g}$:

$$[c_1 \otimes u, c_2 \otimes v] = c_1c_2 \otimes [u, v] = -c_2c_1 \otimes [v, u] = -[c_2 \otimes v, c_1 \otimes u]$$

Finally we need to show that the Jacobi identity holds in $\hat{g}^c$.

$$\begin{gathered}
[c_1 \otimes u, [c_2 \otimes v, c_3 \otimes w]]+ \\
[c_3 \otimes w, [c_1 \otimes u, c_2 \otimes v]]+ \\
[c_2 \otimes v, [c_3 \otimes w, c_1 \otimes u]] = \\
[c_1 \otimes u, (c_2c_3) \otimes [v, w]] + [c_3 \otimes w, (c_1c_2) \otimes [u, v]] + [c_2 \otimes v, (c_3c_1) \otimes [w, u]] = \\
(c_1c_2c_3) \otimes [u, [v, w]] + (c_3c_1c_2) \otimes [w, [u, v]] + (c_2c_3c_1) \otimes [v, [w, u]] = \\
(c_1c_2c_3) \otimes ([u, [v, w]] + [w, [u, v]] + [v, [w, u]]) = 0
\end{gathered}$$

since the Jacobi identity is valid in $\hat{g}$.

Thus when we build the complexification of the $\mathbb{C}$-linear Lie algebra $\hat{g}^c$ from the structure of the $\mathbb{R}$-linear Lie algebra $\hat{g}$, we have a canonical way of obtaining it. If the complex dimension of $\hat{g}^c$ is n, then we have immediately the $2n$-real dimensional Lie algebra $(\hat{g}^c)^r = \hat{g} \times \hat{g}$, since we have already identified $\hat{g}$, i.e., we do not need to choose a basis in $\hat{g}^c$ to identify $\hat{g}$ in $\hat{g}^c$. Also, from the following calculation, we see that for any decomposition of $\hat{g}^c = \hat{g} \oplus i\hat{g}$ such that $\hat{g} \oplus i0 \subset \hat{g} \oplus i\hat{g}$, we can conclude that $\hat{g}$ is a real Lie subalgebra of real dimension n of $(\hat{g}^c)^r$, which latter algebra has real dimension $2n$. We have for $u = u_{re1} + i(0)$ and $v = v_{re1} + i(0)$ in $\hat{g} \oplus i\hat{g}$

$$\left[\begin{bmatrix} [u_{re1}] \\ [0] \end{bmatrix}, \begin{bmatrix} [v_{re1}] \\ [0] \end{bmatrix}\right] = \begin{bmatrix} [u_{re1}, v_{re1}] - [0, 0] \\ [u_{re1}, 0] + [0, v_{re1}] \end{bmatrix} = \begin{bmatrix} [u_{re1}, v_{re1}] \\ 0 + 0 \end{bmatrix}$$

This also says that $[u, v] = [u_{re1} + i(0), v_{re1} + i(0)] = [u_{re1}, v_{re1}] + i(0)$ in $\hat{g} \oplus i\hat{g}$.

2.12.3 From $\mathbb{C}$ to $\mathbb{R}$: the Linear Space Structure. We now want to move from a $\mathbb{C}$-linear space to an $\mathbb{R}$-linear space. Given a complex linear space, we ask what kind of real linear structures does it determine? This means that the given complex linear space V is regarded as coming from the complexification of some real linear space V_{U_e}, i.e., $V = (V_{U_e})^c$, where V_{U_e} is given a basis $\{e_1, ..., e_n\}$. [CAUTION: the basis $\{e_1, ..., e_n\}$ is in this section <u>not</u> the basis of the canonical vectors $\{(1, 0, \cdots, 0), \cdots, (0, \cdots, 0, 1)\}$, but an arbitrary basis in V_{U_e}.]

It seems evident that, given a complex linear space V, there is no unique V_{U_e} whose complexification is the given complex linear space V. Thus we start with a complex linear space V and chose a V_{U_e} with basis $\{e_1, ..., e_n\}$ such that $V = (V_{U_e})^c$.

This means that any element in x in V_{U_e} can be written as $x = a_1e_1 + \cdot\cdot \cdot + a_ne_n$. Now the complexifcation $V = (V_{U_s})^c$ came from putting a complex structure on the real linear space $\mathbb{C}^r \otimes_{\mathbb{R}} V_{U_e}$. When one starts with a real linear space $\mathbb{C}^r \otimes_{\mathbb{R}} V_{U_e}$, then all of the above analysis is seen to depend on the following definition for making this space into a $\mathbb{C}$-linear space:

$$c_1(c \otimes v) := (c_1c) \otimes v$$

In particular if $c_1 = 1$, we have $1(c \otimes v) = (1c) \otimes v = c \otimes v$; and if $c_1 = i$, we have $i(c \otimes v) = (ic) \otimes v$.

Now since $(V_{U_e})^c$ is the same set as $\mathbb{C} \otimes_{\mathbb{R}} V_{U_e}$, we can form a basis for $(V_{U_e})^c$ over $\mathbb{C}$ to be $\{1 \otimes e_1, ..., 1 \otimes e_n\}$, giving $\dim_{\mathbb{C}}(V_{U_e})^c = n$. We now take the set of linear combinations of this basis and note that

$$\begin{gathered} a_1(1 \otimes e_1) + \cdots + a_n(1 \otimes e_n) = \\ 1 \otimes a_1e_1 + \cdots + 1 \otimes a_ne_n = \\ 1 \otimes (a_1e_1 + \cdots + a_ne_n) \end{gathered}$$

with a_i in $\mathbb{R}$, and this coincides with the subset $\{1 \otimes x | x \in V_U\}$ of $(V_{U_e})^c$. Thus this set is a $\mathbb{R}$-subspace of $(V_{U_e})^c$ and is isomorphic to V_{U_e}. Thus V_{U_e} becomes a $\mathbb{R}$-subspace, and it has the following two properties:

(1) The $\mathbb{C}$-space spanned by V_{U_e} is $(V_{U_e})^c$
(2) Any subset of V_{U_e} which is linearly independent over $\mathbb{R}$ is linearly independent over $\mathbb{C}$

The proofs follow:

(1) is obvious. Any v in $(V_{U_e})^c$ can be written as

$$\begin{gathered} v = c_1(1 \otimes e_1) + \cdots + c_n(1 \otimes e_n) \text{ with } c_i \text{ in } \mathbb{C} = \\ c_1 \cdot 1 \otimes e_1 + \cdots + c_n \cdot 1 \otimes e_n = \\ c_1 \otimes e_1 + \cdots + c_n \otimes e_n \end{gathered}$$

(2) Let W be a subspace of V and let $\{f_1, ..., f_k\}$ be an $\mathbb{R}$ basis for W. For x in W and a_i in $\mathbb{R}$,

if $x = a_1f_1 + \cdots + a_kf_k = 0$, then $a_1 = \cdots = a_k = 0$. Now we take the corresponding $\mathbb{C}$ basis $\{1 \otimes f_1, ..., 1 \otimes f_k\}$ and form

$$\begin{gathered} z = c_1(1 \otimes f_1) + \cdots + c_k(1 \otimes f_k) = 0 \\ z = (a_1 + ib_1)(1 \otimes f_1) + \cdots + (a_k + ib_k)(1 \otimes f_k) = 0 \\ \text{then } z = (a_1(1 \otimes f_1) + \cdots + a_k(1 \otimes f_k)) + \\ i(b_1(1 \otimes f_1) + \cdots + b_k(1 \otimes f_k)) = 0 \end{gathered}$$

$$z = ((1 \otimes a_1 f_1) + \cdots + (1 \otimes a_k f_k)) +$$
$$i((1 \otimes b_1 f_1) + \cdots + (1 \otimes b_k f_k)) = 0$$
$$z = 1 \otimes ((a_1 f_1 + \cdots + a_k f_k) + i(b_1 f_1 + \cdots + b_k f_k)) = 0$$
$$z = 1 \otimes ((a_1 + ib_1) f_1 + \cdots + (a_k + ib_k) f_k) = 0$$
$$z = 1 \otimes ((a_1 f_1 + \cdots + a_k f_k) + i(b_1 f_1 + \cdots + b_k f_k)) = 0$$

Thus $(a_1 f_1 + \cdots + a_k f_k) = 0$ and $(b_1 f_1 + \cdots + b_k f_k) = 0$;
and we conclude that $a_1 = \cdots = a_k = 0$, and $b_1 = \cdots = b_k = 0$,
giving $c_1 = \cdots = c_k = 0$. Thus we conclude that W^c,
a $\mathbb{C}$-subspace of $(V_{U_e})^c$, is k-dimensional over $\mathbb{C}$.

Thus we start with an n-dimensional real linear space V_{U_e} with basis $\{e_1, ..., e_n\}$. From the above we know that the complex linear space $(V_{U_e})^c$ has a basis $\{1 \otimes e_1, ...1 \otimes e_n\}$ over $\mathbb{C}$ and therefore any vector v in $(V_{U_e})^c$ can be written as:

$$v = c_1(1 \otimes e_1) + ... + c_n(1 \otimes e_n) = (a_1 + ib_1)(1 \otimes e_1) + ... + (a_n + ib_n)(1 \otimes e_n) =$$
$$(a_1)(1 \otimes e_1) + ... + (a_n)(1 \otimes e_n) + i(b_1(1 \otimes e_1) + ... + b_n(1 \otimes e_n))$$

In this combination the c_i's are complex numbers and the a_1's and b_i's are real numbers. However if we identify $(1 \otimes f_i)$ with f_i, we can write the combination as:

$$v = (a_1 + ib_1) f_1 + ... + (a_n + ib_n) f_n =$$
$$a_1 f_1 + ... + a_n f_n + i(b_1 f_1 + ... + b_n f_n) =$$
$$a_1 f_1 + \cdots + a_n f_n + b_1(if_1) + \cdots + b_(if_n)$$

This shows that $V_{U_e} + iV_{U_e}$ can be written as a $2n$-dimensional <u>real</u> linear space on a special basis $\{e_1, ..., e_n, ie_1, ..., ie_n\}$. We call this real linear space V_e^r. Thus we have

$$V_e^r = V_{U_e} + iV_{U_e}$$

Given a linear space V over $\mathbb{C}$, we have complex scalar multiplication

$$\mathbb{C} \times V \longrightarrow V$$
$$(c_k, v) \longmapsto c_k v.$$

where the c_k's are complex numbers. Also we can write these complex numbers as $c_k = a_k + b_k i$, where the a_k'sand the b_k's are real numbers. This gives

$$v = c_1 e_1 + \cdots + c_n e_n$$
$$v = (a_1 + b_1 i) e_1 + \cdots + (a_n + b_n i) e_n =$$
$$a_1 e_1 + \cdots + a_n e_n + b_1(ie_1) + \cdots + b_n(ie_n)$$

This says that we can give same <u>set</u> V a basis $\{e_1, \cdots, e_n, ie_1, \cdots, ie_n\}$ over the real numbers, making the set V into a real linear space V_e^r with real dimension twice the complex dimension of V. Thus we know that if V is a complex linear space of complex dimension n, then V_e^r is a real linear space of real dimension $2n$.

However we observe that the real basis given above is not completely arbitrary. We did arbitrarily choose n vectors linearly independent over the reals, but the other n vectors linearly independent over the reals were not arbitrary but were i times the original set of n vectors linearly independent over the reals. Thus we have a very special kind of basis for V_e^r. In fact, the v in V_e^r can now be written as

$$v = (a_1 + b_1 i)e_1 + \cdots + (a_n + b_n i)e_n =$$
$$a_1 e_1 + \cdots + a_n e_n + b_1(ie_1) + \cdots + b_n(ie_n) =$$
$$a_1 e_1 + \cdots + a_n e_n + i(b_1 e_1 + \cdots + b_n e_n)$$

and thus we see that $V_e^r = V_{U_e} \oplus iV_{U_e}$, where V_{U_e} is a real subspace of dimension n of V_e^r.

We thus have a map from the complexification of V_{U_e} to V:

$$(V_{U_e})^c \longrightarrow V$$
$$V_{U_e} \oplus iV_{U_e} \longrightarrow V$$
$$(a_1 e_1 + \cdots + a_n e_n) + i(b_1 e_1 + \cdots + b_e e_n) \longmapsto (a_1 + b_1 i)e_1 + \cdots + (a_n + b_n i)e_n$$

which is obviously an isomorphism of complex linear spaces.

This phenomenon of complexification shows up in some surprising ways. Suppose now that we chose another complexification V_{U_u} of V where V_{U_u} is given a basis $\{u_1, ..., u_n\}$, and suppose that V_{U_u} is not equal V_{U_e} This means that the given complex linear space V comes from the complexification of some real linear space V_{U_u}, i.e., $V = (V_{U_u})^c$, where V_{U_u} is given a basis $\{u_1, ..., u_n\}$. And with respect to the u-basis $V_u^r = V_{U_u} \times V_{U_u}$, that is, V_u^r is represented by two independent copies of V_{U_u}, giving the dimension of V_u^r as a real linear space to be $2n$. However the information communicated by writing V_u^r as such is very much linked together. To see this we do the following. We go to matrix notation. Having chosen a e-basis for V, then each element v in V is written as a column vector $[c_k] = [a_k + ib_k] = [a_k] + i[b_k]$, and this says that the corresponding element in V_e^r can be written as a $2n$ column vector of real numbers.

$$\begin{bmatrix} [a_k] \\ [b_k] \end{bmatrix}$$

We now choose another basis for V: $\{u_1, \cdots, u_n\}$, and write an arbitrary element v in V in this basis:

$$v = d_1u_1 + \cdots + d_nu_n$$

where the d_k's are complex numbers. Now we can write these complex numbers as $d_k = r_k + s_ki$, where the r_k's and the s_k's are real numbers. This gives

$$\begin{gathered} v = (r_1 + s_1i)u_1 + \cdots + (r_n + s_ni)u_n = \\ r_1u_1 + \cdots + r_nu_n + s_1(iu_1) + \cdots + s_n(iu_n) = \\ r_1u_1 + \cdots + r_nu_n + i(s_1u_1 + \cdots + s_nu_n) \end{gathered}$$

Now we let V_{U_u} be the real linear space which is generated by the real basis $\{u_1, \cdots, u_n\}$. Maintaining our choices as above, we have that every element v in V can be expressed also as $v = u_{re1} + iu_{re2}$, and thus V is expressed as $V = V_{U_u} \oplus iV_{U_u}$, and $V_u^r = V_{U_u} \times V_{U_u}$ is just two copies of V_{U_u}. Thus each element of v in V is written as a column vector $[d_k] = [r_k + is_k] = [r_k] + i[s_k]$, and the same element in V_u^r is written as a column vector

$$\begin{bmatrix} [r_k] \\ [s_k] \end{bmatrix}$$

Now the matrix $C = [c_{jk}]$, which changes the e-basis to the u-basis, is a non-singular matrix in $GL(n,\mathbb{C})$, that has n^2 complex entries. Also C can be written as $C = A + iB$, where $A = [a_{jk}]$ and $B = [b_{jk}]$ are in $M_{nxn}(\mathbb{R})$. This now gives us $2n^2$ real entries. We seek the $\mathbb{R}$-matrix K which maps the basis $\{e_1, \cdots, e_n, ie_1, \cdots, ie_n\}$ to the basis $\{u_1, \cdots, u_n, iu_1, \cdots, iu_n\}$. We observe that this matrix must be $2n$x$2n$ and thus it must have $(2n)^2 = 4n^2$ real entries. Since we only have $2n^2$ real entries, this means that the K matrix is not arbitrary, but has a very special form. Of course, this is due to the fact that the second parts of the bases $\{ie_1, \cdots, ie_n\}$ and $\{iu_1, \cdots, iu_n\}$ are not independent of the first parts, but are related to the first parts in a very special way — through multiplication by i.

To obtain this matrix K from the matrix $C = A+iB$, we do the following. We write the vector v in V as an n-column matrix of complex entries $[c_k]$ with respect to the basis $\{e_k\}$. Since $c_k = a_k + b_ki$, this matrix can be written as $[c_k] = [a_k] + i[b_k]$, where $[a_k]$ and $[b_k]$ are n-column real matrices written with respect to the same basis $\{e_k\}$. We do the same with v written with respect to the basis $\{u_k\}$, giving $[d_k] = [r_k] + i[s_k]$, where $[r_k]$ and $[s_k]$ are n-column real matrices written with respect to the same basis $\{u_k\}$. Using matrix notation we get

$$
\begin{gathered}
[d_k] = C[c_k] \\
[r_k] + i[s_k] = (A + iB)([a_k] + i[b_k]) = \\
A[a_k] + A(i[b_k]) + iB[a_k] + iB(i[b_k]) = \\
A[a_k] - B[b_k] + i(A[b_k] + B[a_k])
\end{gathered}
$$

This matrix equation can be rewritten in the following real notation

$$
\begin{bmatrix} A & -B \\ B & A \end{bmatrix} \begin{bmatrix} [a_k] \\ [b_k] \end{bmatrix} = \begin{bmatrix} A[a_k] - B[b_k] \\ B[a_k] + A[b_k] \end{bmatrix} = \begin{bmatrix} [r_k] \\ [s_k] \end{bmatrix}
$$

which shows a $(2n)^2$ real matrix operating on a $2n$-column real matrix, giving a $2n$-column real matrix. We therefore define K to be

$$
\begin{bmatrix} A & -B \\ B & A \end{bmatrix}
$$

This development shows how an n^2 complex matrix C, with $2n^2$ pieces of real information gives a $(2n)^2$ real matrix K with the same amount of real information. The special form of K is most important. One $2n$-column real matrix is written with respect to two copies of the $\{e_k\}$-basis, while the other $2n$-column real matrix is written with respect to two copies of the $\{u_k\}$-basis. Essentially we have shown how to take any $\mathbb{C}$-automorphism ϕ of V and express it as a <u>special</u> $\mathbb{R}$-automorphism ψ of V_e^r to V_u^r.

But we can say more about this structure. We know that we can compose two endomorphisms of V, ϕ_1, represented in some given basis of V by the matrix C_1, and ϕ_2 represented in the same basis of V by the matrix C_2, giving the endomorphism $\phi_2 \circ \phi_1$, represented by the matrix product C_2C_1 with respect to the same basis. We have

$$
\begin{gathered}
C_2C_1 = (A_2 + iB_2)(A_1 + iB_1) = A_2A_1 + A_2(iB_1) + iB_2A_1 + iB_2(iB_1) = \\
A_2A_1 - B_2B_1 + i(A_2B_1 + B_2A_1)
\end{gathered}
$$

Thus C_2C_1 gives the real matrix

$$
\begin{bmatrix} A_2A_1 - B_2B_1 & -(A_2B_1 + B_2A_1) \\ A_2B_1 + B_2A_1 & A_2A_1 - B_2B_1 \end{bmatrix}
$$

while by multiplying the corresponding matrices, we obtain

$$
\begin{bmatrix} A_2 & -B_2 \\ B_2 & A_2 \end{bmatrix} \begin{bmatrix} A_1 & -B_1 \\ B_1 & A_1 \end{bmatrix} = \begin{bmatrix} A_2A_1 - B_2B_1 & -A_2B_1 - B_2A_1 \\ B_2A_1 + A_2B_1 & -B_2B_1 + A_2A_1 \end{bmatrix}
$$

We observe that we arrive at the same matrix, and thus we can conclude multiplication is preserved under the above correspondence.

Using the methods that gave us this conclusion, we would like see how the inverse of a matrix corresponds. Thus we choose C^{-1} so that $C \cdot C^{-1} = I_n$, where I_n is the complex n-dimensional identity matrix. Since $I_n = I_{n;real} + i0_n$, the real $2n$-matrix corresponding to I_n is

$$\begin{bmatrix} I_{n;real} & 0_n \\ 0_n & I_{n;real} \end{bmatrix}$$

which we see is the $2n$-real identity matrix I_{2n}. Now writing $C \cdot C^{-1}$ as

$$C \cdot C^{-1} = (A + iB)(A' + iB') = (AA' - BB') + i(AB' + BA')$$

gives the $2n$-real matrix

$$\begin{bmatrix} AA' - BB' & -(AB' + BA') \\ AB' + BA' & AA' - BB' \end{bmatrix} = \begin{bmatrix} I_n & 0_n \\ 0_n & I_n \end{bmatrix}$$

Therefore we can say that

$$\begin{bmatrix} A & -B \\ B & A \end{bmatrix}^{-1} = \begin{bmatrix} A' & -B' \\ B' & A' \end{bmatrix}$$

We also observe that $AA' - BB' = I_n$ and $AB' = -BA'$. Thus caution must be observed when treating inverses in this construction.

Finally, another caution should be mentioned, namely, one concerning the taking of transposes. If X is a mxn matrix, we define the transpose X^t of X to be the nxm matrix whose rows are the columns of the original matrix X, and thus whose columns are the rows of the matrix X. We know that

$$\text{if } C = A + iB \qquad \text{then} \qquad C^t = (A + iB)^t = A^t + iB^t$$

Thus if the real matrix representing C is

$$\begin{bmatrix} A & -B \\ B & A \end{bmatrix}$$

then the real matrix representing C^t is

$$\begin{bmatrix} A^t & -B^t \\ B^t & A^t \end{bmatrix} \neq \left(\begin{bmatrix} A & -B \\ B & A \end{bmatrix} \right)^t = \begin{bmatrix} A^t & B^t \\ -B^t & A^t \end{bmatrix}$$

Thus the operation of taking transposes is not preserved in this construction.

2.12.4 The J Transformation. Now we said above that V^r has "more" structure than just $V^r = V_{re_e} \times V_{re_e}$, since it comes from V, a complex linear space, by a choice of a $\mathbb{C}$-basis in V. But since our concern now will not be the specific basis chosen, we will simplify our notation to $V^r = V_{re} \times V_{re}$ to indicate that some basis has been chosen. Using this notation, we can now identify what this "more" structure is. We can get to this additional structure by asking how can we take the real vector space $V^r = V_{re} \times V_{re}$ and make it into a complex vector space which is isomorphic to V. Giving V_{re} a real basis, we can write any element in $V_{re} \times V_{re}$ as a column matrix

$$\begin{bmatrix} [a_k] \\ [b_k] \end{bmatrix}$$

Recalling our construction above, we want to operate on what corresponds in $V_{re} \times V_{re}$ to the imaginary component of V by a transformation C which corresponds to the imaginary transformation. Now $0 \times V_{re}$ in $V_{re} \times V_{re}$ corresponds to the imaginary component of V, and the transformation corresponding to the imaginary transformation is $0 + iI_n$. In matrix notation we get

$$\begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix} \begin{bmatrix} [0] \\ [b_k] \end{bmatrix} = \begin{bmatrix} [-b_k] \\ [0] \end{bmatrix}$$

If we let the same transformation operate on what corresponds to the real component of V, we get

$$\begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix} \begin{bmatrix} [a_k] \\ [0] \end{bmatrix} = \begin{bmatrix} [0] \\ [a_k] \end{bmatrix}$$

We now show that if $V_{re} \times V_{re}$ possesses a linear transformation, call it J, defined by

$$\begin{aligned} J : V_{re} \times V_{re} &\longrightarrow V_{re} \times V_{re} \\ (v_1, v_2) &\longmapsto J(v_1, v_2) := (-v_2, v_1) \end{aligned}$$

then we can define a complex linear space structure on $V_{re} \times V_{re}$. In matrix notation we have

$$J = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}$$

We remark that $J^2 = -I_{2n}$ since:

$$J \cdot J = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix} = - \begin{bmatrix} I_n & 0 \\ 0 & I_n \end{bmatrix} = -I_{2n}$$

The additive structure is simply the additive structure in $V_{re} \times V_{re}$, i.e.:

$$\begin{gathered} (V_{re} \times V_{re}) \times (V_{re} \times V_{re}) \longrightarrow V_{re} \times V_{re} \\ ((v_1, v_2), (w_1, w_2)) \longmapsto (v_1, v_2) + (w_1, w_2) := (v_1 + w_1, v_2 + w_2) \end{gathered}$$

But the scalar multiplication is defined using the J transformation.

$$\begin{gathered} \mathbb{C} \times (V_{re} \times V_{re}) \longrightarrow V_{re} \times V_{re} \\ (c, (v_1, v_2)) \longmapsto c(v_1, v_2) = (a + ib)(v_1, v_2) := (aI_{2n} + bJ)((v_1, v_2)) = \\ a(v_1, v_2) + b(J(v_1, v_2)) = (av_1, av_2) + (-bv_2, bv_1) = (av_1 - bv_2, av_2 + bv_1) \end{gathered}$$

We verify that the expected properties of scalar multiplication hold.

$$\begin{gathered} c((v_1, v_2) + (w_1, w_2)) = (a + ib)(v_1 + w_1, v_2 + w_2) = \\ (a(v_1 + w_1) - b(v_2 + w_2), a(v_2 + w_2) + b(v_1 + w_1)) = \\ (av_1 + aw_1 - bv_2 - bw_2, av_2 + aw_2 + bv_1 + bw_1) = \\ ((av_1 - bv_2, av_2 + bv_1) + ((aw_1 - bw_2, aw_2 + bw_1) = \\ c(v_1, v_2) + c(w_1, w_2) \end{gathered}$$

$$\begin{gathered} (c_1 + c_2)(v_1, v_2) = ((a_1 + ib_1) + (a_2 + ib_2))(v_1, v_2) = \\ ((a_1 + a_2) + i(b_1 + b_2))(v_1, v_2) = \\ ((a_1 + a_2)v_1 - (b_1 + b_2)v_2, (a_1 + a_2)v_2 + (b_1 + b_2)v_1) = \\ (a_1v_1 + a_2v_1 - b_1v_2 - b_2v_2, a_1v_2 + a_2v_2 + b_1v_1 + b_2v_1) = \\ (a_1v_1 - b_1v_2, a_1v_2 + b_1v_1) + (a_2v_1 - b_2v_2, a_2v_2 + b_2v_1) = \\ c_1(v_1, v_2) + c_2(v_1, v_2) \end{gathered}$$

$$\begin{gathered} (c_1c_2)(v_1, v_2) = ((a_1 + ib_1)(a_2 + ib_2))(v_1, v_2) = \\ ((a_1a_2 - b_1b_2) + i(a_1b_2 + b_1a_2))(v_1, v_2) = \\ ((a_1a_2 - b_1b_2)v_1 - (a_1b_2 + b_1a_2)v_2, (a_1a_2 - b_1b_2)v_2 + (a_1b_2 + b_1a_2)v_1) = \\ (a_1a_2v_1 - b_1b_2v_1 - a_1b_2v_2 - b_1a_2v_2, a_1a_2v_2 - b_1b_2v_2 + a_1b_2v_1 + b_1a_2v_1) = \\ (a_1a_2v_1 - a_1b_2v_2 - b_1a_2v_2 - b_1b_2v_1, a_1a_2v_2 + a_1b_2v_1 + b_1a_2v_1 - b_1b_2v_2) = \\ (a_1(a_2v_1 - b_2v_2) - b_1(a_2v_2 + b_2v_1), a_1(a_2v_2 + b_2v_1) + b_1(a_2v_1 - b_2v_2) = \\ (a_1 + ib_1)(a_2v_1 - b_2v_2, a_2v_2 + b_2v_1) = (a_1 + ib_1)((a_2 + ib_2)(v_1, v_2)) = \\ c_1(c_2(v_1, v_2)) \end{gathered}$$

$$1(v_1, v_2) = (1 + i0)(v_1, v_2) := 1(v_1, v_2) + 0(J(v_1, v_2)) = (v_1, v_2)$$

This linear transformation J on $V_{re} \times V_{re}$ has been traditionally called an *almost complex structure* on $V_{re} \times V_{re}$. We also remark that it is canonical, i.e., it does not depend on a basis chosen in V_{re}, since it is made up from only the identity matrix I_n.

We now show that the mapping

$$V \longrightarrow V_{re} \times V_{re}$$
$$v = (v_1 + iv_2) \longmapsto (v_1, v_2)$$

is an isomorphism of complex linear space structures. It is clear that the mapping is 1-to-1 and onto; we need only show that addition and scalar multiplication are preserved.

$$v + w = (v_1 + iv_2) + (w_1 + iw_2) = (v_1 + w_1) + i(v_2 + w_2) \longmapsto$$
$$(v_1 + w_1, v_2 + w_2) = (v_1, v_2) + (w_1, w_2)$$

$$cv = (a+ib)(v_1+iv_2) = (av_1-bv_2)+i(bv_1+av_2) \longmapsto ((av_1-bv_2),(bv_1+av_2))$$
$$= (av_1, av_2) + (-bv_2, bv_1) = a(v_1, v_2) + b(-v_2, v_1) = a(v_1, v_2) + b(J(v_1, v_2)) =$$
$$(aI_{2n} + bJ)(v_1, v_2) = c(v_1, v_2)$$

But we can say still more. We have shown that if we give a complex linear space V any basis, we can construct a isomorphic complex linear space $V_{re} \times V_{re}$ from the real linear space $V_{re} \times V_{re}$, which latter cross product we called above V^r, by constructing an almost complex structure J on $V_{re} \times V_{re}$. If we now emphasize the choice of basis, we have the symbols $V^r = V_{re_e} \times V_{re_e}$ and $V^r = V_{re_u} \times V_{re_u}$ for the choice of two bases for V. And we now know that both of these spaces, given a almost complex structure J, possess a unique complex linear structure such that they are are both isomorphic to the complex linear space V. And V is the complexification $(V_{re})^c$ of the real linear space V_{re}.

In sum, we can therefore say the following: given a complex linear space V, the corresponding real linear space V^r loses some of the structure, and to regain the original complex structure we must define on V^r an almost complex structure J that is independent of bases, so that we can relate V^r to V in a compatible manner.

2.12.5 From $\mathbb{C}$ to $\mathbb{R}$: The Lie Algebra Structure.

Now we would like to examine the Lie algebra structures of these constructions. Thus we consider a complex linear space $\hat{g}$ which also has the structure of a Lie algebra. Since we are now not interested in a change of basis, we choose one basis $(u_1, \cdots, u_n)$ in $\hat{g}$ and fix it. With respect to this basis, we write $\hat{g} = \hat{g}_{re} \oplus i\hat{g}_{re}$ and $\hat{g}^r = \hat{g}_{re} \times \hat{g}_{re}$ Now for any x and y in $\hat{g}$, we have a bracket product $[x, y]$. We write $x = u_{re1}(x) + iu_{re2}(x)$ and $y = u_{re1}(y) + iu_{re2}(y)$ according to the above construction. However since we are fixing our basis (u_i) once and for all, we can shorten this notation to

$$x = u_{re1}(x) + iu_{re2}(x) = x_1 + ix_2 \qquad\qquad y = u_{re1}(y) + iu_{re2}(y) = y_1 + iy_2$$

for x_i and y_i in $\hat{g}_{re}$. Using this notation we calculate the bracket in $\hat{g}$:

$$[x,y] = [x_1 + ix_2, y_1 + iy_2] = ([x_1,y_1] - [x_2,y_2]) + i([x_1,y_2] + [x_2,y_1])$$

This gives the bracket in $\hat{g}$ with respect to the (u_i)-basis for $\hat{g}$. Now the brackets in $\hat{g}_{re}$ must be taken in the complex Lie algebra $\hat{g}$ for we have no reason to assert that $\hat{g}_{re}$ is a <u>real</u> Lie algebra with brackets closing in $\hat{g}_{re}$. On the contrary we must write

$$[x_i, y_j] = \alpha_{ij} + i\beta_{ij} \text{ in } \hat{g}, \qquad \alpha_{ij}, \beta_{ij} \text{ in } \hat{g}_{re}$$

Thus we have

$$\begin{aligned}[x,y] = [x_1 + ix_2, y_1 + iy_2] &= ([x_1,y_1] - [x_2,y_2]) + i([x_1,y_2] + [x_2,y_1]) \\ &= (\alpha_{11} + i\beta_{11}) - (\alpha_{22} + i\beta_{22}) + i((\alpha_{12} + i\beta_{12}) + (\alpha_{21} + i\beta_{21})) \\ &= (\alpha_{11} - \alpha_{22} - \beta_{12} - \beta_{21}) + i(\beta_{11} - \beta_{22} + \alpha_{12} + \alpha_{21})\end{aligned}$$

Also since we have chosen a basis to obtain the elements of this space, it will be convenient to express this space as column matrices of real numbers. [To achieve this, we will now use another convention to express matrices. Since we have only one basis under discussion, we will say that the vector x_1 will be represented in this basis by the symbol of a column matrix $[x_1]$, that is, we will put brackets around the vector symbol in order to represent the column matrix representing this vector. Thus even though brackets will mean column matrices as well as the bracket product of two vectors or two matrices, we hope the context will be sufficient to distinguish these two usages.] Thus in this matrix notation we have

$$\left[\begin{bmatrix}[x_1]\\ [x_2]\end{bmatrix}, \begin{bmatrix}[y_1]\\ [y_2]\end{bmatrix}\right] = \begin{bmatrix}[\alpha_{11}] - [\alpha_{22}] - [\beta_{12}] - [\beta_{21}]\\ [\beta_{11}] - [\beta_{22}] + [\alpha_{12}] + [\alpha_{21}]\end{bmatrix}$$

We remark if we just started with x_1 and y_1 in $\hat{g}_{re}$, we would obtain

$$\left[\begin{bmatrix}[x_1]\\ [0]\end{bmatrix}, \begin{bmatrix}[y_1]\\ [0]\end{bmatrix}\right] = \begin{bmatrix}[\alpha_{11}]\\ [\beta_{11}]\end{bmatrix}$$

which again shows that the value of the bracket product in $\hat{g}_{re}$ is not real, but is complex.

Thus we have

$$[x_1 + ix_2, y_1 + iy_2)] := ((\alpha_{11} - \alpha_{22} - \beta_{12} - \beta_{21}) + i(\beta_{11} - \beta_{22} + \alpha_{12} + \alpha_{21})$$

We would now like to confirm that this definition of a Lie bracket in $\hat{g}^r$ does give to $\hat{g}^r = \hat{g}_{re} \times \hat{g}_{re}$ the structure of a complex Lie algebra. In the following calculations we will use $[x_i, z_k] = \gamma_{ik} + i\delta_{ik}$ to make clear that the brackets map what is inside to a complex number.

We have distribution on the left. In the following calculation we are using the fact that all x_i, y_j and z_k belong to V and thus we are using the distribution property in V when we write

$$[x_i, y_j + z_k] = [x_i, y_j] + [x_i, z_k]$$

Now

$$[(x_1, x_2), (y_1, y_2) + (z_1, z_2)] = [(x_1, x_2), (y_1 + z_1, y_2 + z_2)]$$

By our definition we have

$$\begin{gathered}
[x_1, y_1 + z_1] = [x_1, y_1] + [x_1, z_1] = (\alpha_{11}, \beta_{11}) + (\gamma_{11}, \delta_{11}) = \\
(\alpha_{11} + \gamma_{11}, \beta_{11} + \delta_{11}) \\
[x_2, y_2 + z_2] = [x_2, y_2] + [x_2, z_2] = (\alpha_{22}, \beta_{22}) + (\gamma_{22}, \delta_{22}) = \\
(\alpha_{22} + \gamma_{22}, \beta_{22} + \delta_{22}) \\
[x_1, y_2 + z_2] = [x_1, y_2] + [x_1, z_2] = (\alpha_{12}, \beta_{12}) + (\gamma_{12}, \delta_{12}) = \\
(\alpha_{12} + \gamma_{12}, \beta_{12} + \delta_{12}) \\
[x_2, y_1 + z_1] = [x_2, y_1] + [x_2, z_1] = (\alpha_{21}, \beta_{21}) + (\gamma_{21}, \delta_{21}) = \\
(\alpha_{21} + \gamma_{21}, \beta_{21} + \delta_{21})
\end{gathered}$$

Continuing

$$\begin{gathered}
[(x_1, x_2), (y_1, y_2) + (z_1, z_2)] = [(x_1, x_2), (y_1 + z_1, y_2 + z_2)] = \\
(\alpha_{11} + \gamma_{11}, \beta_{11} + \delta_{11}) - (\alpha_{22} + \gamma_{22}, \beta_{22} + \delta_{22}) + \\
J(\alpha_{12} + \gamma_{12}, \beta_{12} + \delta_{12}) + J(\alpha_{21} + \gamma_{21}, \beta_{21} + \delta_{21}) = \\
(\alpha_{11} + \gamma_{11} - \alpha_{22} - \gamma_{22} - \beta_{12} - \delta_{12} - \beta_{21} - \delta_{21}, \\
\beta_{11} + \delta_{11} - \beta_{22} - \delta_{22} + \alpha_{12} + \gamma_{12} + \alpha_{21} + \gamma_{21})
\end{gathered}$$

Now we calculate $[(x_1, x_2), (y_1, y_2)]$ and $[(x_1, x_2), (z_1, z_2)]$.

$$\begin{gathered}
[(x_1, x_2), (y_1, y_2)] = (\alpha_{11}, \beta_{11}) - (\alpha_{22}, \beta_{22}) + J(\alpha_{12}, \beta_{12}) + J(\alpha_{21}, \beta_{21}) = \\
(\alpha_{11} - \alpha_{22} - \beta_{12} - \beta_{21}, \beta_{11} - \beta_{22} + \alpha_{12} + \alpha_{21})
\end{gathered}$$

$$\begin{gathered}
[(x_1, x_2), (z_1, z_2)] = (\gamma_{11}, \delta_{11}) - (\gamma_{22}, \delta_{22}) + J(\gamma_{12}, \delta_{12}) + J(\gamma_{21}, \delta_{21}) = \\
(\gamma_{11} - \gamma_{22} - \delta_{12} - \delta_{21}, \delta_{11} - \delta_{22} + \gamma_{12} + \gamma_{21})
\end{gathered}$$

Adding $[(x_1, x_2), (y_1, y_2)]$ and $[(x_1, x_2), (z_1, z_2)]$, we obtain

$$\begin{gathered}(\alpha_{11}-\alpha_{22}-\beta_{12}-\beta_{21},\beta_{11}-\beta_{22}+\alpha_{12}+\alpha_{21})+\\(\gamma_{11}-\gamma_{22}-\delta_{12}-\delta_{21},\delta_{11}-\delta_{22}+\gamma_{12}+\gamma_{21})=\\(\alpha_{11}-\alpha_{22}-\beta_{12}-\beta_{21}+\gamma_{11}-\gamma_{22}-\delta_{12}-\delta_{21},\\\beta_{11}-\beta_{22}+\alpha_{12}+\alpha_{21}+\delta_{11}-\delta_{22}+\gamma_{12}+\gamma_{21})\end{gathered}$$

Thus we see that

$$[(x_1,x_2),(y_1,y_2)+(z_1,z_2)]=[(x_1,x_2),(y_1,y_2)]+[(x_1,x_2),(z_1,z_2)].$$

and we have distribution on the left. It is evident that if we use the same kind of argument we will also show that we have distribution on the right.

We now have to show that for c in $\mathbb{C}$

$$c[(x_1,x_2),(y_1,y_2)]=[c(x_1,x_2),(y_1,y_2)]=[(x_1,x_2),c(y_1,y_2)].$$

First we calculate $c[(x_1,x_2),(y_1,y_2)]$. We write $c=a+ib$.

$$\begin{gathered}(a+ib)[(x_1,x_2),(y_1,y_2)]=(a+ib)((\alpha_{11}-\alpha_{22}-\beta_{12}-\beta_{21},\beta_{11}-\beta_{22}+\alpha_{12}+\alpha_{21}))=\\(a(\alpha_{11}-\alpha_{22}-\beta_{12}-\beta_{21})-b(\beta_{11}-\beta_{22}+\alpha_{12}+\alpha_{21}),\\a(\beta_{11}-\beta_{22}+\alpha_{12}+\alpha_{21})+b(\alpha_{11}-\alpha_{22}-\beta_{12}-\beta_{21}))\end{gathered}$$

Next we calculate

$$\begin{gathered}[c(x_1,x_2),(y_1,y_2)]=\\ [(a+ib)(x_1,x_2),(y_1,y_2)]=[(ax_1-bx_2,ax_2+bx_1),(y_1,y_2)]\end{gathered}$$

By our definition we have

$$\begin{gathered}[ax_1-bx_2,y_1]=[ax_1,y_1]-[bx_2,y_1]=a[x_1,y_1]-b[x_2,y_1]=\\a(\alpha_{11},\beta_{11})-b(\alpha_{21},\beta_{21})\\ [ax_2+bx_1,y_2]=[ax_2,y_2]+[bx_1,y_2]=a[x_2,y_2]+b[x_1,y_2]=\\a(\alpha_{22},\beta_{22})+b(\alpha_{12},\beta_{12})\\ [ax_1-bx_2,y_2]=[ax_1,y_2]-[bx_2,y_2]=a[x_1,y_2]-b[x_2,y_2]=\\a(\alpha_{12},\beta_{12})-b(\alpha_{22},\beta_{22})\\ [ax_2+bx_1,y_1]=[ax_2,y_1]+[bx_1,y_1]=a[x_2,y_1]+b[x_1,y_1]=\\a(\alpha_{21},\beta_{21})+b(\alpha_{11},\beta_{11})\end{gathered}$$

Continuing

$$\begin{gathered}[c(x_1,x_2),(y_1,y_2)]=[(ax_1-bx_2,ax_2+bx_1),(y_1,y_2)]=\\a(\alpha_{11},\beta_{11})-b(\alpha_{21},\beta_{21})-(a(\alpha_{22},\beta_{22})+b(\alpha_{12},\beta_{12}))\\+J(a(\alpha_{12},\beta_{12})-b(\alpha_{22},\beta_{22}))+J(a(\alpha_{21},\beta_{21})+b(\alpha_{11},\beta_{11}))=\\a(\alpha_{11},\beta_{11})-b(\alpha_{21},\beta_{21})-a(\alpha_{22},\beta_{22})-b(\alpha_{12},\beta_{12})\\+aJ(\alpha_{12},\beta_{12})-bJ(\alpha_{22},\beta_{22}))+aJ(\alpha_{21},\beta_{21})+bJ(\alpha_{11},\beta_{11}))=\end{gathered}$$

$$\begin{gathered}
a(\alpha_{11},\beta_{11})-b(\alpha_{21},\beta_{21})-a(\alpha_{22},\beta_{22})-b(\alpha_{12},\beta_{12})\\
+a(-\beta_{12},\alpha_{12})-b(-\beta_{22},\alpha_{22})+a(-\beta_{21},\alpha_{21})+b(-\beta_{11},\alpha_{11})=\\
(a\alpha_{11}-b\alpha_{21}-a\alpha_{22}-b\alpha_{12}-a\beta_{12}+b\beta_{22}-a\beta_{21}-b\beta_{11},\\
a\beta_{11}-b\beta_{21}-a\beta_{22}-b\beta_{12}+a\alpha_{12}-b\alpha_{22}+a\alpha_{21}+b\alpha_{11})=\\
(a(\alpha_{11}-\alpha_{22}-\beta_{12}-\beta_{21})-b(\alpha_{21}+\alpha_{12}-\beta_{22}+\beta_{11}),\\
a(\beta_{11}-\beta_{22}+\alpha_{12}+\alpha_{21})+b(-\beta_{21}-\beta_{12}-\alpha_{22}+\alpha_{11}))
\end{gathered}$$

Finally, we calculate

$$\begin{gathered}
[(x_1,x_2),c(y_1,y_2)]=\\
[(x_1,x_2),(a+ib)(y_1,y_2)]=[(x_1,x_2),(ay_1-by_2,ay_2+by_1)]
\end{gathered}$$

By our definition we have

$$\begin{gathered}
[x_1,ay_1-by_2]=[x_1,ay_1]-[x_1,by_2]=a[x_1,y_1]-b[x_1,y_2]=\\
a(\alpha_{11},\beta_{11})-b(\alpha_{12},\beta_{12})\\
[x_2,ay_2+by_1]=[x_2,ay_2]+[x_2,by_1]=a[x_2,y_2]+b[x_2,y_1]=\\
a(\alpha_{22},\beta_{22})+b(\alpha_{21},\beta_{21})
\end{gathered}$$

$$\begin{gathered}
[x_1,ay_2+by_1]=[x_1,ay_2]+[x_1,by_1]=a[x_1,y_2]+b[x_1,y_1]=\\
a(\alpha_{12},\beta_{12})+b(\alpha_{11},\beta_{11})\\
[x_2,ay_1-by_2]=[x_2,ay_1]-[x_2,by_2]=a[x_2,y_1]-b[x_2,y_2]=\\
a(\alpha_{21},\beta_{21})-b(\alpha_{22},\beta_{22})
\end{gathered}$$

Continuing

$$\begin{gathered}
[(x_1,x_2),c(y_1,y_2)]=[(x_1,x_2),(ay_1-by_2,ay_2+by_1)]=\\
a(\alpha_{11},\beta_{11})-b(\alpha_{12},\beta_{12})-(a(\alpha_{22},\beta_{22})+b(\alpha_{21},\beta_{21}))\\
+J(a(\alpha_{12},\beta_{12})+b(\alpha_{11},\beta_{11}))+J(a(\alpha_{21},\beta_{21})-b(\alpha_{22},\beta_{22}))=\\
a(\alpha_{11},\beta_{11})-b(\alpha_{12},\beta_{12})-a(\alpha_{22},\beta_{22})-b(\alpha_{21},\beta_{21})\\
+aJ(\alpha_{12},\beta_{12})+bJ(\alpha_{11},\beta_{11}))+aJ(\alpha_{21},\beta_{21})-bJ(\alpha_{22},\beta_{22}))=\\
a(\alpha_{11},\beta_{11})-b(\alpha_{12},\beta_{12})-a(\alpha_{22},\beta_{22})-b(\alpha_{21},\beta_{21})\\
+a(-\beta_{12},\alpha_{12})+b(-\beta_{11},\alpha_{11})+a(-\beta_{21},\alpha_{21})-b(-\beta_{22},\alpha_{22})=\\
(a\alpha_{11}-b\alpha_{12}-a\alpha_{22}-b\alpha_{21}-a\beta_{12}-b\beta_{11}-a\beta_{21}+b\beta_{22},\\
a\beta_{11}-b\beta_{12}-a\beta_{22}-b\beta_{21}+a\alpha_{12}+b\alpha_{11}+a\alpha_{21}-b\alpha_{22})=\\
(a(\alpha_{11}-\alpha_{22}-\beta_{12}-\beta_{21})-b(\alpha_{12}+\alpha_{21}+\beta_{11}-\beta_{22}),\\
a(\beta_{11}-\beta_{22}+\alpha_{12}+\alpha_{21})+b(-\beta_{12}-\beta_{21}+\alpha_{11}-\alpha_{22}))
\end{gathered}$$

Thus we can conclude that scalar multiplication by $\mathbb{C}$ satisfies

$$c[(x_1,x_2),(y_1,y_2)]=[c(x_1,x_2),(y_1,y_2)]=[(x_1,x_2),c(y_1,y_2)].$$

We now verify the anticommutativity of the bracket product. In order to write $[(y_1,y_2),(x_1,x_2)]$, we must return to the calculation of $[y,x]$ in V, which, of course, gives immediately $[y,x]=-[x,y]$. We use this to show $[(y_1,y_2),(x_1,x_2)]=-[(x_1,x_2),(y_1,y_2)]$

$$\begin{gathered}[y,x]=[y_1+iy_2,x_1+ix_2]=([y_1,x_1]-[y_2,x_2])+i([y_1,x_2]+[y_2,x_1])=\\(-[x_1,y_1]+[x_2,y_2])-i([x_2,y_1]+[x_1,y_2])=\\-(\alpha_{11},\beta_{11})+(\alpha_{22},\beta_{22})-i((\alpha_{21},\beta_{21})+(\alpha_{12},\beta_{12}))\end{gathered}$$

Now by definition

$$\begin{gathered}[(y_1,y_2),(x_1,x_2)]:=-(\alpha_{11},\beta_{11})+(\alpha_{22},\beta_{22})-J(\alpha_{12},\beta_{12})-J(\alpha_{21},\beta_{21})=\\-(\alpha_{11},\beta_{11})+(\alpha_{22},\beta_{22})-(-\beta_{12},\alpha_{12})-(-\beta_{21},\alpha_{21})=\\(-\alpha_{11}+\alpha_{22}+\beta_{12}+\beta_{21},-\beta_{11}+\beta_{22}-\alpha_{12}-\alpha_{21})\end{gathered}$$

And we see immediately that

$$\begin{gathered}[(y_1,y_2),(x_1,x_2)]=(-\alpha_{11}+\alpha_{22}+\beta_{12}+\beta_{21},-\beta_{11}+\beta_{22}-\alpha_{12}-\alpha_{21})=\\-(\alpha_{11}-\alpha_{22}-\beta_{12}-\beta_{21},\beta_{11}-\beta_{22}+\alpha_{12}+\alpha_{21}))=\\-[(x_1,x_2),(y_1,y_2)]\end{gathered}$$

Finally, we must verify the Jacobi identity in $\hat{g}^r=\hat{g}_{re}\times\hat{g}_{re}$.

First we calculate $[[(x_1,x_2),(y_1,y_2)],(z_1,z_2)]$. In this calculation we have $[x_i,y_j]=(\alpha_{ij},\beta_{ij})$, $[\alpha_{ij},z_k]=(\alpha_{ijk},\beta_{ijk})$ and $[\beta_{ij},z_k]=(\gamma_{ijk},\delta_{ijk})$.

$$\begin{gathered}[[(x_1,x_2),(y_1,y_2)],(z_1,z_2)]=\\[(\alpha_{11}-\alpha_{22}-\beta_{12}-\beta_{21},\beta_{11}-\beta_{22}+\alpha_{12}+\alpha_{21}),(z_1,z_2)]=\\[(\alpha_{11}-\alpha_{22}-\beta_{12}-\beta_{21}),z_1]-[(\beta_{11}-\beta_{22}+\alpha_{12}+\alpha_{21}),z_2]\\+J([(\alpha_{11}-\alpha_{22}-\beta_{12}-\beta_{21}),z_2])+J([(\beta_{11}-\beta_{22}+\alpha_{12}+\alpha_{21}),z_1])=\\[\alpha_{11},z_1]-[\alpha_{22},z_1]-[\beta_{12},z_1]-[\beta_{21},z_1]\\-[\beta_{11},z_2]+[\beta_{22},z_2]-[\alpha_{12},z_2]-[\alpha_{21}),z_2]\\+J([\alpha_{11},z_2]-[\alpha_{22},z_2]-[\beta_{12},z_2]-[\beta_{21},z_2])\\+J([\beta_{11},z_1]-[\beta_{22},z_1]+[\alpha_{12},z_1]+[\alpha_{21},z_1])=\\(\alpha_{111},\beta_{111})-(\alpha_{221},\beta_{221})-(\gamma_{121},\delta_{121})-(\gamma_{211},\delta_{211})\\-(\gamma_{112},\delta_{112})+(\gamma_{222},\delta_{222})-(\alpha_{122},\beta_{122})-(\alpha_{212},\beta_{212})\\+J((\alpha_{112},\beta_{112})-(\alpha_{222},\beta_{222})-(\gamma_{122},\delta_{122})-(\gamma_{212},\delta_{212}))\\+J((\gamma_{111},\delta_{111})-(\gamma_{221},\delta_{221})+(\alpha_{121},\beta_{121})+(\alpha_{211},\beta_{211}))=\\(\alpha_{111},\beta_{111})-(\alpha_{221},\beta_{221})-(\gamma_{121},\delta_{121})-(\gamma_{211},\delta_{211})\\-(\gamma_{112},\delta_{112})+(\gamma_{222},\delta_{222})-(\alpha_{122},\beta_{122})-(\alpha_{212},\beta_{212})\\+(-\beta_{112},\alpha_{112})-(-\beta_{222},\alpha_{222})-(-\delta_{122},\gamma_{122})-(-\delta_{212},\gamma_{212})\\+(-\delta_{111},\gamma_{111})-(-\delta_{221},\gamma_{221})+(-\beta_{121},\alpha_{121})+(-\beta_{211},\alpha_{211})=\\(\alpha_{111}-\alpha_{221}-\gamma_{121}-\gamma_{211}-\gamma_{112}+\gamma_{222}-\alpha_{122}-\alpha_{212}\\-\beta_{112}+\beta_{222}+\delta_{122}+\delta_{212}-\delta_{111}+\delta_{221}-\beta_{121}-\beta_{211},\\\beta_{111}-\beta_{221}-\delta_{121}-\delta_{211}-\delta_{112}+\delta_{222}-\beta_{122}-\beta_{212}\\+\alpha_{112}-\alpha_{222}-\gamma_{122}-\gamma_{212}+\gamma_{111}-\gamma_{221}+\alpha_{121}+\alpha_{211})\end{gathered}$$

Next we calculate $[[(z_1,z_2),(x_1,x_2)],(y_1,y_2)]$. In this calculation we have $[z_i,x_j]=-(\gamma_{ij},\delta_{ij})$, $[\gamma_{ij},y_k]=(\kappa_{ijk},\lambda_{ijk})$ and $[\delta_{ij},y_k]=(\mu_{ijk},\nu_{ijk})$.

$$
\begin{gathered}
[[(z_1,z_2),(x_1,x_2)],(y_1,y_2)] = \\
[(-\gamma_{11}+\gamma_{22}+\delta_{12}+\delta_{21},-\delta_{11}+\delta_{22}-\gamma_{12}-\gamma_{21}),(y_1,y_2)] = \\
[(-\gamma_{11}+\gamma_{22}+\delta_{12}+\delta_{21}),y_1]-[(-\delta_{11}+\delta_{22}-\gamma_{12}-\gamma_{21}),y_2] \\
+J([(-\gamma_{11}+\gamma_{22}+\delta_{12}+\delta_{21}),y_2])+J([(-\delta_{11}+\delta_{22}-\gamma_{12}-\gamma_{21}),y_1]) = \\
[-\gamma_{11},y_1]+[\gamma_{22},y_1]+[\delta_{12},y_1]+[\delta_{21},y_1] \\
+[\delta_{11},y_2]-[\delta_{22},y_2]+[\gamma_{12},y_2]+[\gamma_{21}),y_2] \\
+J([-\gamma_{11},y_2]+[\gamma_{22},y_2]+[\delta_{12},y_2]+[\delta_{21},y_2]) \\
+J(-[\delta_{11},y_1]+[\delta_{22},y_1]-[\gamma_{12},y_1]-[\gamma_{21},y_1]) = \\
-(\kappa_{111},\lambda_{111})+(\kappa_{221},\lambda_{221})+(\mu_{121},\nu_{121})+(\mu_{211},\nu_{211}) \\
+(\mu_{112},\nu_{112})-(\mu_{222},\nu_{222})+(\kappa_{122},\lambda_{122})+(\kappa_{212},\lambda_{212}) \\
J(-(\kappa_{112},\lambda_{112})+(\kappa_{222},\lambda_{222})+(\mu_{122},\nu_{122})+(\mu_{212},\nu_{212})) \\
J(-(\mu_{111},\nu_{111})+(\mu_{221},\nu_{221})-(\kappa_{121},\lambda_{121})-(\kappa_{211},\lambda_{211})) = \\
-(\kappa_{111},\lambda_{111})+(\kappa_{221},\lambda_{221})+(\mu_{121},\nu_{121})+(\mu_{211},\nu_{211}) \\
+(\mu_{112},\nu_{112})-(\mu_{222},\nu_{222})+(\kappa_{122},\lambda_{122})+(\kappa_{212},\lambda_{212}) \\
+(\lambda_{112},-\kappa_{112})+(-\lambda_{222},\kappa_{222})+(-\nu_{122},\mu_{122})+(-\nu_{212},\mu_{212}) \\
+(\nu_{111},-\mu_{111})+(-\nu_{221},\mu_{221})-(-\lambda_{121},\kappa_{121})-(-\lambda_{211},\kappa_{211}) = \\
(-\kappa_{111}+\kappa_{221}+\mu_{121}+\mu_{211}+\mu_{112}-\mu_{222}+\kappa_{122}+\kappa_{212} \\
\lambda_{112}-\lambda_{222}-\nu_{122}-\nu_{212}+\nu_{111}-\nu_{221}+\lambda_{121}+\lambda_{211}, \\
-\lambda_{111}+\lambda_{221}+\nu_{121}+\nu_{211}+\nu_{112}-\nu_{222}+\lambda_{122}+\lambda_{212} \\
-\kappa_{112}+\kappa_{222}+\mu_{122}+\mu_{212}-\mu_{111}+\mu_{221}-\kappa_{121}-\kappa_{211})
\end{gathered}
$$

Lastly, we calculate $[[(y_1,y_2),(z_1,z_2)],(x_1,x_2)]$. In this calculation we have $[y_i,z_j]=(\kappa_{ij},\lambda_{ij})$, $[\kappa_{ij},x_k]=(\pi_{ijk},\rho_{ijk})$ and $[\lambda_{ij},x_k]=(\sigma_{ijk},\tau_{ijk})$

$$
\begin{gathered}
[(y_1,y_2),(z_1,z_2)],(x_1,x_2)] = \\
[(\kappa_{11}-\kappa_{22}-\lambda_{12}-\lambda_{21},\lambda_{11}-\lambda_{22}+\kappa_{12}+\kappa_{21}),(x_1,x_2)] = \\
[(\kappa_{11}-\kappa_{22}-\lambda_{12}-\lambda_{21}),x_1]-[(\lambda_{11}-\lambda_{22}+\kappa_{12}+\kappa_{21}),x_2] \\
+J([(\kappa_{11}-\kappa_{22}-\lambda_{12}-\lambda_{21}),x_2])+J([(\lambda_{11}-\lambda_{22}+\kappa_{12}+\kappa_{21}),x_1]) = \\
[\kappa_{11},x_1]-[\kappa_{22},x_1]-[\lambda_{12},x_1]-[\lambda_{21},x_1] \\
-[\lambda_{11},x_2]+[\lambda_{22},x_2]-[\kappa_{12},x_2]-[\kappa_{21},x_2] \\
+J([\kappa_{11},x_2]-[\kappa_{22},x_2]-[\lambda_{12},x_2]-[\lambda_{21},x_2]) \\
+J([\lambda_{11},x_1]-[\lambda_{22},x_1]+[\kappa_{12},x_1]+[\kappa_{21},x_1]) = \\
(\pi_{111},\rho_{111})-(\pi_{221},\rho_{221})-(\sigma_{121},\tau_{121})-(\sigma_{211},\tau_{211}) \\
-(\sigma_{112},\tau_{112})+(\sigma_{222},\tau_{222})-(\pi_{122},\rho_{122})-(\pi_{212},\rho_{212}) \\
+J((\pi_{112},\rho_{112})-(\pi_{222},\rho_{222})-(\sigma_{122},\tau_{122})-(\sigma_{212},\tau_{212})) \\
+J((\sigma_{111},\tau_{111})-(\sigma_{221},\tau_{221})+(\pi_{121},\rho_{121})+(\pi_{211},\rho_{211})) = \\
(\pi_{111},\rho_{111})-(\pi_{221},\rho_{221})-(\sigma_{121},\tau_{121})-(\sigma_{211},\tau_{211}) \\
-(\sigma_{112},\tau_{112})+(\sigma_{222},\tau_{222})-(\pi_{122},\rho_{122})-(\pi_{212},\rho_{212}) \\
+(-\rho_{112},\pi_{112})-(-\rho_{222},\pi_{222})-(-\tau_{122},\sigma_{122})-(-\tau_{212},\sigma_{212}) \\
+(-\tau_{111},\sigma_{111})-(-\tau_{221},\sigma_{221})+(-\rho_{121},\pi_{121})+(-\rho_{211},\pi_{211}) = \\
(\pi_{111}-\pi_{221}-\sigma_{121}-\sigma_{211}-\sigma_{112}+\sigma_{222}-\pi_{122}-\pi_{212} \\
-\rho_{112}+\rho_{222}+\tau_{122}+\tau_{212}-\tau_{111}+\tau_{221}-\rho_{121}-\rho_{211}, \\
\rho_{111}-\rho_{221}-\tau_{121}-\tau_{211}-\tau_{112}+\tau_{222}-\rho_{122}-\rho_{212}
\end{gathered}
$$

$$+\pi_{112} - \pi_{222} - \sigma_{122} - \sigma_{212} + \sigma_{111} - \sigma_{221} + \pi_{121} + \pi_{211})$$

To complete the proof we must show that the sums of the following three expressions are each equal to 0.

$$\begin{gathered}(\alpha_{111} - \alpha_{221} - \gamma_{121} - \gamma_{211} - \gamma_{112} + \gamma_{222} - \alpha_{122} - \alpha_{212} \\ -\beta_{112} + \beta_{222} + \delta_{122} + \delta_{212} - \delta_{111} + \delta_{221} - \beta_{121} - \beta_{211}, \\ \beta_{111} - \beta_{221} - \delta_{121} - \delta_{211} - \delta_{112} + \delta_{222} - \beta_{122} - \beta_{212} \\ +\alpha_{112} - \alpha_{222} - \gamma_{122} - \gamma_{212} + \gamma_{111} - \gamma_{221} + \alpha_{121} + \alpha_{211})\end{gathered}$$

$$\begin{gathered}(-\kappa_{111} + \kappa_{221} + \mu_{121} + \mu_{211} + \mu_{112} - \mu_{222} + \kappa_{122} + \kappa_{212} \\ \lambda_{112} - \lambda_{222} - \nu_{122} - \nu_{212} + \nu_{111} - \nu_{221} + \lambda_{121} + \lambda_{211}, \\ -\lambda_{111} + \lambda_{221} + \nu_{121} + \nu_{211} + \nu_{112} - \nu_{222} + \lambda_{122} + \lambda_{212} \\ -\kappa_{112} + \kappa_{222} + \mu_{122} + \mu_{212} - \mu_{111} + \mu_{221} - \kappa_{121} - \kappa_{211})\end{gathered}$$

and

$$\begin{gathered}(\pi_{111} - \pi_{221} - \sigma_{121} - \sigma_{211} - \sigma_{112} + \sigma_{222} - \pi_{122} - \pi_{212} \\ -\rho_{112} + \rho_{222} + \tau_{122} + \tau_{212} - \tau_{111} + \tau_{221} - \rho_{121} - \rho_{211}, \\ \rho_{111} - \rho_{221} - \tau_{121} - \tau_{211} - \tau_{112} + \tau_{222} - \rho_{122} - \rho_{212} \\ +\pi_{112} - \pi_{222} - \sigma_{122} - \sigma_{212} + \sigma_{111} - \sigma_{221} + \pi_{121} + \pi_{211})\end{gathered}$$

Now, knowing that x_i,y_j, and z_k are all in V, we can use the Jacobi identity in V to make the following calculations:

$$\begin{gathered}[[x_i, y_j], z_k] = [(\alpha_{ij}, \beta_{ij}), (z_k, 0)] = \\ [\alpha_{ij}, z_k] - [\beta_{ij}, 0] + J([\alpha_{ij}, 0]) + J([\beta_{ij}, z_k]) = \\ (\alpha_{ijk}, \beta_{ijk}) + J(\gamma_{ijk}, \delta_{ijk}) = (\alpha_{ijk}, \beta_{ijk}) + (-\delta_{ijk}, \gamma_{ijk}) = \\ (\alpha_{ijk} - \delta_{ijk}, \beta_{ijk} + \gamma_{ijk})\end{gathered}$$

$$\begin{gathered}[[z_i, x_j], y_k] = [-(\gamma_{ij}, \delta_{ij}), (y_k, 0)] = \\ [-\gamma_{ij}, y_k] - [-\delta_{ij}, 0] + J([-\gamma_{ij}, 0]) + J([-\delta_{ij}, y_k]) = \\ -(\kappa_{ijk}, \lambda_{ijk}) - J(\mu_{ijk}, \nu_{ijk}) = -(\kappa_{ijk}, \lambda_{ijk}) - (-\nu_{ijk}, \mu_{ijk}) = \\ (-\kappa_{ijk} + \nu_{ijk}, -\lambda_{ijk} - \mu_{ijk})\end{gathered}$$

$$\begin{gathered}[[y_i, z_j], x_k] = [(\kappa_{ij}, \lambda_{ij}), (x_k, 0)] = \\ [\kappa_{ij}, x_k] - [\lambda_{ij}, 0] + J([\kappa_{ij}, 0]) + J([\lambda_{ij}, x_k]) = \\ (\pi_{ijk}, \rho_{ijk}) + J(\sigma_{ijk}, \tau_{ijk}) = (\pi_{ijk}, \rho_{ijk}) + (-\tau_{ijk}, \sigma_{ijk}) = \\ (\pi_{ijk} - \tau_{ijk}, \rho_{ijk} + \sigma_{ijk})\end{gathered}$$

Thus by the Jacobi identity in V we have:

$$\begin{gathered}(\alpha_{ijk} - \delta_{ijk}, \beta_{ijk} + \gamma_{ijk}) + (-\kappa_{ijk} + \nu_{ijk}, -\lambda_{ijk} - \mu_{ijk}) + (\pi_{ijk} - \tau_{ijk}, \rho_{ijk} + \sigma_{ijk}) = 0 \\ (\alpha_{ijk} - \delta_{ijk} - \kappa_{ijk} + \nu_{ijk} + \pi_{ijk} - \tau_{ijk}, \beta_{ijk} + \gamma_{ijk} - \lambda_{ijk} - \mu_{ijk} + \rho_{ijk} + \sigma_{ijk}) = 0\end{gathered}$$

Thus we have

$$\begin{array}{c}(\alpha_{111} - \delta_{111} - \kappa_{111} + \nu_{111} + \pi_{111} - \tau_{111} = 0 \\ \beta_{111} + \gamma_{111} - \lambda_{111} - \mu_{111} + \rho_{111} + \sigma_{111}) = 0\end{array}$$

and, using similar arguments, we have the same result for the indices {112}, {121}, {122}, {211}, {212}, {221}, and {222}. giving us that the sum of the above three expressions are indeed each equal to 0. This proves the Jacobi identity in $\hat{g}^r = \hat{g}_{re} \times \hat{g}_{re}$.

Finally, we show that there is an isomorphism of Lie algebras between V and $\hat{g}_{re} \times \hat{g}_{re}$. We already know as $\mathbb{C}$-linear spaces that they are isomorphic. Thus we only need to show that brackets go to brackets. Now for v and w in V we have

$$\begin{array}{c}V \longrightarrow V_{re} \times V_{re} \\ v = (v_1 + iv_2) \longmapsto (v_1, v_2) \\ w = (w_1 + iw_2) \longmapsto (w_1, w_2)\end{array}$$

If in the following calculation we let $[v_i, w_j] = \alpha_{ij} + i\beta_{ij}$ then we have

$$\begin{array}{c}[v, w] = [v_1 + iv_2, w_1 + iw_2] = [v_1, w_1] - [v_2, w_2] + i([v_1, w_2] + [v_2, w_1]) = \\ \alpha_{11} + i\beta_{11} - (\alpha_{22} + i\beta_{22}) + i((\alpha_{12} + i\beta_{12}) + (\alpha_{21} + i\beta_{21})) = \\ (\alpha_{11} - \alpha_{22} - \beta_{12} - \beta_{21}) + i(\beta_{11} - \beta_{22} + \alpha_{12} + \alpha_{21}) \longmapsto \\ (\alpha_{11} - \alpha_{22} - \beta_{12} - \beta_{21}, \beta_{11} - \beta_{22} + \alpha_{12} + \alpha_{21})\end{array}$$

Now we calculate $[(v_1, v_2), (w_1, w_2)]$ in $\hat{g}_{re} \times \hat{g}_{re}$.

$$\begin{array}{c}[(v_1, v_2), (w_1, w_2)] = (\alpha_{11}, \beta_{11}) - (\alpha_{22}, \beta_{22}) + J(\alpha_{12}, \beta_{12}) + J(\alpha_{21}, \beta_{21}) = \\ (\alpha_{11}, \beta_{11}) - (\alpha_{22}, \beta_{22}) + (-\beta_{12}, \alpha_{12}) + (-\beta_{21}, \alpha_{21}) = \\ (\alpha_{11} - \alpha_{22} - \beta_{12} - \beta_{21}, \beta_{11} - \beta_{22} + \alpha_{12} + \alpha_{21})\end{array}$$

Thus we can conclude that we have an isomorphism of $\mathbb{C}$-Lie algebras.

At this point, it would be good to recall that the reason for making all these calculations is to be able to move from $\mathbb{C}$-Lie algebras to $\mathbb{R}$-Lie algebras. We saw that just expressing V as $\hat{g}^r = \hat{g}_{re} \times \hat{g}_{re}$ was just not enough. However our search might still be able to find $\mathbb{R}$-Lie algebras in the $\mathbb{C}$-Lie algebra $\hat{g}_{re} \times \hat{g}_{re}$. For example, if perchance we found a basis for V such that $\hat{g}_{re} = \hat{g}_{re} \times 0 \subset \hat{g}_{re} \times \hat{g}_{re}$ is real with the Lie bracket as defined above, then we could write the following.

For x, y in V, we write (x_1, x_2) and (y_1, y_2) in $\hat{g}_{re} \times \hat{g}_{re}$. Now $[x, y]$ in V is written as

$$[x_1 + ix_2, y_1 + iy_2] = [x_1, y_1] - [x_2, y_2] + i([x_1, y_2] + [x_2, y_1])$$

Now we wrote $[x_i, y_j] = \alpha_{ij} + i\beta_{ij}$, since the bracket was still a complex quantity. However if we knew that $[x_i, y_j]$ was real, then we would have $[x_i, y_j] = \alpha_{ij} + i0$, and we would have no need for the complex notation and then the following expression becomes

$$[(x_1, x_2), (y_1, y_2)] = ([x_1, y_1], 0) - ([x_2, y_2], 0) + J(([x_1, y_2], 0)) + J(([x_2, y_1], 0)) =$$
$$([x_1, y_1], 0) - ([x_2, y_2], 0) + (-0, [x_1, y_2]) + (-0, [x_2, y_1]) =$$
$$([x_1, y_1] - [x_2, y_2], [x_1, y_2] + [x_2, y_1])$$

and if we restrict to just $\hat{g}_{re} = \hat{g}_{re} \times 0 \subset \hat{g}_{re} \times \hat{g}_{re}$, we would obtain

$$[(x_1, 0), (y_1, 0)] = ([x_1, y_1] - 0, 0 + 0) = ([x_1, y_1], 0)$$

which just says that in $\hat{g}_{re}$ we have a real valued bracket product making it into a real Lie algebra, which is what we assumed in the beginning.

However we can say more, namely that $\hat{g}_{re} \times \hat{g}_{re}$, which is a $2n$-dimensional real linear space, can be made into a real Lie algebra by letting

$$[(x_1, x_2), (y_1, y_2)] = ([x_1, y_1] - [x_2, y_2], [x_1, y_2] + [x_2, y_1])$$

And it is these kinds of real Lie algebras that we are searching for in this study. They are called the real Lie subalgebras of a simple complex Lie algebra. We shall remark more on this point later.

For now, however, we let V be a real Lie algebra, i.e., $V = \hat{g}$, where $\hat{g}$ is a real Lie algebra. We form $\hat{g}^c$ and we wish to give it the structure of a complex Lie algebra. Thus we need to define a bracket in $\hat{g}^c$:

$$\hat{g}^c \times \hat{g}^c = (\mathbb{C} \otimes_{\mathbb{R}} \hat{g}) \times (\mathbb{C} \otimes_{\mathbb{R}} \hat{g}) \longrightarrow \mathbb{C} \otimes_{\mathbb{R}} \hat{g} = \hat{g}^c$$
$$(c_1 \otimes u, c_2 \otimes v) \longmapsto [c_1 \otimes u, c_2 \otimes v] := (c_1 c_2) \otimes [u, v]$$

where, of course, $[u, v]$ is the bracket in $\hat{g}$.

We then write the following.

$$[c_1 \otimes u, c_2 \otimes v] = [(a_1 + ib_1) \otimes u, (a_2 + ib_2) \otimes v] = [a_1 \otimes u + (ib_1) \otimes u, a_2 \otimes v + (ib_2) \otimes v]$$

Continuing, we get

$$(c_1 c_2) \otimes [u, v] = ((a_1 + ib_1)(a_2 + ib_2)) \otimes [u, v] =$$
$$((a_1 a_2 - b_1 b_2) + i(a_1 b_2 + b_1 a_2)) \otimes [u, v] =$$
$$(a_1 a_2 - b_1 b_2) \otimes [u, v] + (i(a_1 b_2 + b_1 a_2)) \otimes [u, v]$$

Now

$$\begin{aligned}(a_1a_2)\otimes[u,v] &= [a_1\otimes u, a_2\otimes v]\\ -(b_1b_2)\otimes[u,v] &= -[b_1\otimes u, b_2\otimes v]\\ (a_1b_2)\otimes[u,v] &= [a_1\otimes u, b_2\otimes v]\\ (b_1a_2)\otimes[u,v] &= [b_1\otimes u, a_2\otimes v]\end{aligned}$$

Thus we see that by just using $\mathbb{C}$-linearity in $\hat{g}^c$ we have a natural definition of a Lie bracket in $\hat{g}^c$.

It is straightforward that scalar multiplication is bilinear with respect to this multiplication. For c in $\mathbb{C}$ and $c_1\otimes u$ and $c_2\otimes v$ in $(\mathbb{C}\otimes_{\mathbb{R}}\hat{g})$, we have

$$\begin{gathered}c[c_1\otimes u, c_2\otimes v] = c(c_1c_2\otimes[u,v]) = (c(c_1c_2))\otimes[u,v] =\\ ((cc_1)c_2)\otimes[u,v] = [(cc_1)\otimes u, c_2\otimes v] = [c(c_1\otimes u), c_2\otimes v]\\ c[c_1\otimes u, c_2\otimes v] = c(c_1c_2\otimes[u,v]) = (c(c_1c_2))\otimes[u,v] =\\ (c_1(cc_2))\otimes[u,v] = [c_1\otimes u, (cc_2)\otimes v] = [c_1\otimes u, c(c_2\otimes v)]\end{gathered}$$

We also need to show that this multiplication distributes on the right and on the left, that is, it is bilinear with respect to addition:

$$[c_1\otimes u, c_2\otimes v + c_3\otimes w] = [c_1\otimes u, c_2\otimes v] + [c_1\otimes u), c_3\otimes w]$$

and

$$[c_1\otimes u + c_2\otimes v, c_3\otimes w] = [c_1\otimes u), c_3\otimes w] + [c_2\otimes v, c_3\otimes w]$$

For left distribution we want to show

$$[c_1\otimes u, c_2\otimes v + c_3\otimes w] = [c_1\otimes u, c_2\otimes v] + [c_1\otimes u, c_3\otimes w]$$

We reduce first the righthand side. We work with a basis $(v_1, ..., v_n)$ in $\hat{g}$.

$$\begin{gathered}[c_1\otimes u, c_2\otimes v] + [c_1\otimes u, c_3\otimes w] = c_1c_2\otimes[u,v] + c_1c_3\otimes[u,w] =\\ (a_1+b_1i)(a_2+b_2i)\otimes[\textstyle\sum_{i=1}^n(r_iv_i), \sum_{j=1}^n(s_jv_j)]+\\ (a_1+b_1i)(a_3+b_3i)\otimes[\textstyle\sum_{i=1}^n(r_iv_i), \sum_{k=1}^n(t_kv_k)] =\\ ((a_1a_2-b_1b_2)+(a_1b_2+b_1a_2)i)\otimes\textstyle\sum_{i=1}^n\sum_{j=1}^n r_is_j[v_i,v_j]+\\ ((a_1a_3-b_1b_3)+(a_1b_3+b_1a_3)i)\otimes\textstyle\sum_{i=1}^n\sum_{k=1}^n r_it_k[v_i,v_k] =\\ \textstyle\sum_{i=1}^n\sum_{j=1}^n((a_1a_2-b_1b_2)+(a_1b_2+b_1a_2)i)\otimes r_is_j[v_i,v_j]+\\ \textstyle\sum_{i=1}^n\sum_{k=1}^n((a_1a_3-b_1b_3)+(a_1b_3+b_1a_3)i)\otimes r_it_k[v_i,v_k]\end{gathered}$$

The lefthand side is more delicate. We work first with $c_2\otimes v + c_3\otimes w$.

$$\begin{gathered}c_2\otimes v + c_3\otimes w = (a_2+b_2i)\otimes\textstyle\sum_{j=1}^n(s_jv_j) + (a_3+b_3i)\otimes\sum_{k=1}^n(t_kv_k) =\\ \textstyle\sum_{j=1}^n((a_2+b_2i)\otimes(s_jv_j)) + \sum_{k=1}^n((a_3+b_3i)\otimes(t_kv_k))\end{gathered}$$

Now we have

$$[c_1 \otimes u, (c_2 \otimes v + c_3 \otimes w)] =$$
$$[c_1 \otimes u, \sum_{j=1}^{n}((a_2 + b_2 i) \otimes (s_j v_j)) + \sum_{k=1}^{n}((a_3 + b_3 i) \otimes (t_k v_k))]$$

Once again we see that we have nothing that makes legitimate moving brackets across addition in $\hat{g}^c = (\mathbb{C} \otimes_{\mathbb{R}} \hat{g})$. Again let us *assume* that we can. This gives

$$[c_1 \otimes u, \sum_{j=1}^{n}((a_2 + b_2 i) \otimes (s_j v_j)) + \sum_{k=1}^{n}((a_3 + b_3 i) \otimes (t_k v_k))] =$$
$$[c_1 \otimes u, \sum_{j=1}^{n}((a_2 + b_2 i) \otimes (s_j v_j))] + [c_1 \otimes u, \sum_{k=1}^{n}((a_3 + b_3 i) \otimes (t_k v_k))] =$$
$$\sum_{j=1}^{n}[c_1 \otimes u, ((a_2 + b_2 i) \otimes (s_j v_j))] + \sum_{k=1}^{n}[c_1 \otimes u, ((a_3 + b_3 i) \otimes (t_k v_k))] =$$

Now we can apply the definition of the Lie bracket in $\hat{g}^c = (\mathbb{C} \otimes_{\mathbb{R}} \hat{g})$.

$$\sum_{j=1}^{n}[c_1 \otimes u, ((a_2 + b_2 i) \otimes (s_j v_j))] + \sum_{k=1}^{n}[c_1 \otimes u, ((a_3 + b_3 i) \otimes (t_k v_k))] =$$
$$\sum_{j=1}^{n} c_1(a_2 + b_2 i) \otimes [u, s_j v_j] + \sum_{k=1}^{n} c_1(a_3 + b_3 i) \otimes [u, t_k v_k]$$

We now expand c_1 and u.

$$\sum_{j=1}^{n} c_1(a_2 + b_2 i) \otimes [u, s_j v_j] + \sum_{k=1}^{n} c_1(a_3 + b_3 i) \otimes [u, t_k v_k] =$$
$$\sum_{j=1}^{n}(a_1 + b_1 i)(a_2 + b_2 i) \otimes [\sum_{i=1}^{n}(r_i v_i), s_j v_j] +$$
$$\sum_{k=1}^{n}(a_1 + b_1 i)(a_3 + b_3 i) \otimes [\sum_{i=1}^{n}(r_i v_i), t_k v_k]$$

Now since we know that brackets in $\hat{g}$ are bilinear with respect to addition and real scalars, we have

$$\sum_{j=1}^{n}(a_1 + b_1 i)(a_2 + b_2 i) \otimes [\sum_{i=1}^{n}(r_i v_i), s_j v_j] +$$
$$\sum_{k=1}^{n}(a_1 + b_1 i)(a_3 + b_3 i) \otimes [\sum_{i=1}^{n}(r_i v_i), t_k v_k] =$$
$$\sum_{i=1}^{n} \sum_{j=1}^{n}(a_1 a_2 - b_1 b_2 + (a_1 b_2 + b_1 a_2) i) \otimes r_i s_j [v_i, v_j] +$$
$$\sum_{i=1}^{n} \sum_{k=1}^{n}(a_1 a_3 - b_1 b_3 + (a_1 b_3 + b_1 a_3) i) \otimes r_i t_k [v_i, v_k]$$

Now we wanted to show that

$$[c_1 \otimes u, c_2 \otimes v + c_3 \otimes w] = [c_1 \otimes u, c_2 \otimes v] + [c_1 \otimes u, c_3 \otimes w]$$

We calculated

$$[c_1 \otimes u, c_2 \otimes v] + [c_1 \otimes u, c_3 \otimes w] =$$
$$\sum_{i=1}^{n} \sum_{j=1}^{n}((a_1 a_2 - b_1 b_2) + (a_1 b_2 + b_1 a_2) i) \otimes r_i s_j [v_i, v_j] +$$
$$\sum_{i=1}^{n} \sum_{k=1}^{n}((a_1 a_3 - b_1 b_3) + (a_1 b_3 + b_1 a_3) i) \otimes r_i t_k [v_i, v_k]$$

Assuming that we can move brackets across addition in $\hat{g}^c = (\mathbb{C} \otimes_{\mathbb{R}} \hat{g})$, we obtained

$$
\begin{gathered}
[c_1 \otimes u, (c_2 \otimes v + c_3 \otimes w)] = \\
\sum_{j=1}^{n}(a_1+b_1 i)(a_2+b_2 i) \otimes [\sum_{i=1}^{n}(r_i v_i), s_j v_j] + \\
\sum_{k=1}^{n}(a_1+b_1 i)(a_3+b_3 i) \otimes [\sum_{i=1}^{n}(r_i v_i), t_k v_k] = \\
\sum_{i=1}^{n}\sum_{j=1}^{n}(a_1a_2 - b_1b_2 + (a_1b_2 + b_1a_2)i) \otimes r_i s_j [v_i, v_j] + \\
\sum_{i=1}^{n}\sum_{k=1}^{n}(a_1a_3 - b_1b_3 + (a_1b_3 + b_1a_3)i) \otimes r_i t_k [v_i, v_k]
\end{gathered}
$$

We see that these two expressions are identical. Thus it is *reasonable to define* that brackets are linear with respect to addition in $\hat{g}^c = (\mathbb{C} \otimes_{\mathbb{R}} \hat{g})$. And obviously the same conclusion is true for right distribution.

We also need to show the anticommutativity of the bracket product in $\hat{g}^c$. This is easy since we have this property in $\hat{g}$.

$$
[c_1 \otimes u, c_2 \otimes v] = c_1 c_2 \otimes [u, v] = -c_2 c_1 \otimes [v, u] = -[c_2 \otimes v, c_1 \otimes u]
$$

Finally we need to show that the Jacobi identity holds in $\hat{g}^c$.

$$
\begin{gathered}
[c_1 \otimes u, [c_2 \otimes v, c_3 \otimes w]] + \\
[c_3 \otimes w, [c_1 \otimes u, c_2 \otimes v]] + \\
[c_2 \otimes v, [c_3 \otimes w, c_1 \otimes u]] = \\
[c_1 \otimes u, (c_2 c_3) \otimes [v, w]] + [c_3 \otimes w, (c_1 c_2) \otimes [u, v]] + [c_2 \otimes v, (c_3 c_1) \otimes [w, u]] = \\
(c_1 c_2 c_3) \otimes [u, [v, w]] + (c_3 c_1 c_2) \otimes [w, [u, v]] + (c_2 c_3 c_1) \otimes [v, [w, u]] = \\
(c_1 c_2 c_3) \otimes ([u, [v, w]] + [w, [u, v]] + [v, [w, u]]) = 0
\end{gathered}
$$

since the Jacobi identity is valid in $\hat{g}$.

In light of all the structures created and combinations computed , we can make the following observation. Moving from $\mathbb{C}$ to $\mathbb{R}$, we needed a basis for the $\mathbb{C}$-linear Lie algebra $\hat{g}$ in order to expose the structure of a $\mathbb{R}$-linear Lie algebra $\hat{g}^r$. On the contrary moving from $\mathbb{R}$ to $\mathbb{C}$ by the process of complexification, no such choice was necessary to expose the structure of the $\mathbb{C}$-linear Lie algebra $\hat{g}^c$ from the structure of the $\mathbb{R}$-linear Lie algebra $\hat{g}$.

Thus when we build the complexification of the $\mathbb{C}$-linear Lie algebra $\hat{g}^c$ from the structure of the $\mathbb{R}$-linear Lie algebra $\hat{g}$, we have a canonical way of obtaining it. If the complex dimension of $\hat{g}^c$ is n, then we have immediately the $2n$-real dimensional Lie algebra $(\hat{g}^c)^r = \hat{g} \times \hat{g}$, since we have already identified $\hat{g}$ in $\hat{g}^c$, i.e., we do not need to choose a basis in $\hat{g}^c$ to identify $\hat{g}$ in $\hat{g}^c$. Also for any decomposition of $\hat{g}^c = \hat{g} \oplus i\hat{g}$ we have for $u = u_{re1} + i(0)$ and $v = v_{re1} + i(0)$ in $\hat{g} \oplus i\hat{g}$

$$
\left[\begin{bmatrix} [u_{re1}] \\ [0] \end{bmatrix}, \begin{bmatrix} [v_{re1}] \\ [0] \end{bmatrix}\right] = \begin{bmatrix} [u_{re1}, v_{re1}] - [0, 0] \\ [u_{re1}, 0] + [0, v_{re1}] \end{bmatrix} = \begin{bmatrix} [u_{re1}, v_{re1}] \\ 0 + 0 \end{bmatrix}
$$

and since $\hat{g}$ is a real Lie algebra, we have $[u, v] = [u_{re1} + i(0), v_{re1} + i(0)] = [u_{re1}, v_{re1}] + i(0)$ in $\hat{g} \oplus i\hat{g}$ since $[u_{re1}, v_{re1}]$ is a real number because $\hat{g}$ is a real Lie algebra.

2.12.6 Real Forms.

Now the $\hat{g}_1^c = \hat{g}_1 \oplus i\hat{g}_1$ may not be unique, i.e., there may be another real $\hat{g}_2$ whose complexification $\hat{g}_2^c = \hat{g}_1^c$. Thus, given a complex Lie algebra $\hat{g}$, there may be more than one real Lie algebra whose complexification is $\hat{g}$. These real Lie algebras are called the *real forms* of $\hat{g}$. For instance, we can take $\hat{sl}(n,\mathbb{C})$ and form matrices using only real numbers giving $\hat{sl}(n,\mathbb{R}) \subset \hat{sl}(n,\mathbb{C})$, thus making $\hat{sl}(n,\mathbb{R})$ a real subalgebra of $\hat{sl}(n,\mathbb{C})$. And we know that the complexification of $\hat{sl}(n,\mathbb{R})$ is $\hat{sl}(n,\mathbb{C})$, i.e.,

$$\hat{sl}(n,\mathbb{C}) = \mathbb{C} \otimes \hat{sl}(n,\mathbb{R}) = \hat{sl}(n,\mathbb{R}) \oplus i(\hat{sl}(n,\mathbb{R}))$$

Thus $\hat{sl}(n,\mathbb{R})$ is a real form of $\hat{sl}(n,\mathbb{C})$, called the *split form* of $\hat{sl}(n,\mathbb{C})$. For instance for $n = 2$, $\hat{sl}(2,\mathbb{C})$ are the $2x2$ complex matrices with trace 0; and $\hat{sl}(2,\mathbb{R})$ are the $2x2$ real matrices with trace 0, giving

$$\hat{sl}(2,\mathbb{C}) = \hat{sl}(2,\mathbb{R}) \oplus i(\hat{sl}(2,\mathbb{R}))$$

$$\begin{bmatrix} u & t \\ s & -u \end{bmatrix} = \begin{bmatrix} a_1 & c_1 \\ b_1 & -a_1 \end{bmatrix} + i \begin{bmatrix} a_2 & c_2 \\ b_2 & -a_2 \end{bmatrix}$$

However, we also have

$$\hat{su}_n = \{A \in \hat{sl}_n(\mathbb{C}) | \overline{A^t} = -A\}$$

which, even though the matrices are not real, we will show that it is also a real form of $\hat{sl}(n,\mathbb{C})$.

Again for dimension n= 2 we have

$$A = \begin{bmatrix} u & t \\ s & -u \end{bmatrix} = \begin{bmatrix} a_1 + ia_2 & c_1 + ic_2 \\ b_1 + ib_2 & -a_1 - ia_2 \end{bmatrix}$$

$$A^t = \begin{bmatrix} a_1 + ia_2 & b_1 + ib_2 \\ c_1 + ic_2 & -a_1 - ia_2 \end{bmatrix} \qquad \overline{A^t} = \begin{bmatrix} a_1 - ia_2 & b_1 - ib_2 \\ c_1 - ic_2 & -a_1 + ia_2 \end{bmatrix}$$

Now

$$-A = \begin{bmatrix} -a_1 - ia_2 & -c_1 - ic_2 \\ -b_1 - ib_2 & a_1 + ia_2 \end{bmatrix}$$

Thus we see that $\overline{A^t} \neq -A$, and thus A is not in $\hat{su}_n$. But we now let A in $\hat{sl}_n(\mathbb{R})$ have the form

$$A = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}$$

then iA is in $\hat{su}_n$:

$$iA = \begin{bmatrix} ia & ib \\ ib & -ia \end{bmatrix}$$

Now

$$(iA)^t = \begin{bmatrix} ia & ib \\ ib & -ia \end{bmatrix}$$

and

$$\overline{(iA)^t} = \begin{bmatrix} -ia & -ib \\ -ib & ia \end{bmatrix}$$

Now

$$-iA = \begin{bmatrix} -ia & -ib \\ -ib & ia \end{bmatrix}$$

and thus we see that $\overline{(iA)^t} = -iA$, which says that iA is in $\hat{su}_n$. This gives

$$-iA + i(-iA) = -iA + A =$$

$$\begin{bmatrix} -ia & -ib \\ -ib & ia \end{bmatrix} + \begin{bmatrix} a & b \\ b & -a \end{bmatrix} = \begin{bmatrix} -ia + a & -ib + b \\ -ib + b & ia - a \end{bmatrix}$$

and we can conclude that

$$\begin{bmatrix} -ia + a & -ib + b \\ -ib + b & ia - a \end{bmatrix}$$

is in $\hat{sl}_n(\mathbb{C})$.

Thus when we begin with a $\mathbb{C}$ Lie algebra $\hat{g}$ and ask what real structures it determines, we are actually in the context of determining the real forms of the complex Lie algebra $\hat{g}$. This pursuit is a major endeavor in the study of Lie algebras and leads to some beautiful mathematics, but we will pursue it no further in these pages.

2.12.7 Changing Fields Preserves Solvability. Recall that we took this long detour through the topic of "Changing Scalar Fields" in order to address the proofs of theorems such as

> A Lie algebra $\hat{g}$ over $\mathbb{R}$ or $\mathbb{C}$ is solvable if and only if the Killing form $B(x,x) = 0$ for all x in $D^1\hat{g}$

The reason for doings so is the fact that we will need below the following result to achieve our proof:

> If a Lie algebra $\hat{s}$ over $\mathbb{R}$ is solvable, then its complexification $\hat{s}^c = \mathbb{C}^r \otimes_{\mathbb{R}} \hat{s}$ is solvable; and if a Lie algebra $\hat{s}$ over $\mathbb{C}$ is solvable and if $\hat{s} = \hat{s}_{re} \oplus i\hat{s}_{re}$ and $\hat{s}^r = s_{re} \times s_{re}$, where $\hat{s}_{re}$ is a real Lie algebra, then the real Lie algebra $\hat{s}^r$ and its subalgebra s_{re} are solvable.

First we make the following remarks. Recall that $D^1\hat{s}^c$ is the linear space generated by brackets in $\hat{s}^c$. Let $c \otimes v$ be one summand in $D^1\hat{s}^c$, where c is in $\mathbb{C}$ and v is in $\hat{s}$. Thus there exists a $c_1 \otimes v_1$ and a $c_2 \otimes v_2$ in $\hat{s}^c$ such that $[c_1 \otimes v_1, c_2 \otimes v_2] = c \otimes v$, where c_1 and c_2 are in $\mathbb{C}$ and v_1 and v_2 are in $\hat{s}$. But $[c_1 \otimes v_1, c_2 \otimes v_2] = c_1c_2 \otimes [v_1, v_2]$. Thus $c = c_1c_2$ and $v = [v_1, v_2]$, and we can conclude that every summand in $D^1\hat{s}^c$ comes from an element in $D^1\hat{s}$ by complexification. Likewise, by induction, we can assert that every summand in $D^k\hat{s}^c$ comes from an element v in $D^k\hat{s}$ by complexification.

On the other hand (given $\hat{s}$, a complex Lie algebra such that the subalgebras s_{re} and $\hat{s}^r$ are real Lie algebras) and after having chosen a decomposition $\hat{s} = \hat{s}_{re} \oplus i\hat{s}_{re}$, we know that $D^1\hat{s}^r \subset \hat{s}_{re} \times \hat{s}_{re}$ is generated by brackets in $\hat{s}_{re} \times \hat{s}_{re}$. Thus each summand comes from an element in $D^1\hat{s}^r \subset \hat{s}_{re} \times \hat{s}_{re}$. Let (u_1, u_2) be a summand in $D^1\hat{s}^r \subset \hat{s}_{re} \times \hat{s}_{re}$. Then there exist (x_1, x_2) and (y_1, y_2) in $\hat{s}_{re} \times \hat{s}_{re}$ such that

$$[(x_1, x_2), (y_1, y_2)] =$$
$$([x_1, y_1] - [x_2, y_2], [x_1, y_2] + [x_2, y_1]) = (u_1, u_2)$$

To see this we let u be in $\hat{s}$, giving $u = u_1 + iu_2$ with u_1 and u_2 in $\hat{s}_{re}$. Now let $x = x_1 + ix_2$ and $y = y_1 + iy_2$. They are in $\hat{s}$. Then we have

$$[x, y] = [x_1 + ix_2, y_1 + iy_2] =$$
$$([x_1, y_1] - [x_2, y_2]) + i([x_1, y_2] + [x_2, y_1]) = u_1 + iu_2 = u$$

Thus any summand (u_1, u_2) in $D^1\hat{s}^r$ comes from an element u in $D^1\hat{s}$. Likewise, by induction, we can assert that any summand in $D^k\hat{s}^r$ comes from an element in $D^k\hat{s}$.

Now let us assume that the real Lie algebra $\hat{s}$ is solvable. This means for some k, $D^{k-1}\hat{s} \neq 0$ and $D^k\hat{s} = 0$. We complexify $\hat{s}$ and obtain $\hat{s}^c$. Now we know that any summand $c \otimes v$ in $D^k\hat{s}^c$ comes from an element v in $D^k\hat{s}$. But $D^k\hat{s} = 0$, and thus $D^k\hat{s}^c$ is also equal to 0 and this says that the complex Lie algebra $\hat{s}^c$ is solvable. We now assume that the complex Lie algebra $\hat{s}$ such that the subalgebras s_{re} and $\hat{s}^r$ are real Lie algebras is solvable. This means that for some k, $D^{k-1}\hat{s} \neq 0$ and $D^k\hat{s} = 0$. We choose a decomposition of $\hat{s} = \hat{s}_{re} \oplus i\hat{s}_{re}$. Now any summand of $D^k\hat{s}^r$ comes from an element in $D^k\hat{s}$. But $D^k\hat{s} = 0$. Thus we conclude that $D^k\hat{s}^r = 0$ and that the real Lie algebra $\hat{s}^r$ is solvable. Since $\hat{s}_{re}$ is a subalgebra of $\hat{s}^r$, we can also conclude that the real Lie algebra $\hat{s}_{re}$ is solvable.

Since this proof just uses the properties of the Lie bracket, it is obvious that if the real Lie algebra $\hat{n}$ is a nilpotent Lie algebra, then $\hat{n}^c$ is a nilpotent complex Lie algebra; and if the complex Lie algebra $\hat{n}$ such that the subalgebras n_{re} and $\hat{n}^r$ are real Lie algebras is a nilpotent Lie algebra, then $\hat{n}^r$ and $\hat{n}_{re}$ are nilpotent real Lie algebras for any decomposition of $\hat{n} = \hat{n}_{re} \oplus i\hat{n}_{re}$.

2.12.8 Solvable Lie Algebra ŝ implies $\mathbf{D^1}$ŝ is a Nilpotent Lie Algebra. We already know this statement is true for a Lie Algebra $\hat{s}$ over $\mathbb{C}$ (see 2.11.2). Thus we start now with a solvable Lie algebra $\hat{s}$ over $\mathbb{R}$. From the above we know that $\hat{s}^c$ is a solvable Lie algebra over $\mathbb{C}$. Thus $D^1\hat{s}^c$ is a nilpotent Lie algebra. But again from the above we know that $D^1\hat{s}$ is a nilpotent Lie algebra over $\mathbb{R}$. And thus we have our conclusion that for any solvable Lie algebra over $\mathbb{F}$, where $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$, we know that $D^1\hat{s}$ is a nilpotent Lie algebra.

2.13 The Killing Form (2)

We are now ready to give the proof of

> A Lie algebra $\hat{g}$ over $\mathbb{R}$ or $\mathbb{C}$ is solvable if and only if the Killing form $B(x,x) = 0$ for all x in $D^1\hat{g}$.

2.13.1 From Solvability to the Killing Form. The easy direction of this proof is that if $\hat{g}$ is solvable over $\mathbb{R}$ or $\mathbb{C}$, then the Killing form $B(x,x) = 0$ for all x in $D^1\hat{g}$. Note that for either field $\mathbb{R}$ or $\mathbb{C}$, the definition of the Killing form is the same.

$$B(x,y) = trace(ad(x) \circ ad(y))$$

Since $\hat{g}$ is a solvable Lie algebra, we have shown that $D^1\hat{g}$ is nilpotent. Thus we know that $ad(x)$ is a nilpotent linear transformation in $\widehat{gl}(\hat{g})$ for every $x \in D^1\hat{g}$ (see 2.7.1). Also, we know that the composition of a nilpotent linear transformations is still nilpotent and that the trace of the nilpotent linear transformations with respect to the basis found by Engel's Theorem is zero, since these matrices are upper triangular with a zero diagonal. Now since the trace is independent of the basis used to calculate it, we can conclude that $B(x,y) = 0$ for all x, y in $D^1\hat{g}$. Thus the weaker conclusion is also true, $B(x,x) = 0$ for all x in $D^1\hat{g}$.

At this point we would like to expose another technique using complexification. We would like to show how we can arrive at the above conclusion for a Lie algebra over $\mathbb{R}$, a field of characteristic zero which is not algebraically closed, but now on the assumption that the conclusion is true for a Lie algebra over $\mathbb{C}$. Thus If $\hat{g}$ is a Lie algebra over $\mathbb{R}$, and it is solvable, we would like to conclude again that $B(x,x) = 0$ for all x in $D^1\hat{g}$. But we know that if $\hat{g}$ is solvable as a real Lie Algebra, then its complexification $\hat{g}^c$ is also solvable (see 2.12.7). Thus, by assumption, we can apply the theorem to $\hat{g}^c$. We choose and fix a summand x in $D^1\hat{g}$. Complexifying $\hat{g}$, we obtain $\hat{g}^c = \hat{g} \oplus i\hat{g}$, and we know that for any $c \neq 0$, $cx = c \otimes x$ is in $D^1\hat{g}^c$. [To distinguish the Killing form in $\hat{g}^c$ and in $\hat{g}$, we use $B^{\mathbb{C}}$ and $B^{\mathbb{R}}$ respectively.] Using the theorem we have

$$B^{\mathbb{C}}(cx, cx) = B^{\mathbb{C}}((a+ib)x, (a+ib)x) = ((a^2 - b^2) + i(ab + ba))B^{\mathbb{C}}(x,x) = 0$$

But we observe that if we restrict $B^{\mathbb{C}}$ in $\hat{g}^c$ to $\hat{g} \subset \hat{g}^c$, we obtain $B^{\mathbb{R}}$. We have for x, y in $\hat{g}$, $B^{\mathbb{C}}(x,y) = trace(ad(x)ad(y))$, and we see only $ad(\hat{g})$ appearing. Thus we can conclude that in this case only $B^{\mathbb{C}}(x,y) = B^{\mathbb{R}}(x.y)$. Since $c \neq 0$,

$$B^{\mathbb{C}}(cx, cx) = ((a^2 - b^2) + i(ab + ba))B^{\mathbb{R}}(x,x) = 0$$

giving $B^{\mathbb{R}}(x,x) = 0$ for any x in $D^1\hat{g}$.

2.13.2. From the Killing Form to Solvability over $\mathbb{C}$. We now want to prove the converse, that if the Killing form $B(x,x) = 0$ for all x in $D^1\hat{g}$ in a Lie algebra $\hat{g}$ over $\mathbb{R}$ or $\mathbb{C}$, then $\hat{g}$ is solvable. We first treat the case for the scalar field $\mathbb{C}$.

We first observe that we can assume that $D^1\hat{g} \neq 0$. Since if $D^1\hat{g} = 0$, then $B(x,x) = 0$ is immediately satisfied and this means that $\hat{g}$ is abelian, and thus $\hat{g}$ is solvable and the theorem is verified.

Next we want to remark that $B(x,x) = 0$ for all x in a Lie algebra $\hat{g}$ is equivalent to $B(x,y) = 0$ for all x and y in that Lie algebra $\hat{g}$. [In 2.11.1, where we were treating a similar result for $\hat{B}$, we saw that this property just depended on the symmetry of the form $\hat{B}$ and the fact that our field was not of characteristic 2. Since the same conditions apply to B, the equivalence is true here as well.]

Of course, we will use our *Theorem* $\hat{B}$ (cf 2.11.1). In order to avoid ambiguities for the Lie algebra in the Theorem we will use the symbol $\hat{g}'$, which, we recall, is a subalgebra of the Lie algebra $\widehat{gl}(V)$ for some complex linear space V.] *Theorem* $\hat{B}$ then becomes:

Let the form $\hat{B}$ satisfy the condition that for all X in $D^1\hat{g}'$ and all Y in $\hat{g}'$, $\hat{B}(X,Y) = 0$. Then X is a nilpotent linear transformation.

Our hypothesis now is that $B(x,x) = 0$ for all x in $D^1\hat{g}$, where $\hat{g}$ is a Lie algebra over $\mathbb{C}$. This means by the definition of B that $tr(ad(x), ad(x)) = 0$ with $ad(x)$ in $ad(D^1\hat{g}) \subset \widehat{gl}(\hat{g})$. However this is the same as $\hat{B}(ad(x), ad(x)) = 0$ for $\hat{B}$ defined on the subalgebra $ad(D^1\hat{g})$ of the Lie algebra $\widehat{gl}(\hat{g})$ for the complex Lie algebra $\hat{g}$ as a complex linear space. We now take y in $D^2\hat{g}$, and note that $ad(y)$ is in $ad(D^2\hat{g})$. We have

$$ad(D^2\hat{g}) = ad([D^1\hat{g}, D^1\hat{g}]) \subset [ad(D^1\hat{g}), ad(D^1\hat{g})] = D^1(ad(D^1\hat{g}))$$

Now we have $ad(y)$ in $D^1(ad(D^1\hat{g}))$ and $ad(x)$ in $ad(D^1\hat{g})$ and thus we can conclude that $\hat{B}(ad(y), ad(x)) = 0$, which is the same as $B(y,x) = 0$, for all x in $D^1\hat{g}$ and all y in $D^2\hat{g}$. And we have reached the conclusion that $ad(y)$ is a nilpotent linear transformation for all y in $D^2\hat{g}$. Now 2.8.2 says that in this case $D^2\hat{g}$ is a nilpotent Lie algebra. And we know that nilpotent Lie algebras are also solvable (see 2.5.1). This says then that $D^k(D^2\hat{g}) = 0$ for some k. Thus $D^{k+2}\hat{g} = D^k(D^2\hat{g}) = 0$, giving us our conclusion that $\hat{g}$ is solvable. And therefore we can assert that

A Lie algebra $\hat{g}$ over $\mathbb{C}$ is solvable if the Killing form $B(x,x) = 0$ for all x in $D^1\hat{g}$.

2.13.3 From the Killing Form to Solvability over $\mathbb{R}$. To complete this part of the exposition we would like to prove that for a Lie algebra $\hat{g}$ over $\mathbb{R}$, if the Killing form $B(x,x) = 0$ for each x in $D^1\hat{g}$, then the Lie algebra is solvable. Now in order to prove this we should first complexify the real Lie algebra $\hat{g}$, apply the theorem just proven to this case, and then move back down to the real Lie algebra.

Thus, we assume that for a Lie algebra $\hat{g}$ over $\mathbb{R}$ the Killing form $B^{\mathbb{R}}(x,x) = 0$ for each x in $D^1\hat{g}$. We complexify $\hat{g}$ to $\hat{g}^c$ and consequently we know that any z in $D^1\hat{g}^c$ can be written as $z = cx$ for some x in $\hat{g}$ and some c in $\mathbb{C}$. Now z is a linear combination of brackets $z_{ij} = [z_i, z_j]$ for z_i and z_j in $\hat{g}^c$. But $z_i = c_i x_i$ for some c_i in $\mathbb{C}$ and some x_i in $\hat{g}$; and likewise for z_j. Thus $[z_i, z_j] = [c_i x_i, c_j x_j] = c_i c_j [x_i, x_j] = c_{ij} x_{ij}$. We can conclude that x_{ij} is in $D^1\hat{g}$. Calculating the Killing form $B^{\mathbb{C}}$ in $\hat{g}^c$, we obtain

$$B^{\mathbb{C}}(c_{ij}x_{ij}, c_{ij}x_{ij}) = B^{\mathbb{C}}((a_{ij} + ib_{ij})x_{ij}, ((a_{ij} + ib_{ij})x_{ij}) = \\ ((a_{ij}^2 - b_{ij}^2) + i(a_{ij}b_{ij} + b_{ij}a_{ij}))B^{\mathbb{R}}(x_{ij}, x_{ij})$$

But by hypothesis $B^{\mathbb{R}}((x_{ij}, x_{ij})) = 0$, which gives the desired conclusion that $B^{\mathbb{C}}(c_{ij}x_{ij}, c_{ij}x_{ij}) = B^{\mathbb{C}}(z_{ij}, z_{ij}) = 0$. But z is a linear combination of the z_{ij}, and since $B^{\mathbb{C}}$ is bilinear, we can conclude that $B^{\mathbb{C}}(z,z) = 0$ for z in $D^1\hat{g}^c$. Thus we know that $\hat{g}^c$ is solvable. And we have already shown that this means that $\hat{g}$ is solvable. (See 2.12.7.)

2.14 Some Remarks on Semisimple Lie Algebras (3)

Recall that, after defining the Killing form of a Lie algebra, we affirmed that with this new tool we could prove two remarkable theorems:

> A Lie algebra $\hat{g}$ over $\mathbb{R}$ or $\mathbb{C}$ is solvable if and only if the Killing form $B(x,x) = 0$ for all x in $D^1\hat{g}$.
>
> A Lie algebra $\hat{g}$ over $\mathbb{R}$ or $\mathbb{C}$ is semisimple if and only if its Killing form B is nondegenerate.

2.14.1 Non-Trivial Solvable Ideal Implies the Degeneracy of the Killing Form. We have just given some beautiful proofs for solvable Lie algebras. Now we wish to give the proofs for the semisimple Lie algebras. But rather than using the defining property of semisimple Lie algebras — a semisimple Lie algebra has a trivial radical — we translate this property over to its equivalent one using the Killing form (see 2.11.2), which property is a rather powerful expression for semisimplicity. The bridge among these ideas is the fact that every bilinear form B on a linear space V, e.g. the Killing form, determines a linear map $\mathcal{B}$ between V and its dual V^*. In our case we have [where the scalar field for $\hat{g}$ is indicated by $\mathbb{F}$] that

$$\begin{aligned} &\hat{g} \xrightarrow{\mathcal{B}} \hat{g}^* \\ &u \longmapsto \mathcal{B}(u) : \hat{g} \longrightarrow \mathbb{F} \\ &\qquad\qquad\qquad v \longmapsto \mathcal{B}(u)(v) := B(u,v) \end{aligned}$$

If there is a nonzero element u in the kernel $\hat{k}$ of this map $\mathcal{B}$, this element has the property that for all v in $\hat{g}$, $B(u,v) = 0$. Thus any such u gives the zero map in $\hat{g}^*$, and, of course, this means that there is a degeneracy in the $\mathcal{B}$ map. For non-degeneracy it is demanded that the only zero map in $\hat{g}^*$ comes by $\mathcal{B}$ from the zero element in $\hat{g}$. In other words non-degeneracy demands that the kernel $\hat{k}$ of the map $\mathcal{B}$ be zero. We also note that, in general, the kernel $\hat{k}$ is a Lie subalgebra. To show this we let u_1 and u_2 be in the kernel $\hat{k}$. We show that $[u_1,u_2]$ is also in the kernel. For by associativity of the Killing form, we have for any v in $\hat{g}$, $\mathcal{B}([u_1,u_2])(v) = B([u_1,u_2],v) = B(u_1,[u_2,v]) = \mathcal{B}(u_1)([u_2,v]) = 0$ since u_1 is in the kernel. Thus we can conclude that $[u_1,u_2]$ is also in the kernel, and thus we have that $\hat{k}$ a subalgebra of $\hat{g}$. We remark that since associativity of the Killing form is a consequence of the Jacobi identity, we see that $\hat{k}$ being a subalgebra reflects the structure of the Lie algebra $\hat{g}$.

Now the condition that we placed on a non-trivial solvable Lie algebra $\hat{s}$ — namely that for every u in $D^1\hat{s}$ we have $B(u,u) = 0$ — that condition implies that a degeneracy exists in the map $\mathcal{B}$. We remark that since $\hat{s}$ is just a part of $\hat{g}$, we need a way of extending this condition on $\hat{s}$ to all of $\hat{g}$. This can be done since in our situation $\hat{s}$ is a solvable ideal in $\hat{g}$ and the Killing form is associative. Here is how. For any v in $\hat{g}$ and any u_1 and u_2 in $D^1\hat{s}$, consider $B(v,[u_1,u_2])$. Associativity of B gives $B(v,[u_1,u_2]) = B([v,u_1],u_2)$. Since $D^1\hat{s}$ is an ideal in $\hat{g}$, this means $[v,u_1]$ is in $D^1\hat{s}$. Our condition on B says that $B([v,u_1],u_2) = 0$. [Recall that $B(u,u) = 0$ implies $B(u,v) = 0$ for all u and v in $\hat{g}$.] Thus $B(v,[u_1,u_2]) = 0$, which says $[u_1,u_2]$ is in the kernel of the map $\mathcal{B}$. We can conclude that if $D^1\hat{s}$ is not abelian, then $\mathcal{B}$ has a degeneracy.

If, however, $D^1\hat{s}$ is abelian, we return to the definition of B to affirm that we have a degeneracy also in this case. We take v in $\hat{g}$ and u_1 in $D^1\hat{s}$. Then $B(v,u_1) = trace(ad(v)ad(u_1))$. We choose a basis (a_i) for $D^1\hat{s}$, and a complementary basis (b_i) for a complementary subspace $\hat{d}$ such that $\hat{g} = D^1\hat{s} \oplus \hat{d}$. For any w in $\hat{g}$, we write $w = a + b$, where a is in $D^1\hat{s}$ and b is in $\hat{d}$. Now $ad(u_1)\cdot w = ad(u_1)\cdot(a+b) = ad(u_1)\cdot a + ad(u_1)\cdot b = [u_1,a] + [u_1,b]$. Since u_1 and a are in $D^1\hat{s}$, which is abelian, then $[u_1,a] = 0$. Also since u_1 is in the ideal $D^1\hat{s}$, then $[u_1,b]$ is in $D^1\hat{s}$. Thus the matrix for $ad(u_1)$ written with respect to the above basis takes the form

$$ad(u_1) = \begin{bmatrix} 0 & A_{12} \\ 0 & 0 \end{bmatrix}$$

We see that $ad(u_1)\cdot w = [u_1,w]$ and therefore gives an element u_2 in $D^1\hat{s}$, which fact, of course, is also a conclusion from the ideal structure of $D^1\hat{s}$ in $\hat{g}$. Now for arbitrary matrix B representing $ad(v)$, we have for an arbitrary u in $D^1\hat{s}$ that $ad(v)\cdot u = [v,u]$, which gives again an element in $D^1\hat{s}$ because of the ideal structure of $D^1\hat{s}$. Thus with respect to the above chosen basis, we obtain:

$$ad(v) = \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{bmatrix}$$

Thus the matrix representing $ad(v)ad(u_1)$ is the following

$$ad(v)ad(u_1) = \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{bmatrix} \begin{bmatrix} 0 & A_{12} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & B_{11}A_{12} \\ 0 & 0 \end{bmatrix}$$

which obviously has trace zero. Thus if $D^1\hat{s}$ is abelian, we again have a degeneracy for $\mathcal{B}$. where the scalar field [But this last proof exposes the structure. We know that a non-trivial $\hat{s}$, as a solvable Lie algebra, has a non-zero abelian ideal (which may be $\hat{s}$ itself). Thus we can repeat the proof given above, using this abelian ideal instead of $D^1\hat{s}$, and prove immediately that $\mathcal{B}$ has a degeneracy on $\hat{g}$ – without analyzing the condition of B on $D^1\hat{s}$. This means that the condition on $D^1\hat{s}$ is really superfluous. What matters is the abelian character of the radical. And we will see in the following pages how the fact that $\hat{s}$ itself is abelian, and more particularly when it is the center, demands special attention in the important parts of the proofs of this theory.]

2.14.2 Semisimple Implies Non-Degeneracy of the Killing Form. We now are in a position to prove that

> A Lie algebra $\hat{g}$ over $\mathbb{R}$ or $\mathbb{C}$ is semisimple if and only if its Killing form B is nondegenerate.

We first prove that a semisimple Lie algebra $\hat{g}$ has a non-degenerate Killing form. This means we are affirming that the kernel $\hat{k}$ of the map $\mathcal{B}$ contains only the zero element of $\hat{g}$. Let us suppose that this kernel is not trivial. First , as was proved in general above, we observe that $\hat{k}$ is an ideal in $\hat{g}$. Recall the proof. This fact follows from the associativity of the Killing form. We wish to show that $[\hat{k}, \hat{g}]$ is contained in $\hat{k}$. Let u be in $\hat{k}$ and v be in $\hat{g}$. Now for any w in $\hat{g}$, we have $B([u,v],w) = B(u,[v,w])$. Since u is in $\hat{k}$, this says that $B(u,[v,w]) = 0$. Thus $B([u,v],w) = 0$ for all w in $\hat{g}$, which says that $[u,v]$ is in the kernel of the map $\mathcal{B}$. We conclude that $\hat{k}$ is an ideal in $\hat{g}$. We now show that this ideal is a solvable ideal. We restrict the Killing form now to $\hat{k}$. Let u be in $D^1\hat{k}$. [Clearly, if $D^1\hat{k} = 0$, $\hat{k}$ is solvable.] Then $B(u,u) = \mathcal{B}(u)(u)$. But u is in the kernel of $\mathcal{B}$, giving $B(u,u) = 0$ for all u in $D^1\hat{k}$. Thus we know that $\hat{k}$ is solvable, and that it is also a solvable ideal in $\hat{g}$. But we have assumed that $\hat{g}$ is semisimple. Thus we can conclude that $\hat{k} = 0$. But this is the same as asserting that map $\mathcal{B}$ is an isomorphism, which fact is the definition of non-degeneracy for the Killing form B.

2.14.3 Non-Degeneracy of the Killing Form Implies Semisimple. Now we prove that if a Lie algebra $\hat{g}$ has a non-degenerate Killing form, then it is semisimple. Again suppose that $\hat{g}$ is not semisimple. Then it has a non-trivial solvable ideal $\hat{s}$. Above we showed that this information implied that B was degenerate on $\hat{g}$. But we have assumed that B is non-degenerate, and thus we can conclude that the ideal $\hat{s} = 0$, which means that $\hat{g}$ is semisimple.

2.14.4 A Semisimple Lie Algebra is a Direct Sum of Simple Lie Algebras. We now have arrived at the point to which we have been heading. Assuming the Levi decomposition theorem, we know that any Lie algebra $\hat{g}$

can be written as a direct sum of linear spaces $\hat{g} = \hat{k} \oplus \hat{r}$, where $\hat{k}$ is semisimple ideal and $\hat{r}$ is the radical of $\hat{g}$. Now we want to affirm that any semisimple Lie algebra $\hat{k}$ can be written as a direct sum of simple Lie algebras which are ideals in $\hat{g}$. From this we can affirm the structure theorem for any Lie algebra $\hat{g}$ over $\mathbb{R}$ or $\mathbb{C}$, namely that

$$\hat{g} = \hat{a}_1 \oplus \hat{a}_2 \oplus \cdots \hat{a}_l \oplus \hat{r}$$

where each $\hat{a}_i$ is a simple Lie algebra and $\hat{r}$ is the radical of $\hat{g}$ and where this decomposition is unique up to the order.

To prove this theorem we take advantage of the beautiful criterion for semisimplicity — non-degeneracy of the Killing form. Thus, suppose that $\hat{l}$ is any semisimple Lie algebra which is not simple. Thus it has a proper ideal $\hat{a}$ which is also semisimple. Now we define $\hat{a}^{\perp}$ to be all the elements of $\hat{l}$ which are perpendicular to $\hat{a}$ with respect to the Killing form B on $\hat{l}$, i.e., v in $\hat{l}$ is in $\hat{a}^{\perp}$ if $B(v,u) = 0$ for all u in $\hat{a}$. It is obvious that $\hat{a}^{\perp}$ is a linear subspace of $\hat{l}$. It is also a ideal in $\hat{l}$ and here is the proof. Let v be in $\hat{a}^{\perp}$, w any element in $\hat{l}$, and u in $\hat{a}$. Then $B([v,w],u) = B(v,[w,u])$. Since $\hat{a}$ is an ideal in $\hat{l}$, $[w,u]$ is in $\hat{a}$. Thus $B(v,[w,u]) = 0$, giving $[v,w]$ in $\hat{a}^{\perp}$. We conclude that $\hat{a}^{\perp}$ is an ideal in $\hat{l}$. We also know that the intersection of any two ideals is an ideal. We conclude that $\hat{a} \cap \hat{a}^{\perp}$ is also an ideal in $\hat{l}$. But we remark that this ideal is an abelian ideal in $\hat{l}$. To show this we take two elements w_1 and w_2 in $\hat{a} \cap \hat{a}^{\perp}$ and we calculate $[w_1, w_2]$. Now for any element w in $\hat{l}$, we have $B([w_1, w_2], w) = B(w_1, [w_2, w])$. Since w_2 is in $\hat{a}$, which is an ideal in $\hat{l}$, we know that $[w_2, w]$ is in $\hat{a}$. But w_1 is in $\hat{a}^{\perp}$. Thus $B(w_1, [w_2, w]) = 0$, giving $B([w_1, w_2], w) = 0$. Since w is arbitrary in $\hat{l}$, and since B is nondegenerate, this means that $[w_1, w_2]$ lies in the zero kernel of the map $\mathcal{B}$, giving us our desired conclusion that $[w_1, w_2] = 0$ and that $\hat{a} \cap \hat{a}^{\perp}$ is abelian. Thus $\hat{l}$ has an abelian ideal, which of course, is solvable. But $\hat{l}$ is semisimple, making $\hat{a} \cap \hat{a}^{\perp} = 0$.

To continue, we need to use one of the dimension theorems of linear algebra:

$$\dim(\hat{a}) + \dim(\hat{a}^{\perp}) = \dim(\hat{l}) + \dim(\hat{a} \cap \hat{a}^{\perp})$$

Now since $\hat{a} \cap \hat{a}^{\perp} = 0$, we have that:

$$\hat{a} \oplus \hat{a}^{\perp} = \hat{l}$$

We remark that this also shows $[\hat{a}, \hat{a}^{\perp}] = 0$ since both $\hat{a}$ and $\hat{a}^{\perp}$ are ideals, then $[\hat{a}, \hat{a}^{\perp}]$ is contained in $\hat{a}$ and also in $\hat{a}^{\perp}$, which says that $[\hat{a}, \hat{a}^{\perp}] = 0$.

Finally, we show that $\hat{a}^\perp$ is semisimple. We do this by using the nondegeneracy of the Killing form B. We begin by choosing $(u^\perp)_1$ in $\hat{a}^\perp$ such that $B((u^\perp)_1, u^\perp) = 0$ for an arbitrary $u^\perp$ in $\hat{a}^\perp$. We choose an arbitrary element $v = u + u^\perp$ in $\hat{l}$, where u is in $\hat{a}$. Now

$$B((u^\perp)_1, v) = B((u^\perp)_1, u + u^\perp) = B((u^\perp)_1, u) + B((u^\perp)_1, u^\perp).$$

Since $(u^\perp)_1$ is in $\hat{a}^\perp$ and u is in $\hat{a}$, we have $B((u^\perp)_1, u) = 0$. By hypothesis $B((u^\perp)_1, u^\perp) = 0$. Thus we conclude that $B((u^\perp)_1, v) = 0$. Since $\hat{l}$ is semisimple, this means B on $\hat{l}$ is nondegenerate, making $(u^\perp)_1 = 0$. We now conclude that B on $\hat{a}^\perp$ is nondegenerate, which makes $\hat{a}^\perp$ semisimple.

We can continue in this manner. If $\hat{a}$ has a proper ideal, we can write it also as a direct sum of semisimple ideals $\hat{a} = \hat{a}_1 \oplus \hat{a}_2$. Since these are proper ideals, their dimensions are decreasing, and we must finally arrive at a $\hat{a}_k$ which has no proper ideals. But an a_k that is semisimple ideal with no proper ideals is simple. Thus we can conclude that any semisimple Lie algebra $\hat{l}$ can be written as

$$\hat{l} = \hat{a}_1 \oplus \hat{a}_2 \oplus \cdots \oplus \hat{a}_l$$

where each $\hat{a}_i$ is a simple Lie algebra. [Indeed it can be proven that this decomposition is unique up to the order, and that each summand is uniquely determined, and not only up to isomorphism. But we shall not do so here.]

2.15 The Casimir Operator and the Complete Reducibility of a Representation of a Semisimple Lie Algebra

Before we can give a proof of the Levi Decomposition Theorem, we need the fact of the complete reducibility of a representation of a semisimple Lie algebra. This is one of the deepest results in the theory and its proof is not trivial. The complete reducibility of a representation of a semisimple Lie algebra refers to the following theorem:

> Let V be a representation of a semisimple Lie algebra $\hat{g}$, and let W be an invariant subspace of $\hat{g}$. Then there exists a subspace W' of V invariant by $\hat{g}$ which is complementary.

[Put in other words: complete reducibility of a representation of $\hat{g}$ in a representation space V refers to the fact that given an invariant subspace W of $\hat{g}$ in the representation space V, there exists an invariant subspace W' complementary to W, such that, $V = W \oplus W'$.]

In a matrix representation of $\hat{g}$ acting on V, the invariance of the subspace W means that all the matrices would take the block form

$$\begin{bmatrix} * & * \\ 0 & * \end{bmatrix}$$

What we are affirming is that we can find a basis of V such that all the matrices of the representation take the block form

$$\begin{bmatrix} * & 0 \\ 0 & * \end{bmatrix}$$

The Casimir operator allows us to begin the process of finding such invariant subspaces. But in order to define the Casimir operator we need first to return to the form $\hat{B}$.

2.15.1 The Killing Form Defined on $\widehat{\mathbf{gl}}(\mathbf{V})$. Recall that in 2.11.1 we stated that the form $\hat{B}$ is a bilinear form defined over the Lie algebra $\hat{g}$ [considered as a linear space], which is a Lie subalgebra of $\widehat{gl}(V)$, where V is a linear space over the field $\mathbb{C}$ or $\mathbb{R}$. Choosing a basis for V, we defined form $\hat{B}$ as

$$\begin{aligned} \hat{B} : \hat{g} \times \hat{g} &\longrightarrow \mathbb{F} \\ (X, Y) &\longmapsto \hat{B}(X, Y) := trace(X \circ Y) \end{aligned}$$

and we proved the difficult and beautiful *Theorem* $\hat{B}$ which says that when V is a $\mathbb{C}$-linear space:

> Let the form $\hat{B}$ satisfy the condition that for all X in $D^1\hat{g}$ and all Y in $\hat{g}$, $\hat{B}(X, Y) = 0$. Then X is a nilpotent linear transformation.

Using the form $\hat{B}$, we then defined the traditional *Killing form* B. It was defined as a bilinear form over a Lie algebra $\hat{g}$ [considered as a linear space] over the field $\mathbb{C}$ or $\mathbb{R}$, by using the adjoint map ad from $\hat{g}$ into the Lie subalgebra $ad(\hat{g})$ of $\widehat{gl}(\hat{g})$. Thus after choosing a basis for $\hat{g}$, we defined

$$\begin{aligned} B : \hat{g} \times \hat{g} &\longrightarrow \mathbb{F} \\ (x, y) &\longmapsto B(x, y) := trace(ad(x) \circ ad(y)) \end{aligned}$$

Now we wish to rename form $\hat{B}$ and call it the *Killing form* B_V of a Lie algebra $\hat{g}$ contained in $\widehat{gl}(V)$ over a field $\mathbb{F}$ of characteristic 0. To do this we choose a basis for V, and define the form as:

$$B_V : \hat{g} \times \hat{g} \longrightarrow \mathbb{F}$$

$$(X, Y) \longmapsto B_V(X, Y) := \hat{B}(X, Y) = trace(X \circ Y)$$

We know that B_V is a symmetric bilinear form that also has the associative property, that is, for X, Y, Z in $\hat{g}$

$$B_V([X, Y], Z) = B_V(X, [Y, Z])$$

For a Killing form defined in this way on $\hat{g}$ contained in $\widehat{gl}(V)$, we also have these corresponding theorems:

First we examine the following theorem:

> A Lie subalgebra $\hat{g}$ of $\widehat{gl}(V)$ is solvable if and only if the Killing form $B_V(X,X) = 0$ for all X in $D^1\hat{g}$.

We examine first the case when V is a complex linear space. Assuming that the Lie subalgebra $\hat{g}$ of $\widehat{gl}(V)$ is solvable, we show that $B_V(X,X) = 0$ for all X in $D^1\hat{g}$. Now since $\hat{g}$ is a solvable Lie algebra in $\widehat{gl}(V)$, Lie's theorem says that we can find a basis in V such that all elements X in $\hat{g}$ can be represented simultaneously by upper triangular matrices. But we know that brackets of such matrices yield nilpotent linear transformations, which in this representation means that the matrices are upper triangular matrices with a zero diagonal. Obviously such matrices have a zero trace, and <u>linear</u> products of these matrices will also have a zero trace. Since the trace is independent of the basis used to calculate it, we can conclude that, since B_V is bilinear, $B_V(X,Y) = 0$ for all summands X and Y in $D^1\hat{g}$. [Recall that $D^1\hat{g}$ is the span of the set of all brackets of $\hat{g}$.] Thus the weaker conclusion also holds: $B_V(X,X) = 0$, and we can conclude that for V a complex linear space and $\hat{g}$ a solvable Lie algebra $B_V(X,X) = 0$ for all X in $D^1\hat{g}$.

Suppose now that our field of scalars is $\mathbb{R}$. If $\hat{g}$ is a solvable Lie subalgebra of $\widehat{gl}(V)$, we would like to conclude again that $B_V(X,X) = 0$ for all X in $D^1\hat{g}$. But, of course, we cannot imitate the above proof, which needed an algebraically closed field of characteristic zero.

However, starting with a Lie algebra $\hat{g}$ over $\mathbb{R}$ which is solvable, we know that its complexificaton $\hat{g}^c$ is solvable. Now $\hat{g}$ is a real Lie subalgebra of the real Lie algebra $\widehat{gl}(V)$ which acts on V. On complexification we obtain the complex linear space V^c, and $\hat{g}^c$, a complex Lie subalgebra of the complex Lie algebra $\widehat{gl}(V^c) = (\widehat{gl}(V))^c$, the complexification of $\widehat{gl}(V)$. The set of transformations $\widehat{gl}(V^c)$ acts on V^c. Thus, starting with $\hat{g}$ solvable, we know that $\hat{g}^c$ is also solvable, and we can apply the theorem to $\hat{g}^c$. Taking X in $\hat{g}$, we know that cX is in $\hat{g}^c$ for any c in $\mathbb{C}$. We also know that for X in $D^1\hat{g}$, cX is in $D^1\hat{g}^c$. Using the theorem, we have $B_{V^c}(cX, cX) = 0$ for all cX in $D^1\hat{g}^c$. Now we move back to $\hat{g}$. We take X in $D^1\hat{g}$. Then for $c = a + ib$ in $\mathbb{C}$, we have cX in $D^1\hat{g}^c$. Expanding

$$B_{V^c}(cX, cX) = B_{V^c}((a+ib)X, (a+ib)X) = \\ ((a^2 - b^2) + i(ab + ba))B_{V^c}(X,X) = 0$$

However starting with a Lie algebra $\hat{g}$ over $\mathbb{R}$, we know that its complexificaton $\hat{g}^c = \hat{g} \oplus i\hat{g}$, which is contained in the complexification $(\widehat{gl}(V))^c = \widehat{gl}(V^c)$. Thus for X in $\widehat{gl}(V)$ we also see that X is in $\widehat{gl}(V^c)$, and therefore it makes sense to write $B_{V^c}(X,X)$. However, we know that B_V is B_{V^c} restricted to $\hat{g}$, and we obtain by this restriction the Killing form for B_V that $B_{V^c}(X,Y) = trace(X \circ Y) = B_V(X,Y)$ for X and Y in $\hat{g}$. Thus, using the above relation $B_{V^c}(cX,cX) = 0$ we can conclude that $B_V(X,X) = 0$ and this gives us our theorem.

We now want to prove the converse, namely that for a Lie subalgebra $\hat{g}$ of $\widehat{gl}(V)$ over $\mathbb{R}$ or $\mathbb{C}$, if the Killing form $B_V(X,X) = 0$ for all X in $D^1\hat{g}$, then $\hat{g}$ is solvable.

Again we observe that we can assume that $D^1\hat{g} \neq 0$ for if $D^1\hat{g} = 0$ then $B_V(X,X) = 0$ is immediately satisfied and this means that $\hat{g}$ is abelian and therefore is solvable and the theorem is verified.

We first assume the field of scalars is $\mathbb{C}$. Then under this assumption we will show that for each Y in $D^2\hat{g}$, Y is a nilpotent linear transformation in $\widehat{gl}(V)$, and thus 2.8.2 says that $D^2\hat{g}$ is a nilpotent Lie algebra. But we know that a nilpotent Lie algebra is a solvable Lie algebra (see 2.5.1). This says then that $D^k(D^2\hat{g}) = 0$ for some k. Thus $D^{k+2}\hat{g} = D^k(D^2\hat{g}) = 0$, giving us our conclusion that $\hat{g}$ is solvable.

Thus we are reduced to proving that for each Y in $D^2\hat{g}$, Y is a linear nilpotent transformation in $\widehat{gl}(V)$ because the Killing form $B_V(X,X) = 0$ for all X in $D^1\hat{g}$.

But we know that B_V is just another name for form $\hat{B}$ and for form $\hat{B}$ we have already proven that for a Lie algebra $\hat{g}$ (a Lie subalgebra of $\widehat{gl}(V)$) and for V (a linear space over the field $\mathbb{C}$,) *Theorem* $\hat{B}$ holds. Thus we have

> Let the form $\hat{B}$ satisfy the condition that for all Y in $D^1\hat{g}$ and all X in $\hat{g}$, $\hat{B}(Y,X) = 0$. Then Y is a nilpotent linear transformation.

This translates into

> If the form $B_V(Y,X) = 0$, then Y is a nilpotent linear transformation,

where $\hat{g}$ is a subalgebra of $\widehat{gl}(V)$ and V is a complex linear space. Now our subalgebra is

$$D^1\hat{g} \subset \hat{g} \subset \widehat{gl}(V)$$

Also we have

$$D^2\hat{g} = D^1(D^1\hat{g}) \subset D^1\hat{g} \subset \widehat{gl}(V)$$

Thus we can apply our theorem to prove

> Let the form B_V satisfy the condition that for all Y in $D^2\hat{g}$ and all X in $D^1(\hat{g})$, $B_V(Y,X) = 0$. Then Y is a nilpotent linear transformation in $\widehat{gl}(V)$.

However our hypothesis is for all X in $D^1\hat{g}$, $B_V(X,X) = 0$. Certainly $D^2\hat{g}$ is a subset of $D^1\hat{g}$, and thus for all Y in $D^2\hat{g}$ and all X in $D^1\hat{g}$, $B_V(Y,X) = 0$. We can therefore conclude that Y is a nilpotent linear transformation in $\widehat{gl}(V)$, which is what we were seeking.

To complete this part of the exposition we still need to prove that for a Lie subalgebra $\hat{g}$ over $\mathbb{R}$ of $\widehat{gl}(V)$, if the Killing form $B_V(X,X) = 0$ for each X in $D^1\hat{g}$, then the Lie subalgebra is solvable. Of course, to prove this we first complexify the real Lie subalgebra $\hat{g}$, apply the theorem just proven to this case, and then move back down to the real Lie subalgebra.

Thus we assume that for a Lie subalgebra $\hat{g}$ over $\mathbb{R}$ the Killing form $B_V(X,X) = 0$ for each X in $D^1\hat{g}$ which is contained in $\widehat{gl}(V)$, where V is a n-dimensional real linear space. On complexification we obtain the complex linear space V^c, and $\hat{g}^c$, a complex Lie subalgebra of the complex Lie algebra $\widehat{gl}(V^c) = (\widehat{gl}(V))^c$ (which last mentioned is the complexification of $\widehat{gl}(V)$). Now the set of transformations $\widehat{gl}(V^c)$ acts on V^c. Thus any Z in $D^1\hat{g}^c$ can be written as $Z = cX$ for some X in $\hat{g}$ and some c in $\mathbb{C}$. Now Z is a linear combination of brackets $Z_{ij} = [Z_i, Z_j]$ for Z_i and Z_j in $\hat{g}^c$. But $Z_i = c_i X_i$ for some c_i in $\mathbb{C}$ and some X_i in $\hat{g}$; and likewise for Z_j. Thus $[Z_i, Z_j] = [c_i X_i, c_j X_j] = c_i c_j [X_i, X_j] = c_{ij} X_{ij}$. We can conclude that X_{ij} is in $D^1\hat{g}$. Calculating the Killing form B_{V^c} in $\hat{g}^c$, we obtain

$$B_{V^c}(c_{ij}X_{ij}, c_{ij}X_{ij}) = B_{V^c}((a_{ij} + ib_{ij})X_{ij}, ((a_{ij} + ib_{ij})X_{ij}) = \\ ((a_{ij}^2 - b_{ij}^2) + i(a_{ij}b_{ij} + b_{ij}a_{ij}))B_{V^c}(X_{ij}, X_{ij})$$

[Once again we are using the fact that X in $\widehat{gl}(V)$ is also in $\widehat{gl}(V^c)$. Thus it makes sense to write $B_{V^c}(X,X)$, and that B_V can be defined as the restriction of B_{V^c} to V.] But by hypothesis $B_V(X_{ij}, X_{ij}) = 0$ and this gives the desired conclusion that $B_{V^c}(c_{ij}X_{ij}, c_{ij}X_{ij}) = B_{V^c}(Z_{ij}, Z_{ij}) = 0$. But Z is a linear combination of the Z_{ij}, and since B_{V^c} is bilinear, we can conclude that $B_{V^c}(Z,Z) = 0$ for Z in $D^1\hat{g}^c$. Thus we know that $\hat{g}^c$ is solvable. And we have already shown that this means that $\hat{g}$ is solvable.

We now want to prove the "other" theorems for the Killing form B_V. But here is where a surprise occurs. In one direction the conclusion remains unchanged, i.e.:

> If a Lie subalgebra $\hat{g}$ over $\mathbb{R}$ or $\mathbb{C}$ contained in $\widehat{gl}(V)$ is semisimple, then the Killing form B_V is nondegenerate.

But in the other direction an important modification must be made:

> If a Lie subalgebra $\hat{g}$ over $\mathbb{R}$ or $\mathbb{C}$ contained in $\widehat{gl}(V)$ has a non-degenerate Killing form, then this algebra is either semisimple or it has a non-zero radical $\hat{r}$ which is abelian.

We want to comment on this phenomenon. When we treated the Killing form in the context of an abstract Lie algebra, we had the adjoint representation to move us over to the Lie algebra of matrices $\widehat{gl}(\hat{g})$. Thus the Lie subalgebra of $\widehat{gl}(\hat{g})$ was not arbitrary but came form a well-defined object. But now we begin with a arbitrary Lie subalgebra $\hat{g}$ of $\widehat{gl}(V)$, where there is no connection between $\hat{g}$ and the vector space V. And thus we impose another condition that connects $\hat{g}$ with V, namely that $\hat{g}$ has non-degenerate Killing form. In this sense we are more in the spirit of representation theory of Lie algebras.

Just as before we begin with the dual map. Every bilinear form on a linear space V determines a linear map between V and its dual V^*. In our case we have

$$\begin{aligned}
&\widehat{gl}(V) \xrightarrow{\mathcal{B}_V} (\widehat{gl}(V))^* \\
&X \longmapsto \mathcal{B}_V(X) : \widehat{gl}(V) \longrightarrow \mathbb{F} \\
&\qquad\qquad\qquad\qquad Y \longmapsto \mathcal{B}_V(X)(Y) := B_V(X,Y)
\end{aligned}$$

However, we are only interested in a Lie subalgebra $\hat{g}$ of $\widehat{gl}(V)$. Thus we restrict this dual map to the domain $\hat{g}$:

$$\begin{aligned}
&\hat{g} \xrightarrow{\mathcal{B}_V} (\hat{g})^* \\
&X \longmapsto \mathcal{B}_V(X) : \hat{g} \longrightarrow \mathbb{F} \\
&\qquad\qquad\qquad\qquad Y \longmapsto \mathcal{B}_V(X)(Y) := B_V(X,Y)
\end{aligned}$$

If there is a nonzero element W in the kernel $\hat{k}$ of this restricted map $\mathcal{B}_V$, this element has the property that for all Y in $\hat{g}$, $B_V(W,Y) = 0$. Thus any such W gives the zero map in $(\hat{g})^*$, and, of course, this means that there is a degeneracy in this restricted $\mathcal{B}_V$ map. For non-degeneracy it is demanded that the only zero map in $(\hat{g})^*$ comes from the zero element in $\hat{g}$ by $\mathcal{B}_V$. In other words non-degeneracy demands that the kernel $\hat{k}$ of this restricted map $\mathcal{B}_V$ be zero.

We first prove that a semisimple Lie subalgebra $\hat{g}$ of $\widehat{gl}(V)$ has a non-degenerate Killing form. This means we are affirming that the kernel $\hat{k}$ of the map $\mathcal{B}_V$ contains only the zero element of $\hat{g}$. First we observe that $\hat{k}$ is an ideal in $\hat{g}$. Recall that this follows from the associativity of the Killing form. It was used to show that $[\hat{k},\hat{g}]$ is contained in $\hat{k}$ as follows. Let W be in $\hat{k}$ and Y be in $\hat{g}$. Now for any Z in $\hat{g}$, we have $B_V([W,Y],Z) = B_V(W,[Y,Z])$. Since W is in $\hat{k}$, this says that $B_V(W, [Y, Z]) = 0$. Thus $B_V([W, Y], Z) = 0$ for all Z in $\hat{g}$, which says that $[W, Y]$ is in the kernel of the map $\mathcal{B}_V$. We conclude that $\hat{k}$ is an ideal in $\hat{g}$. We now show that this ideal is a solvable ideal. We restrict the Killing form now to $\hat{k}$. Let W be in $D^1\hat{k}$. Then $B_V(W, W) = ((\mathcal{B}_V)(W))(W)$. But W is in the kernel of $\mathcal{B}_V$, giving $B_V(W, W) = 0$ for all W in $D^1\hat{k}$. Thus we know that $\hat{k}$ is solvable, and a solvable ideal in $\hat{g}$. But we have assumed that $\hat{g}$ is semisimple. Thus we can conclude that $\hat{k} = 0$. But this is the same as asserting that map $\mathcal{B}_V$ is an isomorphism, which is the definition of non-degeneracy for the Killing form B_V.

Now in order to prove the theorem in the other direction, namely to prove that

> If a Lie subalgebra $\hat{g}$ over $\mathbb{R}$ or $\mathbb{C}$ contained in $\widehat{gl}(V)$ has a non-degenerate Killing form, then this algebra is either semisimple or it has a non-zero radical $\hat{r}$ which is abelian

we show that if $\hat{g}$ is not semisimple and has a nontrivial radical which is not abelian, then the Killing form is degenerate. This means that we are assuming that $\hat{g}$ has a non-trivial solvable ideal $\hat{s}$ which is not abelian. Thus we can affirm that for every X in $D^1\hat{s}$ the Killing form $B_V(X, X) = 0$. We show that this implies the existence of a degeneracy in the map $\mathcal{B}_V$ restricted to $\hat{g}$ except in the case when $\hat{s}$ is abelian. We remark that since $\hat{s}$ is just a part of $\hat{g}$, we need a way of extending this information on $\hat{s}$ to all of $\hat{g}$. This can be done in our situation since $\hat{s}$ is a solvable ideal in $\hat{g}$ and since the Killing form is associative. Thus for any Y in $\hat{g}$ and any W_1 and W_2 in $D^1\hat{s}$, we consider $B_V(Y, [W_1, W_2])$. Associativity of B_V gives $B_V(Y, [W_1, W_2]) = B_V([Y, W_1], W_2)$. Since $D^1\hat{s}$ is an ideal in $\hat{g}$, we have that $[Y, W_1]$ is in $D^1\hat{s}$. Our condition on B_V says that $B_V([Y, W_1], W_2) = 0$. [Recall that $B_V(X, X) = 0$ implies $B_V(X, Y) = 0$ for all X and Y in $\hat{g}$.] Thus $B_V([Y, [W_1, W_2]) = 0$, which says $[W_1, W_2]$ is in the kernel of the restricted map $\mathcal{B}_V$. We can therefore conclude that if $D^1\hat{s}$ is not abelian, then the restricted map $\mathcal{B}_V$ has a degeneracy.

If, however, $D^1\hat{s}$ is abelian, we can modify the above proof a little and reach the same conclusion. We choose any Y in $\hat{g}$, any W_1 in $\hat{s}$, and any W_2 in $D^1\hat{s}$. Now $B_V(Y, [W_1, W_2]) = B_V([Y, W_1], W_2)$. We have $[Y, W_1]$ in $\hat{s}$ by the ideal structure of $\hat{s}$ in $\hat{g}$, and we have W_2 in $D^1\hat{s}$. Obviously we cannot say that $B_V([Y, W_1], W_2) = 0$ since $[Y, W_1] = Z$ is not known to be in $D^1\hat{s}$, but only in $\hat{s}$. But we know that $B_V([Y, W_1], W_2) = B_V(Z, W_2) = trace(Z \circ W_2)$. Since Z and W_2 are both in $\hat{s}$, and $\hat{s}$ is a solvable Lie subalgebra of $\widehat{gl}(V)$, Lie's Theorem tells us we can find a basis in V such that all the matrices in $\hat{s} \subset \widehat{gl}(V)$ written with respect to that basis are in upper triangular form with the eigenvalues on the diagonal. [Of course, this means we are now working in the field of scalars $\mathbb{C}$]. Now Z is in $\hat{s}$ and thus may have a non-zero diagonal, and W_2 is in $D^1\hat{s}$, which we know is a linear nilpotent transformation in $\widehat{gl}(V)$. Thus its eigenvalues are all zero. Now since the matrix product $(Z \circ W_2)$ has a zero diagonal, the trace of the product equal to zero. Thus we can conclude that $B_V(Y, [W_1, W_2]) = 0$ for all Y in $\hat{g}$, and that $[W_1, W_2]$ is in the kernel of $\mathcal{B}_V$, making $\mathcal{B}_V$ degenerate. This proves the claim when the scalar field is $\mathbb{C}$. Obviously we will need to use a further argument to obtain conclusions when the scalar field is $\mathbb{R}$.

At this point we make the following observation. We could not use the fact that $B_V(X, X) = 0$ for all X in $D^1\hat{s}$ in the above proof. But instead all we needed was that $D^1\hat{s}$ was a non-trivial abelian ideal. Thus our situation is as follows. We have a Lie subalgebra $\hat{g}$ in $\widehat{gl}(V)$ and we have that $\hat{g}$ contains a non-trivial solvable ideal $\hat{s}$. This means, of course, that $\hat{g}$ has a nontrivial radical. But we also know that this means that $\hat{g}$ contains a non-trivial abelian ideal $\hat{a}$, and we assume it is not equal to $\hat{s}$, i.e., $\hat{a}$ is properly contained in $\hat{s}$. Thus we can repeat the above proof, using this abelian ideal $\hat{s}$ instead of $D^1\hat{s}$, and show immediately that the restricted map $\mathcal{B}_\mathcal{V}$ has a degeneracy on $\hat{g}$ without analyzing the condition of B_V on $D^1\hat{s}$.

We now complete the proof by treating the real case. Thus, we are now in the situation where we have a real linear space V and a subalgebra $\hat{g}$ of $\widehat{gl}(V)$. We also have a non-trivial solvable ideal $\hat{s}$ of $\hat{g}$ and for every X in $D^1\hat{s}$, $B_V(X, X) = 0$. And finally $D^1\hat{s}$ is also abelian. We want to show that all this implies that a degeneracy exists in the restricted map $\mathcal{B}_V$. Our first step is to complexify. This gives us a non-trivial solvable ideal $\hat{s}^c$ of $\hat{g}^c$. We choose a Z in $D^1\hat{s}^c$. Now we know that Z is a linear combination of brackets $Z_{ij} = [Z_i, Z_j]$ for Z_i and Z_j in $\hat{s}^c$. But $Z_i = c_i X_i$ for some c_i in $\mathbb{C}$ and some X_i in $\hat{s}$; and likewise for Z_j. Thus $[Z_i, Z_j] = [c_i X_i, c_j X_j] = c_i c_j [X_i, X_j] = c_{ij} X_{ij}$. We can conclude that X_{ij} is in $D^1\hat{s}$. Calculating the Killing form B_{V^c} in $\hat{s}^c$, we obtain

$$B_{V^c}(c_{ij}X_{ij}, c_{ij}X_{ij}) = B_{V^c}((a_{ij} + ib_{ij})X_{ij}, ((a_{ij} + ib_{ij})X_{ij}) = \\ ((a_{ij}^2 - b_{ij}^2) + i(a_{ij}b_{ij} + b_{ij}a_{ij}))B_{V^c}(X_{ij}, X_{ij})$$

[Once again we are using the fact that X in $\widehat{gl}(V)$ is also in $\widehat{gl}(V^c)$. Thus it makes sense to write $B_{V^c}(X, X)$, and it also makes sense that B_V can be defined as the restriction of B_{V^c} to V.] But by hypothesis $B_V(X_{ij}, X_{ij}) = 0$, which gives the desired conclusion that $B_{V^c}(c_{ij}X_{ij}, c_{ij}X_{ij}) = B_{V^c}(Z_{ij}, Z_{ij}) = 0$. But Z is a linear combination of the Z_{ij}, and since B_{V^c} is bilinear, we can conclude that $B_{V^c}(Z, Z) = 0$ for Z in $D^1\hat{s}^c$. Thus we know that $\hat{s}^c$ is solvable. And we have already shown that this means that $\hat{g}$ is solvable.

However the situation changes when $\hat{s}$ itself is a non-trivial abelian ideal. Then the above proof will not produce a non-zero element in $\hat{s}$ which is in the kernel of the restricted map $\mathcal{B}_V$. But this condition holding for $\hat{s}$ means, of course, that we can choose $\hat{s}$ to be the maximal solvable ideal in $\hat{g}$, that is, $\hat{s}$ becomes the radical of $\hat{g}$. Thus when $\hat{g}$ has an abelian radical, we cannot conclude that $\mathcal{B}_V$ is degenerate. In fact in this situation $\mathcal{B}_V$ can be nondegenerate, as the following example shows.

We take the 2x2 matrices over $\mathbb{F}$, where $\mathbb{F}$ can be either $\mathbb{R}$ or $\mathbb{C}$. We choose our subalgebra $\hat{g}$ of Lie algebra $\widehat{gl}(\mathbb{F}^2)$ to be the entire Lie algebra $\widehat{gl}(\mathbb{F}^2)$ itself. Our $\hat{g}$ is then the 4-dimensional Lie algebra $\widehat{gl}(\mathbb{F}^2)$. Now $\hat{g}$ is not semisimple since it has a one-dimensional abelian radical, which consists of all the scalar matrices [diagonal matrices with the same scalar on the diagonal]. The set of 3-dimensional matrices with trace zero is the simple Lie algebra $\widehat{sl}(2,\mathbb{F})$. We calculate the Killing Form $B_{\mathbb{F}^2}$ of $\hat{g}$ and the dual map $\mathcal{B}_{\mathbb{F}^2}$ of $\hat{g}$ to $(\hat{g})^*$.

First we choose the 4-dimensional canonical basis for the matrices in $\widehat{gl}(\mathbb{F}^2) = \hat{g}$: $(E_{11}, E_{21}, E_{12}, E_{22})$, where E_{ij} is the 2x2 matrix with 1 in the i,j position and 0 everywhere else. Since this basis has no natural order, we choose the above order for the four basis vectors of $\hat{g}$: $(E_{11}, E_{21}, E_{12}, E_{22})$. With respect to this basis we write the corresponding 4x4 matrix representing $B_{\mathbb{F}^2}$

For the first column:

$$\begin{aligned}
(B_{\mathbb{F}^2})_{11} &= trace(E_{11} \circ E_{11}) = trace E_{11} = 1 + 0 = 1 \\
(B_{\mathbb{F}^2})_{21} &= trace(E_{21} \circ E_{11}) = trace E_{21} = 0 + 0 = 0 \\
(B_{\mathbb{F}^2})_{31} &= trace(E_{12} \circ E_{11}) = trace\ 0 = 0 \\
(B_{\mathbb{F}^2})_{41} &= trace(E_{22} \circ E_{11}) = trace\ 0 = 0
\end{aligned}$$

For the second column:

$$\begin{aligned}
(B_{\mathbb{F}^2})_{12} &= trace(E_{11} \circ E_{21}) = trace\ 0 = 0 \\
(B_{\mathbb{F}^2})_{22} &= trace(E_{21} \circ E_{21}) = trace\ 0 = 0 \\
(B_{\mathbb{F}^2})_{32} &= trace(E_{12} \circ E_{21}) = trace E_{11} = 1 + 0 = 1 \\
(B_{\mathbb{F}^2})_{42} &= trace(E_{22} \circ E_{21}) = trace E_{21} = 0 + 0 = 0
\end{aligned}$$

For the third column:

$$\begin{aligned}
(B_{\mathbb{F}^2})_{13} &= trace(E_{11} \circ E_{12}) = trace E_{12} = 0 + 0 = 0 \\
(B_{\mathbb{F}^2})_{23} &= trace(E_{21} \circ E_{12}) = trace E_{22} = 0 + 1 = 1 \\
(B_{\mathbb{F}^2})_{33} &= trace(E_{12} \circ E_{12}) = trace\ 0 = 0 \\
(B_{\mathbb{F}^2})_{43} &= trace(E_{22} \circ E_{12}) = trace\ 0 = 0
\end{aligned}$$

For the fourth column:

$$
\begin{aligned}
(B_{\mathbb{F}^2})_{14} &= trace(E_{11} \circ E_{22}) = trace\ 0 = 0 \\
(B_{\mathbb{F}^2})_{24} &= trace(E_{21} \circ E_{22}) = trace\ 0 = 0 \\
(B_{\mathbb{F}^2})_{34} &= trace(E_{12} \circ E_{22}) = trace E_{12} = 0 + 0 = 0 \\
(B_{\mathbb{F}^2})_{44} &= trace(E_{22} \circ E_{22}) = trace E_{22} = 0 + 1 = 1
\end{aligned}
$$

Thus, with respect to the basis chosen for $\hat{g}$, namely $(E_{11}, E_{21}, E_{12}, E_{22})$, the matrix for $B_{\mathbb{F}^2}$ becomes

$$
B_{\mathbb{F}^2} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
$$

From this matrix we see immediately that the bilinear form $B_{\mathbb{F}^2}$ is nondegenerate since the determinant of the matrix is nonzero.

Recalling the definition of $\mathcal{B}_{\mathbb{F}^2}$, we have

$$
E_{ij} \longmapsto \mathcal{B}_{\mathbb{F}^2}(E_{ij})
$$

and

$$
\mathcal{B}_{\mathbb{F}^2}(E_{ij})(E_{kl}) = B_{\mathbb{F}^2}(E_{ij}, E_{kl}) = trace(E_{ij}E_{kl})
$$

Using the basis chosen for $\hat{g}$, we calculate the image of these four elements of $\hat{g}$ by $\mathcal{B}_{\mathbb{F}^2}$. We wish to express these duals in $\hat{g}^*$ as 4x1 row matrices, again using the above basis:

$$
\begin{aligned}
\mathcal{B}_{\mathbb{F}^2}(E_{11})(E_{11}) &= B_{\mathbb{F}^2}(E_{11}, E_{11}) = trace(E_{11}E_{11}) = 1 \\
\mathcal{B}_{\mathbb{F}^2}(E_{11})(E_{21}) &= B_{\mathbb{F}^2}(E_{11}, E_{21}) = trace(E_{11}E_{21}) = 0 \\
\mathcal{B}_{\mathbb{F}^2}(E_{11})(E_{12}) &= B_{\mathbb{F}^2}(E_{11}, E_{12}) = trace(E_{11}E_{12}) = 0 \\
\mathcal{B}_{\mathbb{F}^2}(E_{11})(E_{22}) &= B_{\mathbb{F}^2}(E_{11}, E_{22}) = trace(E_{11}E_{22}) = 0
\end{aligned}
$$

which gives the matrix representation of $\mathcal{B}_{\mathbb{F}^2}(E_{11})$ with respect to this basis as $[1, 0, 0, 0]$. Continuing, we have

$$
\begin{aligned}
\mathcal{B}_{\mathbb{F}^2}(E_{21})(E_{11}) &= B_{\mathbb{F}^2}(E_{21}, E_{11}) = trace(E_{21}E_{11}) = 0 \\
\mathcal{B}_{\mathbb{F}^2}(E_{21})(E_{21}) &= B_{\mathbb{F}^2}(E_{21}, E_{21}) = trace(E_{21}E_{21}) = 0 \\
\mathcal{B}_{\mathbb{F}^2}(E_{21})(E_{12}) &= B_{\mathbb{F}^2}(E_{21}, E_{12}) = trace(E_{21}E_{12}) = 1 \\
\mathcal{B}_{\mathbb{F}^2}(E_{21})(E_{22}) &= B_{\mathbb{F}^2}(E_{21}, E_{22}) = trace(E_{21}E_{22}) = 0
\end{aligned}
$$

which gives the matrix representation of $\mathcal{B}_{\mathbb{F}^2}(E_{21})$ with respect to this basis as $[0, 0, 1, 0]$. Continuing, we have

$$\mathcal{B}_{\mathbb{F}^2}(E_{12})(E_{11}) = B_{\mathbb{F}^2}(E_{12}, E_{11}) = trace(E_{12}E_{11}) = 0$$
$$\mathcal{B}_{\mathbb{F}^2}(E_{12})(E_{21}) = B_{\mathbb{F}^2}(E_{12}, E_{21}) = trace(E_{12}E_{21}) = 1$$
$$\mathcal{B}_{\mathbb{F}^2}(E_{12})(E_{12}) = B_{\mathbb{F}^2}(E_{12}, E_{12}) = trace(E_{12}E_{12}) = 0$$
$$\mathcal{B}_{\mathbb{F}^2}(E_{12})(E_{22}) = B_{\mathbb{F}^2}(E_{12}, E_{22}) = trace(E_{12}E_{22}) = 0$$

which gives the matrix representation of $\mathcal{B}_{\mathbb{F}^2}(E_{12})$ with respect to this basis as $[0, 1, 0, 0]$. Continuing, we have

$$\mathcal{B}_{\mathbb{F}^2}(E_{22})(E_{11}) = B_{\mathbb{F}^2}(E_{22}, E_{11}) = trace(E_{22}E_{11}) = 0$$
$$\mathcal{B}_{\mathbb{F}^2}(E_{22})(E_{21}) = B_{\mathbb{F}^2}(E_{22}, E_{21}) = trace(E_{22}E_{21}) = 0$$
$$\mathcal{B}_{\mathbb{F}^2}(E_{22})(E_{12}) = B_{\mathbb{F}^2}(E_{22}, E_{12}) = trace(E_{22}E_{12}) = 0$$
$$\mathcal{B}_{\mathbb{F}^2}(E_{22})(E_{22}) = B_{\mathbb{F}^2}(E_{22}, E_{22}) = trace(E_{22}E_{22}) = 1$$

which gives the matrix representation of $\mathcal{B}_{\mathbb{F}^2}(E_{11})$ with respect to this basis as $[0, 0, 0, 1]$.

However if we wish to write the map

$$\mathcal{B}_{\mathbb{F}^2} : \hat{g} \longrightarrow (\hat{g})^*$$

as a 4x4 matrix with respect to this basis, we must transpose the above row vectors to be the columns of its matrix and we get:

$$\mathcal{B}_{\mathbb{F}^2} \longmapsto \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

We note that it is this matrix that shows that the map

$$\mathcal{B}_{\mathbb{F}^2} : \hat{g} \longrightarrow (\hat{g})^*$$

is nondegenerate, i.e., that its kernel is 0. Moreover, this gives us the result that we have been seeking — that a non-degenerate Killing form on $\hat{g}$ in $\widehat{gl}(\mathbb{F}^2)$ does not necessarily give the conclusion that $\hat{g}$ is semisimple, but it can give a Lie algebra with an abelian radical. In fact, all $\widehat{gl}(\mathbb{F}^n)$ for all n with $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ have this structure.

In order to complete this discussion we now want to define the Killing form on the Lie algebra $\hat{g} = \widehat{gl}(\mathbb{F}^2)$ and not the Killing form on the linear space $\mathbb{F}^2$ as we did above. The latter Killing form $B_{\mathbb{F}^2}$ we have just shown is non-degenerate, but $\hat{g} = \widehat{gl}(\mathbb{F}^2)$ is not semisimple. Since we now want to focus on the Killing form B of the Lie algebra $\hat{g} = \widehat{gl}(\mathbb{F}^2)$, we pick any two elements X and Y of $\widehat{gl}(\mathbb{F}^2)$ and calculate $B(X, Y) = trace(ad(X), ad(Y))$.

Using the basis $(E_{11}, E_{21}, E_{12}, E_{22})$ for $\widehat{gl}(\mathbb{F}^2)$, we calculate B as follows.

First we need the $ad(E_{ij})$.

$$\begin{gathered} ad(E_{ij}) : \widehat{gl}(\mathbb{F}^2) \longrightarrow \widehat{gl}(\mathbb{F}^2) \\ E_{kl} \longmapsto ad(E_{ij}) \cdot E_{kl} = [E_{ij}, E_{kl}] = E_{ij}E_{kl} - E_{kl}E_{ij} \end{gathered}$$

Thus

$$\begin{gathered} \text{If } j = k,\ i \neq l, \text{ then } ad(E_{ij}) \cdot E_{kl} = E_{il} \\ \text{If } j \neq k,\ i = l, \text{ then } ad(E_{ij}) \cdot E_{kl} = -E_{kj} \\ \text{If } j = k,\ i = l, \text{ then } ad(E_{ij}) \cdot E_{kl} = E_{ii} - E_{jj} \end{gathered}$$

Then we choose the canonical basis for the 16-dimensional space $\widehat{gl}(\widehat{gl}(\mathbb{F}^2))$ in the following manner.

$$(e_{11}, e_{21}, e_{31}, e_{41}, e_{12}, e_{22}, e_{32}, e_{42}, e_{13}, e_{23}, e_{33}, e_{43}, e_{14}, e_{24}, e_{34}, e_{44})$$

and we get the matrix

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

On this basis we have

$$ad(E_{11}) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad ad(E_{21}) = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$ad(E_{12}) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 \end{bmatrix} \qquad ad(E_{22}) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The Killing form B of the Lie algebra $\widehat{gl}(\mathbb{F}^2)$, written as a matrix with respect to the basis $(E_{11}, E_{21}, E_{12}, E_{22})$ is determined by 16 calculations of the following type.

$$B(E_{ij}, E_{kl}) = trace(ad(E_{ij}) \circ ad(E_{kl}))$$

For instance the (1,1)-term of the matrix is

$$B_{11} = B(E_{11}, E_{11}) = trace(ad(E_{11}) \circ ad(E_{11}))$$

$$ad(E_{11}) \circ ad(E_{11}) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$trace(ad(E_{11}) \circ ad(E_{11})) = 2$$

Continuing in this manner with the other matrix entries, we arrive at the B matrix

$$B = \begin{bmatrix} 2 & 0 & 0 & -2 \\ 0 & 0 & 4 & 0 \\ 0 & 4 & 0 & 0 \\ -2 & 0 & 0 & 2 \end{bmatrix}$$

We see that the B matrix is singular. Thus we conclude that $\widehat{gl}(\mathbb{F}^2)$ is not semisimple and has a Killing form as a Lie algebra which is degenerate.

We remark that even though the Killing form B on the Lie algebra $\widehat{gl}(\mathbb{F}^2)$ is degenerate, the Killing form $B_{\mathbb{F}^2}$ on the linear space $\mathbb{F}^2$ is non-degenerate. Our theorems say that for the first assertion that $\widehat{gl}(\mathbb{F}^2)$ is not semisimple; and for the second assertion that $\widehat{gl}(\mathbb{F}^2)$ is either semisimple or has an abelian radical $\neq 0$. Combining these two assertions, we conclude that $\widehat{gl}(\mathbb{F}^2)$ has an abelian radical.

Guided by the other theorems on solvability, we arrive at the following computations and conclusions. First we calculate $D^1(\widehat{gl}(\mathbb{F}^2))$. Using $(E_{11}, E_{21}, E_{12}, E_{22})$ for a basis of $\widehat{gl}(\mathbb{F}^2)$, we have:

$$\begin{gathered} [E_{11}, E_{21}] = E_{11}E_{21} - E_{21}E_{11} = -E_{21} \\ [E_{11}, E_{12}] = E_{11}E_{12} - E_{12}E_{11} = E_{12} \\ [E_{11}, E_{22}] = E_{11}E_{22} - E_{22}E_{11} = 0 \\ [E_{21}, E_{12}] = E_{21}E_{12} - E_{12}E_{21} = E_{22} - E_{11} \\ [E_{21}, E_{22}] = E_{21}E_{22} - E_{22}E_{21} = -E_{21} \\ [E_{12}, E_{22}] = E_{12}E_{22} - E_{22}E_{12} = E_{12} \end{gathered}$$

And thus $D^1(\widehat{gl}(\mathbb{F}^2)) = Sp(E_{12}, E_{11} - E_{22}, E_{12})$. Our first remark is that $D^1(\widehat{gl}(\mathbb{F}^2)) \neq \widehat{gl}(\mathbb{F}^2)$, and thus $\widehat{gl}(\mathbb{F}^2)$ cannot be semisimple. Now

$$\begin{gathered} B(E_{11} - E_{22}, E_{11} - E_{22}) = B(E_{11}, E_{11}) + B(E_{22}, E_{22}) - 2B(E_{11}, E_{22}) = \\ 2 + 2 - 2(-2) = 8 \end{gathered}$$

and this computation gives us sufficient information for us to assert that $B(X,X) \neq 0$ for all X in $D^1(\widehat{gl}(\mathbb{F}^2))$. Thus $\widehat{gl}(\mathbb{F}^2)$ is not solvable and we conclude that $\widehat{gl}(\mathbb{F}^2)$ has a semisimple part and a non-trivial radical.

Similarly we have

$$B_{\mathbb{F}^2}(E_{11} - E_{22}, E_{11} - E_{22}) =$$
$$B_{\mathbb{F}^2}(E_{11}, E_{11}) + B_{\mathbb{F}^2}(E_{22}, E_{22}) - 2B_{\mathbb{F}^2}(E_{11}, E_{22}) = 1 + 1 - 2(0) = 2$$

which is sufficient information for us to assert that $B_{\mathbb{F}^2}(X,X) \neq 0$ for all X in $D^1(\widehat{gl}(\mathbb{F}^2))$, and thus we reach the same conclusion as above.

There are some additional remarks that we would like to make. In section 2.4 we showed that if $\hat{r}$ is the radical of a Lie algebra $\hat{g}$, then $[\hat{g}, \hat{r}] = D^1\hat{g} \cap \hat{r}$. Now suppose the radical is also the center $\hat{z}$ of $\hat{g}$. This would mean that $\hat{r}$ is abelian, but also we would have $[\hat{g}, \hat{z}] = 0 = D^1\hat{g} \cap \hat{z}$. If we assumed the Levi Decomposition Theorem, this would mean that $D^1\hat{g}$ is semisimple. Now the counterexample we examined above has exactly this structure. $\widehat{gl}(\mathbb{F}^2)$ has a semisimple part $\widehat{sl}(2, \mathbb{F})$ and an abelian radical which is the center of $\widehat{gl}(\mathbb{F}^2)$. Thus it is still an open question for us whether the only counterexamples must have this additional structure, i.e., whether $\hat{g}$ must be the direct sum of a semisimple subalgebra and a radical which is abelian. [We add the remark that a Lie algebra whose non-trivial radical is the center is called a <u>reductive</u> Lie algebra.]

2.15.2 The Casimir Operator.

With this new definition of the Killing form we can now define the Casimir operator for a Lie subalgebra $\hat{g}$ of $\widehat{gl}(V)$ whose Killing form is non-degenerate. However, since the understanding of this operator is not very transparent, we begin by giving three examples of how one calculates the Casimir operator. They will help us understand this operator better .

We already have at hand our first example for which to compute this operator. We calculate the Casimir operator for the Lie algebra $\widehat{gl}(\mathbb{F}^2)$. [We have shown this Lie algebra to have a non-degenerate Killing form on the linear space $\mathbb{F}^2$.] We let $\hat{g} = \widehat{gl}(\mathbb{F}^2)$.

First we choose a basis for the four-dimensional $\hat{g}$: $\{E_{11}, E_{21}, E_{12}, E_{22}\}$. We know that $\mathcal{B}_{\mathbb{F}^2}$ maps $\hat{g}$ bijectively onto its dual $\hat{g}^*$ and what we are now seeking is the dual basis of $\{E_{11}, E_{21}, E_{12}, E_{22}\}$. First we give the symbols for elements in $\hat{g}^*$ which will be dual to the basis of $\hat{g}$. We choose $\{E_{11}^*, E_{21}^*, E_{12}^*, E_{22}^*\}$. This means that the dual element E_{ij}^* in $\hat{g}^*$ acting on the "undual" element E_{kl} in $\hat{g}$ will give 1 when $(k,l) = (i,j)$ and will give 0 for the other three choices of basis elements. We recall that if E_{ij} is written with respect to the basis $\{E_{11}, E_{21}, E_{12}, E_{22}\}$, it will be a 4x1 column vector with 1 in the ij-th place and 0 elsewhere. This means that E^*_{ij} written with respect to the basis $\{E^*_{11}, E^*_{21}, E^*_{12}, E^*_{22}\}$ will be a 1x4 row vector with 1 in the ij-th place and 0 elsewhere. In other words each E^*_{ij} is a well-defined element in $\hat{g}^*$. The important observation is that E^*_{ij} is not a 2x2 matrix, while E_{ij} is an element of $\hat{g} = \widehat{gl}(\mathbb{F}^2)$, and thus is 2x2 matrix over $\mathbb{F}$. But by using the inverse of the bijective map $\mathcal{B}_{\mathbb{F}^2}$, which is a map from $\hat{g}$ to $\hat{g}^*$, we can write E^*_{ij} as a 4x4 matrix. We will name these matrices as follows:

$$\mathcal{B}^{-1}_{\mathbb{F}^2}(E^*_{ij}) := E'_{ij}$$

We now determine these matrices. For the matrix E'_{11}, we first write it using the basis $\{E_{11}, E_{21}, E_{12}, E_{22}\}$:

$$E'_{11} = aE_{11} + bE_{21} + cE_{12} + dE_{22}$$

that is,

$$E'_{11} = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

Now

$$\mathcal{B}_{\mathbb{F}^2}(E'_{11}) = E^*_{11}$$

Thus

$$E^*_{11} = \mathcal{B}_{\mathbb{F}^2}(aE_{11} + bE_{21} + cE_{12} + dE_{22}) =$$
$$\mathcal{B}_{\mathbb{F}^2}(aE_{11}) + \mathcal{B}_{\mathbb{F}^2}(bE_{21}) + \mathcal{B}_{\mathbb{F}^2}(cE_{12}) + \mathcal{B}_{\mathbb{F}^2}(dE_{22})$$

We now operate on the basis $(E_{11}, E_{21}, E_{12}, E_{22})$.

$$1 = E^*_{11} \cdot E_{11} =$$
$$\mathcal{B}_{\mathbb{F}^2}(aE_{11}) \cdot E_{11} + \mathcal{B}_{\mathbb{F}^2}(bE_{21}) \cdot E_{11} + \mathcal{B}_{\mathbb{F}^2}(cE_{12}) \cdot E_{11} + \mathcal{B}_{\mathbb{F}^2}(dE_{22}) \cdot E_{11}$$

But by definition of $\mathcal{B}_{\mathbb{F}^2}$, we have

$$\mathcal{B}_{\mathbb{F}^2}(E_{ij}) \cdot E_{kl} = B_{\mathbb{F}^2}(E_{ij}, E_{kl}) = trace(E_{ij}E_{kl})$$

Thus

$$1 = aB_{\mathbb{F}^2}(E_{11}, E_{11}) + bB_{\mathbb{F}^2}(E_{21}, E_{11}) + cB_{\mathbb{F}^2}(E_{12}, E_{11}) + dB_{\mathbb{F}^2}(E_{22}, E_{11}) =$$
$$a(trace(E_{11}E_{11})) + b(trace(E_{21}E_{11}))+$$
$$c(trace(E_{12}E_{11})) + d(trace(E_{22}E_{11})) = a$$

Continuing

$$0 = E_{11}^* \cdot E_{21} =$$
$$\mathcal{B}_{\mathbb{F}^2}(aE_{11}) \cdot E_{21} + \mathcal{B}_{\mathbb{F}^2}(bE_{21}) \cdot E_{21} + \mathcal{B}_{\mathbb{F}^2}(cE_{12}) \cdot E_{21} + \mathcal{B}_{\mathbb{F}^2}(dE_{22}) \cdot E_{21}$$

Thus

$$0 = aB_{\mathbb{F}^2}(E_{11}, E_{21}) + bB_{\mathbb{F}^2}(E_{21}, E_{21}) + cB_{\mathbb{F}^2}(E_{12}, E_{21}) + dB_{\mathbb{F}^2}(E_{22}, E_{21}) =$$
$$a(trace(E_{11}E_{21})) + b(trace(E_{21}E_{21}))+$$
$$c(trace(E_{12}E_{21})) + d(trace(E_{22}E_{21})) = c$$

Continuing

$$0 = E_{11}^* \cdot E_{12} =$$
$$\mathcal{B}_{\mathbb{F}^2}(aE_{11}) \cdot E_{12} + \mathcal{B}_{\mathbb{F}^2}(bE_{21}) \cdot E_{12} + \mathcal{B}_{\mathbb{F}^2}(cE_{12}) \cdot E_{12} + \mathcal{B}_{\mathbb{F}^2}(dE_{22}) \cdot E_{12}$$

Thus

$$0 = aB_{\mathbb{F}^2}(E_{11}, E_{12}) + bB_{\mathbb{F}^2}(E_{21}, E_{12}) + cB_{\mathbb{F}^2}(E_{12}, E_{12}) + dB_{\mathbb{F}^2}(E_{22}, E_{12}) =$$
$$a(trace(E_{11}E_{12})) + b(trace(E_{21}E_{12}))+$$
$$c(trace(E_{12}E_{12})) + d(trace(E_{22}E_{12})) = b$$

Continuing

$$0 = E_{11}^* \cdot E_{22} =$$
$$\mathcal{B}_{\mathbb{F}^2}(aE_{11}) \cdot E_{22} + \mathcal{B}_{\mathbb{F}^2}(bE_{21}) \cdot E_{22} + \mathcal{B}_{\mathbb{F}^2}(cE_{12}) \cdot E_{22} + \mathcal{B}_{\mathbb{F}^2}(dE_{22}) \cdot E_{22}$$

Thus

$$0 = aB_{\mathbb{F}^2}(E_{11}, E_{22}) + bB_{\mathbb{F}^2}(E_{21}, E_{22}) + cB_{\mathbb{F}^2}(E_{22}, E_{12}) + dB_{\mathbb{F}^2}(E_{22}, E_{22}) =$$
$$a(trace(E_{11}E_{22})) + b(trace(E_{21}E_{22}))+$$
$$c(trace(E_{12}E_{22})) + d(trace(E_{22}E_{22})) = d$$

Therefore we conclude that

$$E_{11}' = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

We observe that we are essentially calculating the Killing form $B_{\mathbb{F}^2}$ with respect to the basis $(E_{11}, E_{21}, E_{12}, E_{22})$.

Recall that the *Killing form* B_V of a Lie algebra $\hat{g}$ contained in $\widehat{gl}(V)$ over a field $\mathbb{F}$ of characteristic 0 is defined in the following manner. After choosing a basis for V, then the following traces define B_V:

$$B_V : \hat{g} \times \hat{g} \longrightarrow \mathbb{F}$$
$$(X, Y) \longmapsto B_V(X, Y) := trace(X \circ Y)$$

Now in the above calculations we have computed these traces:

$$1 = a(trace(E_{11}E_{11})) + b(trace(E_{21}E_{11})) + c(trace(E_{12}E_{11})) + d(trace(E_{22}E_{11})) = a$$

$$0 = a(trace(E_{11}E_{21})) + b(trace(E_{21}E_{21})) + c(trace(E_{12}E_{21})) + d(trace(E_{22}E_{21})) = c$$

$$0 = a(trace(E_{11}E_{12})) + b(trace(E_{21}E_{12})) + c(trace(E_{12}E_{12})) + d(trace(E_{22}E_{12})) = b$$

$$0 = a(trace(E_{11}E_{22})) + b(trace(E_{21}E_{22})) + c(trace(E_{12}E_{22})) + d(trace(E_{22}E_{22})) = d\ .$$

This gives the first column of the matrix for the Killing form of $B_{\mathbb{C}^2}$:

$$B_{\mathbb{C}^2} = \begin{bmatrix} 1 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{bmatrix}$$

Continuing in this manner we calculate the entire matrix for the Killing form of $B_{\mathbb{C}^2}$. However we do know that the Killing form is a symmetric matrix, and thus we already know the matrix has the following form.

$$B_{\mathbb{C}^2} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{bmatrix}$$

Thus we need only compute six more entries in the Killing form. Now we have for the matrix E'_{21}

$$E'_{21} = eE_{11} + fE_{21} + gE_{12} + hE_{22}$$

that is,

$$E'_{21} = \begin{bmatrix} e & g \\ f & h \end{bmatrix}$$

Now

$$\mathcal{B}_{\mathbb{F}^2}(E'_{21}) = E^*_{21}$$

Thus

$$E^*_{21} = \mathcal{B}_{\mathbb{F}^2}(eE_{11} + fE_{21} + gE_{12} + hE_{22}) =$$
$$\mathcal{B}_{\mathbb{F}^2}(eE_{11}) + \mathcal{B}_{\mathbb{F}^2}(fE_{21}) + \mathcal{B}_{\mathbb{F}^2}(gE_{12}) + \mathcal{B}_{\mathbb{F}^2}(hE_{22})$$

We now operate on the basis $(E_{11}, E_{21}, E_{12}, E_{22})$. But we already know $trace(E_{21}E_{11}) = 0$. Thus we only compute the following:

$$1 = E^*_{21} \cdot E_{21}$$
$$1 = e(trace(E_{11}E_{21})) + f(trace(E_{21}E_{21}))+$$
$$g(trace(E_{12}E_{21})) + h(trace(E_{22}E_{21})) = g$$

Continuing

$$0 = E^*_{21} \cdot E_{12}$$
$$0 = e(trace(E_{11}E_{12})) + f(trace(E_{21}E_{12}))+$$
$$g(trace(E_{12}E_{12})) + h(trace(E_{22}E_{12})) = f$$

Continuing

$$0 = E^*_{21} \cdot E_{22}$$
$$0 = e(trace(E_{11}E_{21})) + f(trace(E_{21}E_{21}))+$$
$$g(trace(E_{12}E_{21})) + h(trace(E_{22}E_{21})) = h$$

Thus we conclude that

$$E'_{21} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Thus this gives the second column and second row of the Killing form of $B_{\mathbb{C}^2}$:

$$B_{\mathbb{C}^2} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & * & * \\ 0 & 0 & * & * \end{bmatrix}$$

Now we have for the matrix E'_{12}

$$E'_{12} = iE_{11} + jE_{21} + kE_{12} + lE_{22}$$

that is,

$$E'_{12} = \begin{bmatrix} i & k \\ j & l \end{bmatrix}$$

Now

$$\mathcal{B}_{\mathbb{F}^2}(E'_{12}) = E^*_{12}$$

Thus

$$E^*_{12} = \mathcal{B}_{\mathbb{F}^2}(iE_{11} + jE_{21} + kE_{12} + lE_{22}) =$$
$$\mathcal{B}_{\mathbb{F}^2}(iE_{11}) + \mathcal{B}_{\mathbb{F}^2}(jE_{21}) + \mathcal{B}_{\mathbb{F}^2}(kE_{12}) + \mathcal{B}_{\mathbb{F}^2}(lE_{22})$$

We now operate on the basis $(E_{11}, E_{21}, E_{12}, E_{22})$. But we already know $trace(E_{11}E_{12}) = 0$, $trace(E_{21}E_{12}) = 1$. Thus we only compute the following:

$$1 = E^*_{12} \cdot E_{21}$$
$$1 = i(trace(E_{11}E_{12})) + j(trace(E_{21}E_{12})) +$$
$$k(trace(E_{12}E_{12})) + l(trace(E_{22}E_{12})) = j$$

Continuing

$$0 = E^*_{12} \cdot E_{22}$$
$$0 = i(trace(E_{11}E_{22})) + j(trace(E_{21}E_{22})) +$$
$$k(trace(E_{12}E_{22})) + l(trace(E_{22}E_{22})) = l$$

Thus we conclude that

$$E'_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

Thus this gives the third column and third row of the Killing form of $B_{\mathbb{C}^2}$:

$$B_{\mathbb{C}^2} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & * \end{bmatrix}$$

We remark that symmetry immediately gave the {2,3} entry in the matrix of the Killing form; and now in calculating the E'_{21} matrix we confirm that entry.

The only entry that is now missing in the Killing form $B_{\mathbb{C}^2}$ is the {4,4} entry which is equal to $trace(E_{44}E_{44}) = trace(E_{44}) = 1$, giving

$$B_{\mathbb{C}^2} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Thus we conclude that

$$E'_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

and the Killing form is

$$B_{\mathbb{F}^2} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Thus we may remark immediately that the Killing form gives us the four matrices:

$$E'_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \; E'_{21} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad E'_{12} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad E'_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

We also observe that

$$E'_{11} = E_{11} \qquad E'_{21} = E_{12} \qquad E'_{12} = E_{21} \qquad E'_{22} = E_{22}$$

We now define the Casimir operator for $\hat{g}$. It is a linear transformation on $\mathbb{F}^2$ in $\hat{g} = \widehat{gl}(\mathbb{F}^2)$ and it is given as follows.

$$C_{\mathbb{F}^2} : \mathbb{F}^2 \longrightarrow \mathbb{F}^2$$
$$v \longmapsto C_{\mathbb{F}^2}(v) := (E_{11}E'_{11} + E_{21}E'_{21} + E_{12}E'_{12} + E_{22}E'_{22})(v)$$

(The motivation for this formula will be given later in this chapter.)

Given the identifications above, we see that

$$C_{\mathbb{F}^2} = E_{11}E_{11} + E_{21}E_{12} + E_{12}E_{21} + E_{22}E_{22} = E_{11} + E_{22} + E_{11} + E_{22} = 2E_{11} + 2E_{22}$$

Writing this as a matrix, we have

$$C_{\mathbb{F}^2} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

We can immediately see that the trace of $C_{\mathbb{F}^2}$ is 4, which, we also note, is the dimension of $\hat{g}$. Also, since $C_{\mathbb{F}^2}$ is a scalar matrix, it is in the center of $\hat{g}$. In other words $C_{\mathbb{F}^2}$ commutes with every element of $\hat{g}$ in the sense that for every X in $\hat{g}$ and every v in $\mathbb{F}^2$

$$C_{\mathbb{F}^2}(Xv) = X(C_{\mathbb{F}^2}(v))$$

Later we will also prove that the Casimir operator is independent of the choice of basis for its definition and thus only depends on $\hat{g}$ in $\widehat{gl}(\mathbb{F}^2)$.

For our next example, we will take the simple Lie algebra $\widehat{sl}(2,\mathbb{C})$ contained in $\widehat{gl}(2,\mathbb{C})$. We seek a representation ρ in $\widehat{sl}(2,\mathbb{C})$ of the canonical 3-dimensional simple Lie algebra $\hat{a}_1$ that is defined as follows. [For a brief summary of the facts we are using about Lie algebras (which we are not developing in this set of Notes) refer to Appendix 2, p. 265 and following.] We recall that $\hat{a}_1$ has a basis $\{h, e, f\}$ such that its brackets are

$$[h,e] = 2e \qquad [h,f] = -2f \qquad [e,f] = h.$$

Thus we assume that we have a representation ρ of $\hat{a}_1$ in $\widehat{gl}(2,\mathbb{C})$. Since $\hat{a}_1$ is simple, we know that the Killing form $\mathcal{B}_{\mathbb{C}^2}$ of $\rho(\hat{a}_1)$ is non-degenerate, and thus it will give the $2x2$ matrices which form a basis for the representation of $\hat{a}_1$ in $\widehat{gl}(2,\mathbb{C})$.

We calculate the Killing form $\mathcal{B}_{\mathbb{C}^2}$ of $\rho(\hat{a}_1)$ exactly as we did for the Lie algebra treated in the previous example. Since $\hat{a}_1$ is 3-dimensional, we have three $2x2$ matrices with $trace = 0$ to calculate. We map the basis $\{h, e, f\}$ of $\hat{a}_1$ to $trace = 0$ matrices $\{H, E, F\}$ in $\widehat{gl}(2,\mathbb{C})$, giving

$$H = \begin{bmatrix} a & c \\ b & -a \end{bmatrix} \qquad E = \begin{bmatrix} d & f \\ e & -d \end{bmatrix} \qquad F = \begin{bmatrix} g & i \\ h & -g \end{bmatrix}$$

Now the Killing form $\mathcal{B}_{\mathbb{C}^2}$ is given by

$$\begin{aligned} \mathcal{B}_{\mathbb{C}^2} : \widehat{sl}(2,\mathbb{C}) \times \widehat{sl}(2,\mathbb{C}) &\longrightarrow \mathbb{C} \\ (X,Y) &\longmapsto \mathcal{B}_{\mathbb{C}^2}(X,Y) := trace(X \circ Y) \end{aligned}$$

We compute these traces and get the first column of matrix of the Killing form:

$$\begin{gathered} 1 = a(trace(HH)) + b(trace(EH)) + c(trace(FH) = \\ a(trace(E_{11} - E_{22})(E_{11} - E_{22}))+ \\ b(trace(E_{12}(E_{11} - E_{22})) + c(trace(E_{21}(E_{11} - E_{22})) = \\ a(trace(E_{11}E_{11}) - trace(E_{11}E_{22}) - trace(E_{22}E_{11}) + trace(E_{22}E_{22}))+ \\ b(trace(E_{12}E_{11}) - trace(E_{12}E_{22})) + c(trace(E_{21}E_{11}) - trace(E_{21}E_{22})) = \\ a(1 - 0 - 0 + 1) + b(0 - 0) + c(0 - 0) = 2a + 0b + oC = 2a \end{gathered}$$

$$0 = a(trace(HE)) + b(trace(EE)) + c(trace(FE) =$$
$$a(trace(E_{11} - E_{22})(E_{12} + b(trace(E_{12}E_{12} + c(trace(E_{21}E_{12} =$$
$$a(trace(E_{11}(E_{12} - E_{22})) + b(trace(E_{12}E_{12} + c(trace(E_{21}E_{12} =$$
$$a(trace(E_{11}E_{12}) - trace(E_{11}E_{22})) + b(trace(E_{12}E_{12}) + c(trace(E_{21}E_{12}) =$$
$$a(0 - 0) + b(0) + c(1) = 0a + 0b + c = c$$

$$0 = a(trace(HF)) + b(trace(EF)) + c(trace(FF) =$$
$$a(trace(E_{11} - E_{22})(E_{21} + b(trace(E_{21}E_{12} + c(trace(E_{21}E_{21} =$$
$$a(trace(E_{11}(E_{21} - E_{11})(E_{21} + b(trace(E_{21}E_{12} + c(trace(E_{21}E_{21} =$$
$$a(trace(E_{11}E_{21}) - trace(E_{11}E_{21})) + b(trace(E_{21}E_{12}) + c(trace(E_{21}E_{21}) =$$
$$a(0 - 0) + b(1) + c(0) = 0a + b + 0c = b$$

giving $a = \frac{1}{2}$, $b = 0$, $c = 0$. Thus we have the first column of the matrix for the Killing form of $\mathcal{B}_{\mathbb{C}^2}$:

$$\mathcal{B}_{\mathbb{C}^2} = \begin{bmatrix} \frac{1}{2} & * & * \\ 0 & * & * \\ 0 & * & * \end{bmatrix}$$

and thus the H matrix is

$$H = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{bmatrix}$$

Continuing we now seek the matrix E:

$$E = \begin{bmatrix} d & f \\ e & -d \end{bmatrix}$$

Also we know that the Killing form is symmetric and thus we have

$$\mathcal{B}_{\mathbb{C}^2} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{bmatrix}$$

This gives $d = 0$. Thus in order to complete the second column we need only calculate the $(2, 2)$ term e and the $(3, 2)$ term f:

$$1 = d(trace(HE)) + e(trace(EE)) + f(trace(FE) =$$
$$d(trace(E_{11}E_{12} - E_{22}E_{12})) + e(trace(E_{12}E_{21})) + f(trace(E_{21}E_{12})) =$$
$$d(trace(E_{11}E_{12})) - d(trace(E_{11}E_{22})) + e(trace(E_{12}E_{12})) + f(trace(E_{21}E_{12})) =$$
$$d(0 - 0) + e(0) + f(1) = 0d + 0e + f = f$$

$$0 = d(trace(HF)) + e(trace(EF)) + f(trace(FF)) =$$
$$d(trace(E_{11}E_{21} - E_{22}E_{21})) + e(trace(E_{12}E_{21})) + f(trace(E_{21}E_{21})) =$$
$$d(trace(E_{11}E_{21}) - trace(E_{11}E_{21})) + e(trace(E_{21}E_{12})) + f(trace(E_{21}E_{21})) =$$
$$d(0 - 0) + e(1) + f(0) = 0d + e + 0f = e$$

This gives $f = 1$, $e = 0$, and the second column of the matrix for the Killing form of $\mathcal{B}_{\mathbb{C}^2}$:

$$\mathcal{B}_{\mathbb{C}^2} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 0 & * \\ 0 & 1 & * \end{bmatrix}$$

and the E matrix

$$E = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Continuing we now seek the matrix F:

$$F = \begin{bmatrix} g & i \\ h & -g \end{bmatrix}$$

Now symmetry of the Killing form gives

$$\mathcal{B}_{\mathbb{C}^2} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & * \end{bmatrix}$$

This gives $g = 0$ and $h = 1$ and thus, in order to complete the third column, we need only calculate the $(3, 3)$ term i in the Killing form.

$$1 = g(trace(HF)) + h(trace(EF)) + i(trace(FF)) =$$
$$g(trace(E_{11}E_{21} - E_{22}E_{21})) + h(trace(E_{12}E_{21})) + i(trace(E_{21}E_{21})) =$$
$$g(trace(E_{11}E_{21}) - trace(E_{22}E_{21})) + h(trace(E_{12}E_{21})) + i(trace(E_{21}E_{21})) =$$
$$g(0 - 0) + h(1) + i(0) = 0g + h + 0i = h$$

We see however that the above calculation evaluates h once again, which value we already know is equal to 1. Thus to compute i we go back to the calculation

$$0 = g(trace(HE)) + h(trace(EE)) + i(trace(FE)) =$$
$$g(trace(E_{11}E_{12} - E_{22}E_{12})) + h(trace(E_{12}E_{12})) + i(trace(E_{21}E_{12})) =$$
$$g(trace(E_{11}E_{12}) - trace(E_{22}E_{12})) + h(trace(E_{12}E_{12})) + i(trace(E_{21}E_{12})) =$$
$$g(0 - 0) + h(0) + i(1) = 0g + 0h + i = i$$

giving $i = 0$, and the third column of the matrix for the Killing form of $\mathcal{B}_{\mathbb{C}^2}$:

$$\mathcal{B}_{\mathbb{C}^2} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

and the F matrix

$$F = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

Now {H,E,F} is a basis for the set of matrices of trace 0 [called the Special Linear Algebra: $\widehat{sl}(2,\mathbb{C})$] . It is contained in the General Linear Algebra $\widehat{gl}(2,\mathbb{C})$ and has dimension 3. These matrices, of course, have the form:

$$\begin{bmatrix} a & c \\ b & -a \end{bmatrix}$$

The image of the basis $\{h, e, f\}$ of $\hat{a}_1$ is $\{H, E, F\}$ in $\widehat{sl}(2,\mathbb{C})$. It is given by the three matrices which we calculated above:

$$H = \tfrac{1}{2}(E_{11} - E_{22}) = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{bmatrix} \qquad E = E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$F = E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

We check the brackets to verify that there is an isomorphism of Lie algebras:

$$[H, E] = [\tfrac{1}{2}(E_{11} - E_{22}), E_{12}] = \tfrac{1}{2}(E_{11} - E_{22})E_{12} - E_{12}(\tfrac{1}{2})(E_{11} - E_{22}) = \tfrac{1}{2}E_{12} + \tfrac{1}{2}E_{12} = E_{12}$$

$$[H, F] = [\tfrac{1}{2}(E_{11} - E_{22}), E_{21}] = \tfrac{1}{2}(E_{11} - E_{22})E_{21} - E_{21}(\tfrac{1}{2})(E_{11} - E_{22}) = -\tfrac{1}{2}E_{21} - \tfrac{1}{2}E_{21} = -E_{21}$$

$$[E, F] = [E_{12}, E_{21}] = E_{12}E_{21} - E_{21}E_{12} = E_{11} - E_{22} = 2H$$

Now recall again that $\hat{a}_1$ has a basis $\{h, e, f\}$ such that its brackets are

$$[h, e] = 2e \qquad [h, f] = -2f \qquad [e, f] = h.$$

Thus we see that we do have an isomorphism of Lie algebras if we define $H' := 2H$. This gives

$$[H', E] = [2H, E] = 2[H, E] = 2E_{12} = 2E$$
$$[H', F] = [2H, F] = 2[H, F] = -2E_{21} = -2F$$
$$[E, F] = 2H = H'$$

Thus the basis in $\widehat{sl}(2,\mathbb{C})$ maps isomorphically onto the basis in $\hat{a}_1$ that is $\{H', E, F\} = \{2H, E, F\}$.

However when we wish to calculate the Casimir operator, we will need once again the dualized basis of $\{2H, E, F\}$, and thus we seek the dual basis $\{H^*, E^*, F^*\}$ in $\widehat{sl}(2,\mathbb{C})^*$.

Thus we have

$$H^*(2H) = 1 \text{ and } H^*(E) = 0 \text{ and } H^*(F) = 0$$
$$E^*(2H) = 0 \text{ and } E^*(E) = 1 \text{ and } E^*(F) = 0$$
$$F^*(2H) = 0 \text{ and } F^*(E) = 0 \text{ and } F^*(F) = 1$$

Now since the Killing form $B_{\mathbb{C}^2}$ restricted to $\widehat{sl}(2,\mathbb{C})$ is non-degenerate — $\widehat{sl}(2,\mathbb{C})$ being a simple Lie algebra — we have the bijective map

$$\mathcal{B}_{\mathbb{C}^2} : \widehat{sl}(2,\mathbb{C}) \longrightarrow \widehat{sl}(2,\mathbb{C})^*$$

We seek the matrices in $\widehat{sl}(2,\mathbb{C})$ which correspond to H^*, E^*, F^* by $\mathcal{B}^{-1}_{\mathbb{C}^2}$. By using the inverse of the bijective map $\mathcal{B}_{\mathbb{C}^2}$, we will name these matrices as follows.

$$\mathcal{B}^{-1}_{\mathbb{C}^2}(H^*) := H' \qquad \mathcal{B}^{-1}_{\mathbb{C}^2}(E^*) := E' \qquad \mathcal{B}^{-1}_{\mathbb{C}^2}(F^*) := F'$$

We now determine the content of these matrices. For the matrix H', we first write it using the basis $\{2H, E, F\}$, giving

$$2H = (E_{11} - E_{22}) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad E = E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$F = E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

and thus the Killing form is

$$\mathcal{B}_{\mathbb{C}^2} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Thus

$$H' = a(2H) + cE + bF$$

that is,

$$H' = \begin{bmatrix} a & c \\ b & -a \end{bmatrix}$$

Now

$$\mathcal{B}_{\mathbb{C}^2}(H') = H^*$$

Thus

$$\begin{gathered} H^* = \mathcal{B}_{\mathbb{C}^2}(a(2H) + cE + bF) = \\ \mathcal{B}_{\mathbb{C}^2}(2.aH) + \mathcal{B}_{\mathbb{C}^2}(cE) + \mathcal{B}_{\mathbb{C}^2}(bF) \end{gathered}$$

We now operate on the basis $\{2H, E, F\}$.

$$1 = H^* \cdot 2H = \mathcal{B}_{\mathbb{C}^2}(a(2H) \cdot 2H + \mathcal{B}_{\mathbb{C}^2}(cE) \cdot 2H + \mathcal{B}_{\mathbb{C}^2}(bF) \cdot 2H$$

By definition of $\mathcal{B}_{\mathbb{C}^2}$ we translate these expressions to $B_{\mathbb{C}^2}$ and traces.

Thus

$$\begin{gathered} 1 = aB_{\mathbb{C}^2}(2H, 2H) + cB_{\mathbb{C}^2}(E, 2H) + bB_{\mathbb{C}^2}(F, 2H) = \\ a(trace(2H \circ 2H)) + c(trace(E \circ 2H)) + b(trace(F \circ 2H)) = \\ a(trace((E_{11} - E_{22}) \circ (E_{11} - E_{22}))) + c(trace(E_{12} \circ (E_{11} - E_{22}))) + \\ b(trace(E_{21} \circ (E_{11} - E_{22}))) = \\ a(trace(E_{11} + E_{22})) + c(trace(-E_{12})) + b(trace(E_{21})) = a(1+1) = 2a \end{gathered}$$

We conclude that $a = \frac{1}{2}$.

Continuing we have

$$0 = H^* \cdot E = \mathcal{B}_{\mathbb{C}^2}(2aH) \cdot E + \mathcal{B}_{\mathbb{C}^2}(cE) \cdot E + \mathcal{B}_{\mathbb{C}^2}(bF) \cdot E$$

Thus

$$\begin{gathered} 0 = aB_{\mathbb{C}^2}(2H, E) + cB_{\mathbb{C}^2}(E, E) + bB_{\mathbb{C}^2}(F, E) = \\ a(trace(2H \circ E)) + c(trace(E \circ E)) + b(trace(F \circ E)) = \\ a(trace((E_{11} - E_{22}) \circ (E_{12}))) + c(trace(E_{12} \circ E_{12}))) + b(trace(E_{21} \circ E_{12}))) = \\ a(trace(E_{12})) + c(trace(0)) + b(trace(E_{22})) = b(1) = b \end{gathered}$$

We conclude that $b = 0$.

Continuing we have

$$0 = H^* \cdot F = \mathcal{B}_{\mathbb{C}^2}(2aH) \cdot F + \mathcal{B}_{\mathbb{C}^2}(cE) \cdot F + \mathcal{B}_{\mathbb{C}^2}(bF) \cdot F$$

Thus we have

$$\begin{gathered} 0 = aB_{\mathbb{C}^2}(2H, F) + cB_{\mathbb{C}^2}(E, F) + bB_{\mathbb{C}^2}(F, F) = \\ a(trace(2H \circ F)) + c(trace(E \circ F)) + b(trace(F \circ F)) = \\ a(trace((E_{11} - E_{22}) \circ (E_{21}))) + c(trace(E_{12} \circ E_{21}))) + b(trace(E_{21} \circ E_{21}))) = \\ a(trace(-E_{21})) + c(trace(E_{11})) + b(trace(0)) = c(1) = c \end{gathered}$$

We conclude that $c = 0$.

Thus we have calculated H':

$$H' = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

We observe again that we are essentially calculating the Killing form $B_{\mathbb{C}^2}$ with respect to the basis $\{2H, E, F\}$. Recall that the *Killing form* B_V of a Lie algebra $\hat{g}$ contained in $\widehat{gl}(V)$ over a field $\mathbb{F}$ of characteristic 0 is defined in the following manner. After choosing a basis for V, then the following traces define B_V:

$$\begin{aligned} B_V : \hat{g} \times \hat{g} &\longrightarrow \mathbb{F} \\ (X, Y) &\longmapsto B_V(X, Y) := trace(X \circ Y) \end{aligned}$$

Now in the above calculations we have computed these traces:

$$\begin{gathered} a(trace(2H \circ 2H)) + c(trace(E \circ 2H)) + b(trace(F \circ 2H)) = a(2) + c(0) + b(0) \\ a(trace(2H \circ E)) + c(trace(E \circ E)) + b(trace(F \circ E)) = a(0) + c(0) + b(1) \\ a(trace(2H \circ F)) + c(trace(E \circ F)) + b(trace(F \circ F)) = a(0) + c(1) + b(0) \end{gathered}$$

giving the following matrix for $B_{\mathbb{C}^2}$:

$$B_{\mathbb{C}^2} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Thus we can read immediately

$$E' = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad F' = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

We observe that

$$H' = 2H \qquad E' = F \qquad F' = E$$

We now define the Casimir operator for $\widehat{sl}(2,\mathbb{C})$. It is a linear transformation on $\mathbb{C}^2$ in $\widehat{gl}(2,\mathbb{C})$ and is given as follows:

$$C_{\mathbb{C}^2} : \mathbb{C}^2 \longrightarrow \mathbb{C}^2$$
$$v \longmapsto C_{\mathbb{C}^2}(v) := (HH' + EE' + FF')(v)$$

Given the identifications above, we see that

$$C_{\mathbb{C}^2} = H(2H) + EF + FE = \tfrac{1}{2}(E_{11} - E_{22})(E_{11} - E_{22}) + E_{12}E_{21} + E_{21}E_{12} = \tfrac{1}{2}(E_{11} + E_{22}) + E_{11} + E_{22} = \tfrac{3}{2}E_{11} + \tfrac{3}{2}E_{22}$$

Writing this as a matrix, we have

$$C_{\mathbb{C}^2} = \begin{bmatrix} \frac{3}{2} & 0 \\ 0 & \frac{3}{2} \end{bmatrix}$$

We can immediately see that the trace of $C_{\mathbb{C}^2}$ is 3, which is the dimension of $\widehat{sl}(2,\mathbb{C})$. Also, since $C_{\mathbb{C}^2}$ is a scalar matrix, it is in the center of $\widehat{gl}(2,\mathbb{C})$. In other words, $C_{\mathbb{C}^2}$ commutes with every element of $\widehat{gl}(2,\mathbb{C})$ in the sense that for every X in $\widehat{gl}(2,\mathbb{C})$ and every v in $\mathbb{C}^2$

$$C_{\mathbb{C}^2}(Xv) = X(C_{\mathbb{C}^2}(v))$$

We also remark that $C_{\mathbb{C}^2}$ is not an element of $\widehat{sl}(2,\mathbb{C})$ for its trace is not 0. [As we mentioned before, later we will also prove that the Casimir operator is independent of a choice of basis for its definition and thus only depends on $\widehat{sl}(2,\mathbb{C})$ in $\widehat{gl}(2,\mathbb{C})$.]

For our third example, we want to examine a more complicated simple Lie algebra, yet one whose structure is simple enough so that the calculations are still doable. We choose the eight-dimensional simple complex Lie algebra with code symbol $\hat{a}_2$. Since it is simple, it is also semisimple. It has a basis $(h_1, h_2, e_1, e_2, e_3, f_1, f_2, f_3)$ with the following brackets:

$$[h_1, h_2] = 0 \quad [h_1, e_1] = 2e_1 \quad [h_1, e_2] = e_2 \quad [h_1, e_3] = -e_3 \quad [h_1, f_1] = -2f_1$$
$$[h_1, f_2] = -f_2 \quad [h_1, f_3] = f_3 \quad [h_2, e_1] = -e_1 \quad [h_2, e_2] = e_2$$
$$[h_2, e_3] = 2e_3 \quad [h_2, f_1] = f_1 \quad [h_2, f_2] = -f_2 \quad [h_2, f_3] = -2f_3$$
$$[e_1, e_2] = 0 \quad [e_1, e_3] = e_2 \quad [e_1, f_1] = h_1 \quad [e_1, f_2] = -f_3 \quad [e_1, f_3] = 0$$
$$[e_2, e_3] = 0 \quad [e_2, f_1] = -e_3 \quad [e_2, f_2] = h_1 + h_2 \quad [e_2, f_3] = e_1 \quad [e_3, f_1] = 0$$
$$[e_3, f_2] = f_1 \quad [e_3, f_3] = h_2 \quad [f_1, f_2] = 0 \quad [f_1, f_3] = -f_2 \quad [f_2, f_3] = 0$$

We choose a representation of $\hat{a}_2$ in the 3-dimensional complex linear space $\mathbb{C}^3$. We give $\mathbb{C}^3$ its canonical basis (e_1, e_2, e_3). We write the complex Lie algebra $\widehat{gl}(\mathbb{C}^3) = \widehat{gl}(3,\mathbb{C})$ as 3x3 complex matrices with respect to the canonical basis (E_{ij}) in $\widehat{gl}(\mathbb{C}^3)$. The representation map ρ takes $\hat{a}_2$ onto $\widehat{sl}(3,\mathbb{C})$ in $\widehat{gl}(\mathbb{C}^3)$, giving:

$$\begin{aligned}
h_1 &\longmapsto \rho(h_1) = H_1 = E_{11} - E_{22} \\
h_2 &\longmapsto \rho(h_2) = H_2 = E_{22} - E_{33} \\
e_1 &\longmapsto \rho(e_1) = E_{12} \\
e_2 &\longmapsto \rho(e_2) = E_{13} \\
e_3 &\longmapsto \rho(e_3) = E_{23} \\
f_1 &\longmapsto \rho(f_1) = E_{21} \\
f_2 &\longmapsto \rho(f_2) = E_{31} \\
f_3 &\longmapsto \rho(f_3) = E_{32}
\end{aligned}$$

We see immediately the $(H_1, H_2, E_{12}, E_{13}, E_{23}, E_{21}, E_{31}, E_{32})$ are independent in $\widehat{gl}(\mathbb{C}^3)$, and thus we have an 8-dimensional subspace in the 9-dimensional complex Lie algebra $\widehat{gl}(\mathbb{C}^3)$. We calculate its brackets and using the above morphism, we compare them with the brackets in $\hat{a}_2$:

$$[H_1, H_2] = [E_{11} - E_{22}, E_{22} - E_{33}] = -E_{22} + E_{22} = 0$$

$$[h_1, h_2] = 0$$

$$* * * * * * * * * **$$

$$[H_1, E_{12}] = [E_{11} - E_{22}, E_{12}] = E_{12} + E_{12} = 2E_{12}$$

$$[h_1, e_1] = 2e_1$$

$$e_1 \longmapsto \rho(e_1) = E_{12}$$

$$* * * * * * * * * **$$

$$[H_1, E_{13}] = [E_{11} - E_{22}, E_{13}] = E_{13}$$

$$[h_1, e_2] = e_2$$

$$e_2 \longmapsto \rho(e_2) = E_{13}$$

$$* * * * * * * * * **$$

$$[H_1, E_{23}] = [E_{11} - E_{22}, E_{23}] = -E_{23}$$

$$[h_1, e_3] = -e_3$$

$$e_3 \longmapsto \rho(e_3) = E_{23}$$

$$* * * * * * * * * **$$

$$[H_1, E_{21}] = [E_{11} - E_{22}, E_{21}] = -E_{21} - E_{21} = -2E_{21}$$

$$[h_1, f_1] = -2f_1$$

$$f_1 \longmapsto \rho(f_1) = E_{21}$$

$$* * * * * * * * * **$$

$$[H_1, E_{31}] = [E_{11} - E_{22}, E_{31}] = -E_{31}$$

$$[h_1, f_2] = -f_2$$

$$f_2 \longmapsto \rho(f_2) = E_{31}$$

$$* * * * * * * * * **$$

$$[H_1, E_{32}] = [E_{11} - E_{22}, E_{32}] = E_{32}$$

$$[h_1, f_3] = f_3$$

$$f_3 \longmapsto \rho(f_3) = E_{32}$$

$$* * * * * * * * * **$$

$$[H_2, E_{12}] = [E_{22} - E_{33}, E_{12}] = -E_{12}$$
$$[h_2, e_1] = -e_1$$
$$e_1 \longmapsto \rho(e_1) = E_{12}$$
* * * * * * * * * **
$$[H_2, E_{13}] = [E_{22} - E_{33}, E_{13}] = E_{13}$$
$$[h_2, e_2] = e_2$$
$$e_2 \longmapsto \rho(e_2) = E_{13}$$
* * * * * * * * * **
$$[H_2, E_{23}] = [E_{22} - E_{33}, E_{23}] = E_{23} + E_{23} = 2E_{23}$$
$$[h_2, e_3] = 2e_3$$
$$e_3 \longmapsto \rho(e_3) = E_{23}$$
* * * * * * * * * **
$$[H_2, E_{21}] = [E_{22} - E_{33}, E_{21}] = E_{21}$$
$$[h_2, f_1] = f_1$$
$$f_1 \longmapsto \rho(f_1) = E_{21}$$
* * * * * * * * * **
$$[H_2, E_{31}] = [E_{22} - E_{33}, E_{31}] = -E_{31}$$
$$[h_2, f_2] = -f_2$$
$$f_2 \longmapsto \rho(f_2) = E_{31}$$
* * * * * * * * * **
$$[H_2, E_{32}] = [E_{22} - E_{33}, E_{32}] = -E_{32} - E_{32} = -2E_{32}$$
$$[h_2, f_3] = -2f_3$$
$$f_3 \longmapsto \rho(f_3) = E_{32}$$
* * * * * * * * * **
$$[E_{12}, E_{13}] = 0$$
$$[e_1, e_2] = 0$$
* * * * * * * * * **
$$[E_{12}, E_{23}] = E_{13}$$
$$[e_1, e_3] = e_2$$
$$e_2 \longmapsto \rho(e_2) = E_{13}$$
* * * * * * * * * **
$$[E_{12}, E_{21}] = E_{11} - E_{22} = H_1$$
$$[e_1, f_1] = h_1$$
$$h_1 \longmapsto \rho(h_1) = H_1$$
* * * * * * * * * **
$$[E_{12}, E_{31}] = -E_{32}$$
$$[e_1, f_2] = -f_3$$
$$f_3 \longmapsto \rho(f_3) = E_{32}$$
* * * * * * * * * **
$$[E_{12}, E_{32}] = 0$$
$$[e_1, f_3] = 0$$
* * * * * * * * * **

$$[E_{13}, E_{23}] = 0$$
$$[e_2, e_3] = 0$$
* * * * * * * * * **
$$[E_{13}, E_{21}] = -E_{23}$$
$$[e_2, f_1] = -e_3$$
$$e_3 \longmapsto \rho(e_3) = E_{23}$$
* * * * * * * * * **
$$[E_{13}, E_{31}] = H_1 + H_2$$
$$[e_2, f_2] = h_1 + h_2$$
$$h_1 \longmapsto \rho(h_1) = H_1$$
$$h_2 \longmapsto \rho(h_2) = H_2$$
* * * * * * * * * **
$$[E_{13}, E_{32}] = E_{12}$$
$$[e_2, f_3] = e_1$$
$$e_1 \longmapsto \rho(e_1) = E_{12}$$
* * * * * * * * * **
$$[E_{23}, E_{21}] = 0$$
$$[e_3, f_1] = 0$$
* * * * * * * * * **
$$[E_{23}, E_{31}] = E_{21}$$
$$[e_3, f_2] = f_1$$
$$f_1 \longmapsto \rho(f_1) = E_{21}$$
* * * * * * * * * **
$$[E_{23}, E_{32}] = H_2$$
$$[e_3, f_3] = h_2$$
$$h_2 \longmapsto \rho(h_2) = H_2$$
* * * * * * * * * **
$$[E_{21}, E_{31}] = 0$$
$$[f_1, f_2] = 0$$
* * * * * * * * * **
$$[E_{21}, E_{32}] = -E_{31}$$
$$[f_1, f_3] = -f_2$$
$$f_2 \longmapsto \rho(f_2) = E_{31}$$
* * * * * * * * * **
$$[E_{31}, E_{32}] = 0$$
$$[f_2, f_3] = 0$$
* * * * * * * * * **

Thus we see that ρ takes brackets in $\hat{a}_2$ to brackets in $\widehat{gl}(\mathbb{C}^3)$, and we have a representation of $\hat{a}_2$ in $\mathbb{C}^3$. Of course, this image of $\hat{a}_2$ by ρ in $\widehat{gl}(\mathbb{C}^3)$ is $\widehat{sl}(3,\mathbb{C})$, the 3x3 complex matrices with trace zero.

We now want to dualize the basis $(H_1, H_2, E_{12}, E_{13}, E_{23}, E_{21}, E_{31}, E_{32})$, i.e., we want a basis $((H_1)^*, (H_2)^*, (E_{12})^*, (E_{13})^*, (E_{23})^*, (E_{21})^*, (E_{31})^*, (E_{32})^*)$ in $\widehat{sl}(3,\mathbb{C})^*$ such that

$$\begin{gathered}(H_1)^*(H_1) = 1 \text{ and } (H_1)^* \text{ operating on all 7 other basis vectors} = 0\\ (H_2)^*(H_2) = 1 \text{ and } (H_2)^* \text{ operating on all 7 other basis vectors} = 0\\ (E_{12})^*(E_{12}) = 1 \text{ and } (E_{12})^* \text{ operating on all 7 other basis vectors} = 0\\ \text{etc.}\end{gathered}$$

Now since the Killing form $B_{\mathbb{C}^3}$ restricted to $\widehat{sl}(3,\mathbb{C})$ is non-degenerate — again $\widehat{sl}(3,\mathbb{C})$ being a simple Lie algebra — we have the bijective map

$$\mathcal{B}_{\mathbb{C}^3} : \widehat{sl}(3,\mathbb{C}) \longrightarrow \widehat{sl}(3,\mathbb{C})^*$$

We seek the matrices in $\widehat{sl}(3,\mathbb{C})$ which correspond to dual basis vectors

$$((H_1)^*, (H_2)^*, (E_{12})^*, (E_{13})^*, (E_{23})^*, (E_{21})^*, (E_{31})^*, (E_{32})^*)$$

by $\mathcal{B}_{\mathbb{C}^3}^{-1}$. By using the inverse of the bijective map $\mathcal{B}_{\mathbb{C}^3}$, we will name these matrices as follows.

$$\begin{array}{ll}\mathcal{B}_{\mathbb{C}^3}^{-1}((H_1)^*) := H_1' & \mathcal{B}_{\mathbb{C}^3}^{-1}((H_2)^*) := H_2'\\ \mathcal{B}_{\mathbb{C}^3}^{-1}((E_{12})^*) := E_{12}' & \mathcal{B}_{\mathbb{C}^3}^{-1}((E_{13})^*) := E_{13}'\\ \mathcal{B}_{\mathbb{C}^3}^{-1}((E_{23})^*) := E_{23}' & \mathcal{B}_{\mathbb{C}^3}^{-1}((E_{21})^*) := E_{21}'\\ \mathcal{B}_{\mathbb{C}^3}^{-1}((E_{31})^*) := E_{31}' & \mathcal{B}_{\mathbb{C}^3}^{-1}((E_{32})^*) := E_{32}'\end{array}$$

We now determine the content matrices $(H_1', H_2', E_{12}', E_{13}', E_{23}', E_{21}', E_{31}', E_{32}')$ in $\widehat{sl}(3,\mathbb{C})$. We write these matrices in terms of the basis

$$(H_1 = E_{11} - E_{22}, H_2 = E_{22} - E_{33}, E_{12}, E_{13}, E_{23}, E_{21}, E_{31}, E_{32})$$

If we let A_i be any one of these matrices, we have

$$A_i = r_iH_1 + s_iH_2 + b_iE_{21} + c_iE_{31} + d_iE_{32} + f_iE_{12} + g_iE_{13} + h_iE_{23}$$
$$A_i = r_i(E_{11} - E_{22}) + s_i(E_{22} - E_{33}) + b_iE_{21} + c_iE_{31} + d_iE_{32} + f_iE_{12} + g_iE_{13} + h_iE_{23}$$

which gives

$$A_i = \begin{bmatrix} r_i & f_i & g_i \\ b_i & -r_i + s_i & h_i \\ c_i & d_i & -s_i \end{bmatrix}$$

We note that these matrices have trace 0, and thus are in $\widehat{sl}(3,\mathbb{C})$.

Now we know that the Killing form of $\widehat{sl}(3,\mathbb{C})$ written with respect to the above basis gives an $8x8$ non-singular symmetric matrix whose entries are $A_{ij},\ i,j = 1,\cdots,8$, and whose values are defined as $B_{\mathbb{C}^3}(A_{ij}, A_{kl}) := trace(A_{ij} \circ A_{kl})$.

For the first column we have:

$$
\begin{aligned}
&(\mathcal{B}_{\mathbb{C}^3}(H_1))(H_1) = B_{\mathbb{C}^3}(H_1, H_1) = trace(H_1 \circ H_1) = \\
&\qquad trace((E_{11} - E_{22}) \circ (E_{11} - E_{22})) = trace(E_{11} + E_{22}) = 2 \\
&(\mathcal{B}_{\mathbb{C}^3}(H_2))(H_1) = B_{\mathbb{C}^3}(H_2, H_1) = trace(H_2 \circ H_1) = \\
&\qquad trace((E_{22} - E_{33}) \circ (E_{11} - E_{22})) = trace(-E_{22}) = -1 \\
&(\mathcal{B}_{\mathbb{C}^3}(E_{21}))(H_1) = B_{\mathbb{C}^3}(E_{21}, H_1) = trace(E_{21} \circ H_1) = \\
&\qquad trace(E_{21} \circ (E_{11} - E_{22})) = trace(E_{21}) = 0 \\
&(\mathcal{B}_{\mathbb{C}^3}(E_{31}))(H_1) = B_{\mathbb{C}^3}(E_{31}, H_1) = trace(E_{31} \circ H_1) = \\
&\qquad trace(E_{31} \circ (E_{11} - E_{22})) = trace(E_{31}) = 0 \\
&(\mathcal{B}_{\mathbb{C}^3}(E_{32}))(H_1) = B_{\mathbb{C}^3}(E_{32}, H_1) = (trace(E_{32} \circ H_1) = \\
&\qquad trace(E_{32} \circ (E_{11} - E_{22})) = trace(-E_{32}) = 0 \\
&(\mathcal{B}_{\mathbb{C}^3}(E_{12}))(H_1) = B_{\mathbb{C}^3}(E_{12}, H_1) = trace(E_{12} \circ H_1) = \\
&\qquad trace((E_{12}) \circ (E_{11} - E_{22}))) = trace(-E_{12}) = 0 \\
&(\mathcal{B}_{\mathbb{C}^3}(E_{13}))(H_1) = B_{\mathbb{C}^3}(E_{13}, H_1) = trace(E_{13} \circ H_1) = \\
&\qquad trace((E_{13}) \circ (E_{11} - E_{22})) = (trace(0)) = 0 \\
&(\mathcal{B}_{\mathbb{C}^3}(E_{23}))(H_1) = B_{\mathbb{C}^3}(E_{23}, H_1) = trace(E_{23} \circ H_1) = \\
&\qquad trace(E_{23} \circ (E_{11} - E_{22})) = trace(0)) = 0
\end{aligned}
$$

For the second column we have:

By symmetry we have

$$
\begin{aligned}
&(\mathcal{B}_{\mathbb{C}^3}(H_1))(H_2) = -1 \\
&(\mathcal{B}_{\mathbb{C}^3}(H_2))(H_2) = B_{\mathbb{C}^3}(H_2, H_2) = trace(H_2 \circ H_2) = \\
&\qquad trace((E_{22} - E_{33}) \circ (E_{22} - E_{33})) = trace(E_{22} + E_{33}) = 2 \\
&(\mathcal{B}_{\mathbb{C}^3}(E_{21}))(H_2) = B_{\mathbb{C}^3}(E_{21}, H_2) = trace(E_{21} \circ H_2) = \\
&\qquad trace(E_{21} \circ (E_{22} - E_{33})) = trace(0) = 0 \\
&(\mathcal{B}_{\mathbb{C}^3}(E_{31}(H_2) = B_{\mathbb{C}^3}(E_{31}, H_2) = trace(E_{31} \circ H_2) = \\
&\qquad trace(E_{31} \circ (E_{22} - E_{33})) = trace(0) = 0 \\
&(\mathcal{B}_{\mathbb{C}^3}(E_{32}))(H_2) = B_{\mathbb{C}^3}(E_{32}, H_2) = trace(E_{32} \circ H_2) = \\
&\qquad trace(E_{32} \circ (E_{22} - E_{33})) = trace(E_{32}) = 0 \\
&(\mathcal{B}_{\mathbb{C}^3}(E_{12}))(H_2) = B_{\mathbb{C}^3}(E_{12}, H_2) = trace(E_{12} \circ H_2) = \\
&\qquad trace(E_{12} \circ (E_{22} - E_{33})) = trace(E_{12}) = 0
\end{aligned}
$$

$(\mathcal{B}_{\mathbb{C}^3}(E_{13}))(H_2) = B_{\mathbb{C}^3}(E_{13}, H_2) = trace(E_{13} \circ H_2) =$
$trace(E_{13} \circ (E_{22} - E_{33})) = trace(-E_{13}) = 0$
$(\mathcal{B}_{\mathbb{C}^3}(E_{23}))(H_2) = B_{\mathbb{C}^3}(E_{23}, H_2) = trace(E_{23} \circ H_2) =$
$trace(E_{23} \circ (E_{22} - E_{33})) = trace(-E_{23}) = 0$

For the third column we have:

By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(H_1)(E_{12}) = 0$
By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(H_2)(E_{12}) = 0$
$(\mathcal{B}_{\mathbb{C}^3}(bE_{21}))(E_{12}) = B_{\mathbb{C}^3}(E_{21}, E_{12}) = trace(E_{21} \circ E_{12}) = (trace(E_{22})) = 1$
$(\mathcal{B}_{\mathbb{C}^3}(E_{31}))(E_{12}) = B_{\mathbb{C}^3}(E_{31}, E_{12}) = trace(E_{31} \circ E_{12}) = (trace(E_{32})) = 0$
$(\mathcal{B}_{\mathbb{C}^3}(E_{32}))(E_{12}) = B_{\mathbb{C}^3}(E_{32}, E_{12}) = trace(E_{32} \circ E_{12}) = (trace(0)) = 0$
$(\mathcal{B}_{\mathbb{C}^3}(E_{12}))(E_{12}) = B_{\mathbb{C}^3}(E_{12}, E_{12}) = trace(E_{12} \circ E_{12}) = (trace(0)) = 0$
$(\mathcal{B}_{\mathbb{C}^3}(E_{13}))(E_{12}) = B_{\mathbb{C}^3}(E_{13}, E_{12}) = trace(E_{13} \circ E_{12}) = (trace(0)) = 0$
$(\mathcal{B}_{\mathbb{C}^3}(hE_{23}))(E_{12}) = B_{\mathbb{C}^3}(E_{23}E_{12}) = trace(E_{23} \circ E_{12}) = (trace(0)) = 0$

For the fourth column we have:

By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(H_1)(E_{13}) = 0$
By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(H_2)(E_{13}) = 0$
By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(E_{12})(E_{13}) = 0$
$(\mathcal{B}_{\mathbb{C}^3}(E_{21}))(E_{13}) = B_{\mathbb{C}^3}(E_{21}, E_{13}) = trace(E_{21} \circ E_{13}) = trace(E_{23}) = 0$
$(\mathcal{B}_{\mathbb{C}^3}(E_{31}))(E_{13}) = B_{\mathbb{C}^3}(E_{31}, E_{13}) = trace(E_{31} \circ E_{13}) = trace(E_{33}) = 1$
$(\mathcal{B}_{\mathbb{C}^3}(E_{32}))(E_{13}) = B_{\mathbb{C}^3}(E_{32}, E_{13}) = trace(E_{32} \circ E_{13}) = trace(0) = 0$
$(\mathcal{B}_{\mathbb{C}^3}(E_{13}))(E_{13}) = B_{\mathbb{C}^3}(E_{13}, E_{13}) = trace(E_{13} \circ E_{13}) = trace(0) = 0$
$(\mathcal{B}_{\mathbb{C}^3}(E_{23}))(E_{13}) = B_{\mathbb{C}^3}(E_{23}, E_{13}) = trace(E_{23} \circ E_{13}) = trace(0) = 0$
?

For the fifth column we have:

By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(H_1)(E_{23}) = 0$
By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(H_2)(E_{23}) = 0$
By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(E_{12})(E_{23}) = 0$
By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(E_{13})(E_{23}) = 0$
$(\mathcal{B}_{\mathbb{C}^3}(E_{23}))(E_{23}) = B_{\mathbb{C}^3}(E_{23}, E_{23}) = trace(E_{23} \circ E_{23}) = trace(0) = 0$
$(\mathcal{B}_{\mathbb{C}^3}(E_{21}))(E_{23}) = B_{\mathbb{C}^3}(E_{21}, E_{23}) = trace(E_{21} \circ E_{23}) = trace(0) = 0$
$(\mathcal{B}_{\mathbb{C}^3}(E_{31}))(E_{23}) = B_{\mathbb{C}^3}(E_{31}, E_{23}) = trace(E_{31} \circ E_{23}) = trace(0) = 0$
$(\mathcal{B}_{\mathbb{C}^3}(E_{32}))(E_{23}) = B_{\mathbb{C}^3}(E_{32}, E_{23}) = trace(E_{32} \circ E_{23}) = trace(E_{33}) = 1$

For the sixth column we have:

By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(H_1)(E_{21}) = 0$
By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(H_2)(E_{21}) = 0$
By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(E_{12})(E_{21}) = 1$
By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(E_{13})(E_{21}) = 0$
By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(E_{23})(E_{21}) = 0$
$(\mathcal{B}_{\mathbb{C}^3}(E_{21}))(E_{21}) = B_{\mathbb{C}^3}(E_{21}, E_{21}) = trace(E_{21} \circ E_{21}) = trace(0) = 0$
$(\mathcal{B}_{\mathbb{C}^3}(E_{31}))(E_{21}) = B_{\mathbb{C}^3}(E_{31}, E_{21}) = trace(E_{31} \circ E_{21}) = trace(0) = 0$
$(\mathcal{B}_{\mathbb{C}^3}(E_{32}))(E_{21}) = B_{\mathbb{C}^3}(E_{32}, E_{21}) = trace(E_{32} \circ E_{21}) = trace(E_{31}) = 0$

For the seventh column we have:

By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(H_1)(E_{31}) = 0$
By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(H_2)(E_{31}) = 0$
By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(E_{12})(E_{31}) = 0$
By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(E_{13})(E_{31}) = 1$
By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(E_{23})(E_{31}) = 0$
By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(E_{21})(E_{31}) = 0$
$(\mathcal{B}_{\mathbb{C}^3}(E_{31}))(E_{31}) = B_{\mathbb{C}^3}(E_{31}, E_{31}) = trace(E_{31} \circ E_{31}) = trace(0) = 0$
$(\mathcal{B}_{\mathbb{C}^3}(E_{32}))(E_{31}) = B_{\mathbb{C}^3}(E_{32}, E_{31}) = trace(E_{32} \circ E_{31}) = trace(0) = 0$

For the eighth column we have:

By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(H_1)(E_{32}) = 0$
By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(H_2)(E_{32}) = 0$
By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(E_{12})(E_{32}) = 0$
By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(E_{13})(E_{32}) = 0$
By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(E_{23})(E_{32}) = 1$
By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(E_{21})(E_{32}) = 0$
By symmetry we have $(\mathcal{B}_{\mathbb{C}^3}(E_{31})(E_{32}) = 0$
$(\mathcal{B}_{\mathbb{C}^3}(E_{32}))(E_{32}) = B_{\mathbb{C}^3}(E_{32}, E_{32}) = trace(E_{32} \circ E_{32}) = trace(0) = 0$

This gives immediately the non-singular symmetric $8x8$ matrix of the Killing form written respect to the ordered basis in $\widehat{sl}(3,\mathbb{C})$:

$$(H_1 = E_{11} - E_{22}, H_2 = E_{22} - E_{33}, E_{12}, E_{13}, E_{23}, E_{21}, E_{31}, E_{32})$$

$$B_{\mathbb{C}^3} = \begin{bmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

We can now write the 8 matrices in $\widehat{sl}(3,\mathbb{C})$ which we are seeking:

$$H_1', H_2', E_{12}', E_{13}', E_{23}', E_{21}', E_{31}', E_{32}'.$$

Now again we let A_i be any one of these 8 matrices.

We know that $(A_i)^* = \mathcal{B}_{\mathbb{C}^3}(A_i)$ maps A_j to δ_{ij}. Now if

$$A_k = r_kH_1 + s_kH_2 + b_kE_{21} + c_kE_{31} + d_kE_{32} + f_kE_{12} + g_kE_{13} + h_kE_{23}$$

then

$$\begin{gathered} \delta_{ij} = \mathcal{B}_{\mathbb{C}^3}(A_i)(A_j) = \\ \mathcal{B}_{\mathbb{C}^3}(r_iH_1 + s_iH_2 + b_iE_{21} + c_iE_{31} + d_iE_{32} + f_iE_{12} + g_iE_{13} + h_iE_{23})(A_j) = \\ r_i\mathcal{B}_{\mathbb{C}^3}(H_1, A_j) + s_i\mathcal{B}_{\mathbb{C}^3}(H_2, A_j) + b_i\mathcal{B}_{\mathbb{C}^3}(E_{21}, A_j) + c_i\mathcal{B}_{\mathbb{C}^3}(E_{31}, A_j) + \\ d_i\mathcal{B}_{\mathbb{C}^3}(E_{32}, A_j) + f_i\mathcal{B}_{\mathbb{C}^3}(E_{12}, A_j) + g_i\mathcal{B}_{\mathbb{C}^3}(E_{13}, A_j) + h_i\mathcal{B}_{\mathbb{C}^3}(E_{23}, A_j) \end{gathered}$$

Thus for the first row of the Killing form $(1, i)$ we have

$$\begin{gathered} \delta_{11} = 1 = \mathcal{B}_{\mathbb{C}^3}(H_1')(H_1) = \\ r_1\mathcal{B}_{\mathbb{C}^3}(H_1, H_1) + s_1\mathcal{B}_{\mathbb{C}^3}(H_2, H_1) + b_1\mathcal{B}_{\mathbb{C}^3}(E_{21}, H_1) + c_1\mathcal{B}_{\mathbb{C}^3}(E_{31}, H_1) + \\ d_1\mathcal{B}_{\mathbb{C}^3}(E_{32}, H_1) + f_1\mathcal{B}_{\mathbb{C}^3}(E_{12}, H_1) + g_1\mathcal{B}_{\mathbb{C}^3}(E_{13}, H_1) + h_1\mathcal{B}_{\mathbb{C}^3}(E_{23}, H_1) \end{gathered}$$

giving

$$1 = 2r_1 - s_1 + 0b_1 + 0c_1 + 0d_1 + 0f_1 + 0g_1 + 0h_1 = 2r_1 - s_1$$

$$\begin{gathered} \delta_{12} = 0 = \mathcal{B}_{\mathbb{C}^3}(H_1')(H_2) = \\ r_1\mathcal{B}_{\mathbb{C}^3}(H_1, H_2) + s_1\mathcal{B}_{\mathbb{C}^3}(H_2, H_2) + b_1\mathcal{B}_{\mathbb{C}^3}(E_{21}, H_2) + c_1\mathcal{B}_{\mathbb{C}^3}(E_{31}, H_2) + \\ d_1\mathcal{B}_{\mathbb{C}^3}(E_{32}, H_2) + f_1\mathcal{B}_{\mathbb{C}^3}(E_{12}, H_2) + g_1\mathcal{B}_{\mathbb{C}^3}(E_{13}, H_2) + h_1\mathcal{B}_{\mathbb{C}^3}(E_{23}, H_2) \end{gathered}$$

giving

$$0 = -r_1 + 2s_1 + 0b_1 + 0c_1 + 0d_1 + 0f_1 + 0g_1 + 0h_1 = -r_1 + 2s_1$$

$$\delta_{13} = 0 = \mathcal{B}_{\mathbb{C}^3}(H_1')(E_{21}) =$$
$$r_1\mathcal{B}_{\mathbb{C}^3}(H_1, E_{21}) + s_1\mathcal{B}_{\mathbb{C}^3}(H_2, E_{21}) + b_1\mathcal{B}_{\mathbb{C}^3}(E_{21}, E_{21}) + c_1\mathcal{B}_{\mathbb{C}^3}(E_{31}, E_{21}) +$$
$$d_1\mathcal{B}_{\mathbb{C}^3}(E_{32}, E_{21}) + f_1\mathcal{B}_{\mathbb{C}^3}(E_{12}, E_{21}) + g_1\mathcal{B}_{\mathbb{C}^3}(E_{13}, E_{21}) + h_1\mathcal{B}_{\mathbb{C}^3}(E_{23}, E_{21})$$

giving

$$0 = 0r_1 + 0s_1 + 0b_1 + 0c_1 + 0d_1 + f_1 + 0g_1 + 0h_1 = f_1$$

$$\delta_{14} = 0 = \mathcal{B}_{\mathbb{C}^3}(H_1', E_{31}) =$$
$$r_1\mathcal{B}_{\mathbb{C}^3}(H_1, E_{31}) + s_1\mathcal{B}_{\mathbb{C}^3}(H_2, E_{31}) + b_1\mathcal{B}_{\mathbb{C}^3}(E_{21}, E_{31}) + c_1\mathcal{B}_{\mathbb{C}^3}(E_{31}, E_{31}) +$$
$$d_1\mathcal{B}_{\mathbb{C}^3}(E_{32}, E_{31}) + f_1\mathcal{B}_{\mathbb{C}^3}(E_{12}, E_{31}) + g_1\mathcal{B}_{\mathbb{C}^3}(E_{13}, E_{31}) + h_1\mathcal{B}_{\mathbb{C}^3}(E_{23}, E_{31})$$

giving

$$0 = 0r_1 + 0s_1 + 0b_1 + 0c_1 + 0d_1 + 0f_1 + g_1 + 0h_1 = g_1$$

$$\delta_{15} = 0 = \mathcal{B}_{\mathbb{C}^3}((H_1', E_{32}) =$$
$$r_1\mathcal{B}_{\mathbb{C}^3}(H_1, E_{32}) + s_1\mathcal{B}_{\mathbb{C}^3}(H_2, E_{32}) + b_1\mathcal{B}_{\mathbb{C}^3}(E_{21}, E_{32}) + c_1\mathcal{B}_{\mathbb{C}^3}(E_{31}, E_{32}) +$$
$$d_1\mathcal{B}_{\mathbb{C}^3}(E_{32}, E_{32}) + f_1\mathcal{B}_{\mathbb{C}^3}(E_{12}, E_{32}) + g_1\mathcal{B}_{\mathbb{C}^3}(E_{13}, E_{32}) + h_1\mathcal{B}_{\mathbb{C}^3}(E_{23}, E_{32})$$

giving

$$0 = 0r_1 + 0s_1 + 0b_1 + 0c_1 + 0d_1 + 0f_3 + 0g_1 + h_1 = h_1$$

$$\delta_{16} = 0 = \mathcal{B}_{\mathbb{C}^3}(H"_1 E_{12}) =$$
$$r_1\mathcal{B}_{\mathbb{C}^3}(H_1, E_{12}) + s_1\mathcal{B}_{\mathbb{C}^3}(H_2, E_{12}) + b_1\mathcal{B}_{\mathbb{C}^3}(E_{21}, E_{12}) + c_1\mathcal{B}_{\mathbb{C}^3}(E_{31}, E_{12}) +$$
$$d_1\mathcal{B}_{\mathbb{C}^3}(E_{32} E_{12}) + f_1\mathcal{B}_{\mathbb{C}^3}(E_{12}, E_{12}) + g_1\mathcal{B}_{\mathbb{C}^3}(E_{13}, E_{12}) + h_1\mathcal{B}_{\mathbb{C}^3}(E_{23} E_{12})$$

giving

$$0 = 0r_1 + 0s_1 + b_1 + 0c_1 + 0d_1 + f_3 + 0g_1 + 0h_1 = b_1$$

$$\delta_{17} = 0 = \mathcal{B}_{\mathbb{C}^3}(H_1', E_{13}) =$$
$$r_1\mathcal{B}_{\mathbb{C}^3}(H_1, E_{13}) + s_1\mathcal{B}_{\mathbb{C}^3}(H_2, E_{13}) + b_1\mathcal{B}_{\mathbb{C}^3}(E_{21}, E_{13}) + c_1\mathcal{B}_{\mathbb{C}^3}(E_{31}, E_{13}) +$$
$$d_1\mathcal{B}_{\mathbb{C}^3}(E_{32}, E_{13}) + f_1\mathcal{B}_{\mathbb{C}^3}(E_{12}, E_{13}) + g_1\mathcal{B}_{\mathbb{C}^3}(E_{13}, E_{13}) + h_1\mathcal{B}_{\mathbb{C}^3}(E_{23}, E_{13})$$

giving

$$0 = 0r_1 + 0s_1 + 0b_1 + c_1 + 0d_1 + 0f_3 + 0g_1 + 0h_1 = c_1$$

$$\delta_{18} = 0 = \mathcal{B}_{\mathbb{C}^3}(H_1', E_{23}) =$$
$$r_1\mathcal{B}_{\mathbb{C}^3}(H_1, E_{23}) + s_1\mathcal{B}_{\mathbb{C}^3}(H_2, E_{23}) + b_1\mathcal{B}_{\mathbb{C}^3}(E_{21}, E_{23}) + c_1\mathcal{B}_{\mathbb{C}^3}(E_{31}, E_{23}) +$$
$$d_1\mathcal{B}_{\mathbb{C}^3}(E_{32}, E_{23}) + f_1\mathcal{B}_{\mathbb{C}^3}(E_{12}, E_{23}) + g_1\mathcal{B}_{\mathbb{C}^3}(E_{13}, E_{23}) + h_1\mathcal{B}_{\mathbb{C}^3}(E_{23}, E_{23})$$

giving

$$0 = 0r_1 + 0s_1 + 0b_1 + 0c_1 + d_1 + 0f_3 + 0g_1 + 0h_1 = d_1$$

Thus we have for H_1':

$$1 = 2r_1 - s_1;\ 0 = -r_1 + 2s_1;\ 0 = b_1;\ 0 = c_1;\ 0 = d_1;\ 0 = f_1;\ 0 = g_1;\ 0 = h_1$$

or

$$\begin{array}{cccc} r_1 = \frac{2}{3} & s_1 = \frac{1}{3} & 0 = b_1 & 0 = c_1 \\ 0 = d_1 & 0 = f_1 & 0 = g_1 & 0 = h_1 \end{array}$$

giving

$$H_1' = \tfrac{2}{3}H_1 + \tfrac{1}{3}H_2 = \tfrac{2}{3}(E_{11} - E_{22}) + \tfrac{1}{3}(E_{22} - E_{33}) = \tfrac{2}{3}E_{11} - \tfrac{1}{3}E_{22} - \tfrac{1}{3}E_{33}$$

Its matrix is

$$H_1' = \begin{bmatrix} \frac{2}{3} & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & -\frac{1}{3} \end{bmatrix}$$

Obviously H_1' has trace 0 .

Repeating for H_2' **we have**:

$$0 = 2r_2 - s_2;\ 1 = -r_2 + 2s_2;\ 0 = b_2;\ 0 = c_2;\ 0 = d_2;\ 0 = f_2;\ 0 = g_2;\ 0 = h_2$$

or

$$\begin{array}{cccc} r_2 = \frac{1}{3} & s_2 = \frac{2}{3} & 0 = b_2 & 0 = c_2 \\ 0 = d_2 & 0 = f_2 & 0 = g_2 & 0 = h_2 \end{array}$$

giving

$$H_2' = \tfrac{1}{3}H_1 + \tfrac{2}{3}H_2 = \tfrac{1}{3}(E_{11} - E_{22}) + \tfrac{2}{3}(E_{22} - E_{33}) = \tfrac{1}{3}E_{11} + \tfrac{1}{3}E_{22} - \tfrac{2}{3}E_{33}$$

Its matrix is

$$H_2' = \begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & -\frac{2}{3} \end{bmatrix}$$

Also H_2' has trace 0 .

Repeating for E_{21}' **we have**:

$$0 = 2r_3 - s_3;\ 0 = -r_3 + 2s_3;\ 1 = b_3;\ 0 = c_3;\ 0 = d_3;\ 0 = f_3;\ 0 = g_3;\ 0 = h_3$$

or

$$\begin{array}{cccc} r_3 = 0 & s_3 = 0 & 1 = b_3 & 0 = c_3 \\ 0 = d_3 & 0 = f_3 & 0 = g_3 & 0 = h_3 \end{array}$$

giving

$$E'_{21} = E_{21}$$

Its matrix is

$$E_{21} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Also E_{21} has trace 0 .

Repeating for E'_{31} we have:

$$0 = 2r_4 - s_4;\ 0 = -r_4 + 2s_4;\ 0 = b_4;\ 1 = c_4;\ 0 = d_4;\ 0 = f_4;\ 0 = g_4;\ 0 = h_4$$

or

$$\begin{array}{cccc} r_4 = 0 & s_4 = 0 & 0 = b_4 & 1 = c_4 \\ 0 = d_4 & 0 = f_4 & 0 = g_4 & 0 = h_4 \end{array}$$

giving

$$E'_{31} = E_{31}$$

Its matrix is

$$E_{31} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Also E_{31} has trace 0 .

Repeating for E'_{32} we have:

$$0 = 2r_5 - s_5;\ 0 = -r_5 + 2s_5;\ 0 = b_5;\ 0 = c_5;\ 1 = d_5;\ 0 = f_5;\ 0 = g_5;\ 0 = h_5$$

or

$$\begin{array}{cccc} r_5 = 0 & s_5 = 0 & 0 = b_5 & 0 = c_5 \\ 1 = d_5 & 0 = f_5 & 0 = g_5 & 0 = h_5 \end{array}$$

giving

$$E'_{32} = E_{32}$$

Its matrix is

$$E_{32} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Also E'_{32} has trace 0 .

Repeating for E'_{12} we have:

$$0 = 2r_6 - s_6;\ 0 = -r_6 + 2s_6;\ 0 = b_6;\ 0 = c_6;\ 0 = d_6;\ 1 = f_6;\ 0 = g_6;\ 0 = h_6$$

or

$$\begin{matrix} r_6 = 0 & s_6 = 0 & 0 = b_6 & 0 = c_6 \\ 0 = d_6 & 1 = f_6 & 0 = g_6 & 0 = h_6 \end{matrix}$$

giving

$$E'_{12} = E_{12}$$

Its matrix is

$$E_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Also E_{12} has trace 0 .

Repeating for E'_{13} we have:

$$0 = 2r_7 - s_7;\ 0 = -r_7 + 2s_7;\ 0 = b_7;\ 0 = c_7;\ 0 = d_7;\ 0 = f_7;\ 1 = g_7;\ 0 = h_7$$

or

$$\begin{matrix} r_7 = 0 & s_7 = 0 & 0 = b_7 & 0 = c_7 \\ 0 = d_7 & 0 = f_7 & 1 = g_7 & 0 = h_7 \end{matrix}$$

giving

$$E'_{13} = E_{13}$$

Its matrix is

$$E_{13} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Also E_{13} has trace 0 .

Repeating for E'_{23} we have:

$$0 = 2r_8 - s_8;\ 0 = -r_8 + 2s_8;\ 0 = b_8;\ 0 = c_8;\ 0 = d_8;\ 0 = f_8;\ 0 = g_8;\ 1 = h_8$$

or

$$\begin{array}{cccc} r_8 = 0 & s_8 = 0 & 0 = b_8 & 0 = c_8 \\ 0 = d_8 & 0 = f_8 & 0 = g_8 & 1 = h_8 \end{array}$$

giving

$$E'_{23} = E_{23}$$

Its matrix is

$$E_{23} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Also E_{23} has trace 0 .

We observe

$$H'_1 = \tfrac{2}{3}H_1 + \tfrac{1}{3}H_2 \qquad H'_2 = \tfrac{1}{3}H_1 + \tfrac{2}{3}H_2$$
$$E'_{12} = E_{21} \quad E'_{21} = E_{12} \quad E'_{13} = E_{31} \quad E'_{31} = E_{13} \quad E'_{23} = E_{23} \quad E'_{32} = E_{32}$$

We now define the Casimir operator for $\widehat{sl}(3,\mathbb{C})$. It is a linear transformation on $\mathbb{C}^3$ in $\widehat{gl}(3,\mathbb{C})$ and is given as follows.

$$\begin{gathered}
C_{\mathbb{C}^3} : \mathbb{C}^3 \longrightarrow \mathbb{C}^3 \\
v \longmapsto C_{\mathbb{C}^3}(v) := \\
(H_1H'_1 + H_2H'_2 + E_{12}E'_{12} + E_{21}E'_{21} + E_{13}E'_{13} + E_{31}E'_{31} + E_{23}E'_{23} + E_{23}E'_{32})(v) = \\
(H_1(\tfrac{2}{3}H_1 + \tfrac{1}{3}H_2) + H_2(\tfrac{1}{3}H_1 + \tfrac{2}{3}H_2)+ \\
E_{12}E_{21} + E_{21}E_{12} + E_{13}E_{31} + E_{31}E_{13} + E_{23}E_{32} + E_{32}E_{23})(v) = \\
(\tfrac{2}{3}(E_{11} - E_{22})(E_{11} - E_{22}) + \tfrac{1}{3}(E_{11} - E_{22})(E_{22} - E_{33})+ \\
\tfrac{1}{3}(E_{22} - E_{33})(E_{11} - E_{22}) + \tfrac{2}{3}(E_{22} - E_{33})(E_{22} - E_{33})+ \\
E_{12}E_{21} + E_{21}E_{12} + E_{13}E_{31} + E_{31}E_{13} + E_{23}E_{32} + E_{32}E_{23})(v) = \\
(\tfrac{2}{3}(E_{11} + E_{22}) + \tfrac{1}{3}(-E_{22}) + \tfrac{1}{3}(-E_{22}) + \tfrac{2}{3}(E_{22} + E_{33})+ \\
E_{11} + E_{22} + E_{11} + E_{33} + E_{22} + E_{33})(v) = \\
(\tfrac{2}{3}(E_{11} + E_{22} + E_{33}) + E_{11} + E_{22} + E_{11} + E_{33} + E_{22} + E_{33})(v) = \\
(\tfrac{8}{3}E_{11} + \tfrac{8}{3}E_{22} + \tfrac{8}{3}E_{33})(v)
\end{gathered}$$

Writing this as a matrix, we have

$$C_{\mathbb{C}^3} = \begin{bmatrix} \frac{8}{3} & 0 & 0 \\ 0 & \frac{8}{3} & 0 \\ 0 & 0 & \frac{8}{3} \end{bmatrix}$$

We see that the trace of $C_{\mathbb{C}^3}$ is 8, which is the dimension of $\widehat{sl}(3,\mathbb{C})$. Also since $C_{\mathbb{C}^3}$ is a scalar matrix, it is in the center of $\widehat{gl}(3,\mathbb{C})$. In other words $C_{\mathbb{C}^3}$ commutes with every element of $\widehat{gl}(3,\mathbb{C})$ in the sense that for every X in $\widehat{gl}(3,\mathbb{C})$ and every v in $\mathbb{C}^3$

$$C_{\mathbb{C}^3}(Xv) = X(C_{\mathbb{C}^3}(v))$$

We also remark that $C_{\mathbb{C}^3}$ is not an element of $\widehat{sl}(3,\mathbb{C})$ for its trace is not 0. And as we mentioned before, later we will prove that the Casimir operator is independent of a choice of basis for its definition and thus only depends on $\widehat{sl}(3,\mathbb{C})$ in $\widehat{gl}(3,\mathbb{C})$.

With the help of these examples we can now examine the Casimir operator in general. We observe that we can define the Casimir operator for a Lie subalgebra $\hat{g}$ of $\widehat{gl}(V)$ whose Killing form on a linear space V restricted to $\hat{g}$ is non-degenerate. We do so as follows. [We remark again that the field $\mathbb{F}$ can be either $\mathbb{R}$ or $\mathbb{C}$ even though all the examples given here were only over $\mathbb{R}$.] Let the dimension of V be n and the dimension of $\hat{g}$ be r. Because the Killing form is non-degenerate, we know that this form defines an isomorphism of the linear space $\hat{g}$ with its dual $\hat{g}^*$.

$$\mathcal{B}_V : \hat{g} \longrightarrow \hat{g}^*.$$

First we choose a basis for $\hat{g}$ in $\widehat{gl}(V)$: $(A_1, \cdots, A_r)$, giving r matrices in the n^2-dimensional linear space of matrices $\widehat{gl}(V)$. Now we dualize this basis by means of the above map, giving us a dual basis for $\hat{g}^*$: $(A_1^*, \cdots, A_r^*)$. This means that the map $\mathcal{B}_V$ takes A_i to A_i^*. where A_i^* is defined in $\hat{g}^*$ by

$$(A_i^*)(A_j) = \delta_{ij}$$

Now A_i^* comes from some matrix A_i' in $\hat{g}$ by $\mathcal{B}_V$, i.e.,

$$(\mathcal{B}_V^{-1})(A_i^*) = A_i'$$

where $(A_1', \cdots, A_r')$ are the matrices in $\hat{g}$ which represent the duals. This means

$$(A_i^*)(A_j) = ((\mathcal{B}_V)(A_i'))(A_j) = B_V(A_i', A_j) = trace(A_i' \circ A_j) = \delta_{ij}$$

[We remark that here we are restricting the form B_V defined on $\widehat{gl}(V)\times\widehat{gl}(V)$ to $\hat{g} \times \hat{g}$.]

Now the *Casimir operator* is a linear transformation on V in $\widehat{gl}(V)$ and is defined as follows:

$$C_V : V \longrightarrow V$$

$$v \longmapsto C_V(v) := (\sum_{i=1}^{r} A_i \circ A_i')(v)$$

We remark that this definition just depends on our being able to dualize a linear subspace of matrices by a non-degenerate bilinear form defined on those matrices.

First, we show that the Casimir operator is independent of the basis chosen for $\hat{g}$. Since A_i' is a matrix in $\hat{g}$, we can express it in the A-basis for $\hat{g}$.

$$A_i' = \sum_k a_{ik} A_k$$

We want to determine this rxr matrix $\mathbf{a} = [a_{ik}]$.

$$\delta_{ij} = A_i^*(A_j) = \mathcal{B}_V(A_i') \cdot A_j = B_V(A_i', A_j) = trace(A_i' \circ A_j) = trace(A_j \circ A_i')$$

Now

$$A_j \circ A_i' = A_j(\sum_k a_{ik} A_k) = \sum_k a_{ik} A_j \circ A_k$$

Thus

$$\delta_{ij} = trace(\sum_k a_{ik} A_j \circ A_k) = \sum_k a_{ik} trace(A_j \circ A_k) = \sum_k a_{ik} B_V(A_j, A_k)$$

Writing B_V in the A-basis, we have

$$B_V(A_j, A_k) = (B_{V(A)})_{jk} = [A_j]_A^T B_{V(A)} [A_k]_A$$

where the term $[A_j]_A^T B_{V(A)} [A_k]_A$ is a product of three matrices written in the A-basis; $[A_j]_A^T$ is a 1xr row matrix corresponding to the dual A_j^* [thus giving the canonical row matrix $[e_j]$; $B_{V(A)}$ is the rxr matrix representing the bilinear form B_V; and $[A_k]_A$, an rx1 column matrix corresponding to the A_k basis vector [thus giving the canonical column matrix $[e_k]$]. The product is a 1x1 matrix corresponding to the jk entry in the matrix $[B_{V(A)})_{jk}]$.

Continuing and using the symmetry of the $[B_{V(A)})_{jk}]$ matrix, we have

$$\delta_{ij} = \sum_k a_{ik} B_V(A_j, A_k) = \sum_k a_{ik}(B_{V(A)})_{jk} = \sum_k a_{ik}(B_{V(A)})_{kj} = (\mathbf{a}B_{V(A)})_{ij}$$

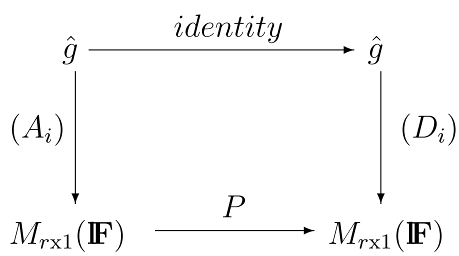
Thus we have

$$\mathbf{a}B_{V(A)} = I_r$$

and we see that the matrix $\mathbf{a}$ is the inverse of the matrix for B_V written with respect to the A-basis.

$$\mathbf{a} = B_{V(A)}^{-1}$$

We now change bases from an A-basis to, say, a D-basis.



Thus, from the commutative diagram, we see that the change of basis matrix is labelled $P = [P_{ij}]$, and in terms of this matrix the j-th column is A_j written in the D-basis is

$$[A_j]_D = P[A_j]_A \qquad \text{or} \qquad A_j = \sum_i P_{ij} D_i$$

We now write B_V in the A-basis in terms of D-basis.

$$(B_{V(A)})_{ij} = B_V(A_i, A_j) = B_V(\sum_k P_{ki} D_k, \sum_l P_{lj} D_l) =$$

$$\sum_{kl} P_{ki} P_{lj} B_V(D_k, D_l) = \sum_{kl} P_{ki} P_{lj} (B_{V(D)})_{kl} = \sum_{kl} P^T_{ik} (B_{V(D)})_{kl} P_{lj}$$

Thus we can conclude that

$$B_{V(A)} = P^T B_{V(D)} P$$

Similarly the D-basis (D_i) in $\hat{g}$ dualizes by the map $\mathcal{B}_V$ to a dual basis (D_i^*) in $\hat{g}^*$, giving the relation

$$D_i^*(D_j) = \delta_{ij}$$

Now D_i^* comes from some matrix D_i' in $\hat{g}$ by $\mathcal{B}_V$, i.e.,

$$(\mathcal{B}_V^{-1})(D_i^*) = D_i'$$

where the (D_i') are the matrices in $\hat{g}$ which represent the duals. If we write these matrices in the basis (D_i), we obtain

$$D_i' = \sum_k d_{ik} D_k$$

Now we know that the matrix $\mathbf{d} = [d_{ik}]$ is the inverse of the matrix representation of B_V in the D-basis:

$$\mathbf{d} = B_{V(D)}^{-1}$$

The Casimir operator defined with respect to the D-basis is

$$C_V := \sum_{i=1}^{r} D_i \circ D_i'$$

We show that these two definitions, one written with respect to the A-basis and other with respect to the D-basis, yield the same result.

$$A_i \circ A_i' = A_i \sum_k a_{ik} A_k = (\sum_j P_{ji} D_j)(\sum_k a_{ik}(\sum_l P_{lk} D_l)) =$$

$$(\sum_j P_{ji} D_j)(\sum_{kl} a_{ik} P_{lk} D_l) = (\sum_j P_{ji} D_j)(\sum_l (\mathbf{a}P^T)_{il} D_l) = \sum_{jl} P_{ji} D_j (\mathbf{a}P^T)_{il} D_l$$

Now we know that

$$B_{V(A)} = P^T B_{V(D)} P \qquad \mathbf{a} = B_{V(A)}^{-1} \qquad \mathbf{d} = B_{V(D)}^{-1}$$

Thus

$$\mathbf{a} = B_{V(A)}^{-1} = P^{-1} B_{V(D)}^{-1} P^{T-1} = P^{-1} \mathbf{d} P^{T-1}$$

Continuing

$$A_i \circ A'_i = \sum_{jl} P_{ji} D_j (\mathbf{a}P^T)_{il} D_l = \sum_{jl} P_{ji} D_j (P^{-1}\mathbf{d}P^{T-1}P^T)_{il} D_l =$$

$$\sum_{jl} P_{ji} D_j (P^{-1}\mathbf{d})_{il} D_l = \sum_{jlk} P_{ji} D_j P^{-1}_{ik} \mathbf{d}_{kl} D_l =$$

$$\sum_{jlk} P_{ji} P^{-1}_{ik} D_j \mathbf{d}_{kl} D_l = \sum_{jk} P_{ji} P^{-1}_{ik} D_j D'_k$$

Finally, we obtain:

$$\sum_i A_i \circ A'_i = \sum_{ijk} P_{ji} P^{-1}_{ik} D_j D'_k = \sum_{jk} \delta_{jk} D_j D'_k = \sum_k D_k \circ D'_k$$

Thus we can conclude that the value of the Casimir operator is basis independent. We also can conclude that a Casimir operator can be defined on a linear space of matrices acting on a linear space V on which a non-degenerate Killing form B_V is defined. We know that such a Killing form exists if the linear space of matrices is a semisimple Lie subalgebra $\hat{g}$ of $\widehat{gl}(V)$.

With these assumptions we can also show that the Casimir operator, which is an element of $\widehat{gl}(V)$ but not necessarily in $\hat{g}$, commutes with every matrix X in $\hat{g}$, i.e.,

$$C_V(Xv) = X(C_V v)$$

for all v in V. To prove this we need the fact that the Killing form associates, i.e.,

$$B_V([X,Y],Z) = B_V(X,[Y,Z])$$

We remark that if C_V were in the center of $\widehat{gl}(V)$, then trivially it would commute with all X in $\hat{g}$. In the examples given above this is what occurred. However we are affirming that C_V commutes only with all X in $\hat{g}$, which may be only a proper subset of $\widehat{gl}(V)$, and thus it is not necessarily in the center of $\widehat{gl}(V)$.

Thus we need to show that

$$C_V(Xv) - X(C_V v) = 0$$

for all X in $\hat{g}$ and all v in V. We write C_V in an A-basis:

$$C_V(Xv) - X(C_V v) = (\sum_i^r A_i A_i')(Xv) - X((\sum_i^r A_i A_i')(v) =$$

$$\sum_i^r (A_i A_i' X - X A_i A_i')(v)$$

We add and subtract $\sum_i^r (A_i X A_i')$ to obtain brackets.

$$\sum_i^r (A_i A_i' X - X A_i A_i')(v) =$$

$$(\sum_i^r (A_i A_i' X - A_i X A_i') + \sum_i^r (A_i X A_i' - X A_i A_i'))(v) =$$

$$(\sum_i^r (A_i [A_i', X]) + \sum_i^r ([A_i, X] A_i'))(v)$$

We now express $[A_i', X]$ in terms of the A'-basis, and $[A_i, X]$ in terms of the A-basis.

$$[A_i', X] = \sum_j^r d_{ij} A_j' \qquad [A_i, X] = \sum_j^r c_{ij} A_j$$

Thus we have

$$(\sum_i^r (A_i [A_i', X]) + \sum_i^r ([A_i, X] A_i'))(v) =$$

$$((\sum_{ij}^r d_{ij} A_i A_j') + (\sum_{ij}^r c_{ij} A_j A_i'))(v)$$

Switching indices in the first sum, we have

$$((\sum_{ij}^r d_{ij} A_i A_j') + (\sum_{ij}^r c_{ij} A_j A_i'))(v) =$$

$$((\sum_{ij}^r d_{ji} A_j A_i') + (\sum_{ij}^r c_{ij} A_j A_i'))(v) =$$

$$(\sum_{ij}^r (d_{ji} + c_{ij}) A_j A_i')(v)$$

Thus we have

$$C_V(Xv) - X(C_V v) = (\sum_{ij}^{r}(d_{ji} + c_{ij})A_j A_i')(v)$$

Now we show that $C_V(Xv) - X(C_V v) = 0$ by showing that $(d_{ji} + c_{ij}) = 0$, or that $d_{ji} = -c_{ij}$. To do this we use the associative property of the Killing form.

By this property we know that

$$B_V([A_i, X], A_j') = B_V(A_i, [X, A_j'])$$

We first work on the left side. We have already expressed the bracket $[A_i, X]$ in the A-basis. Thus

$$B_V([A_i, X], A_j') = B_V(\sum_{k}^{r} c_{ik}A_k, A_j') = \sum_{k}^{r} c_{ik}B_V(A_k, A_j')$$

Now we write A_j' in the A-basis, using the matrix $\mathbf{a}$.

$$\sum_{k}^{r} c_{ik}B_V(A_k, A_j') = \sum_{k}^{r} c_{ik}B_V(A_k, \sum_{l}^{r} a_{jl}A_l) =$$

$$\sum_{kl}^{r} c_{ik}a_{jl}B_V(A_k, A_l) = \sum_{kl}^{r} c_{ik}a_{jl}(B_{V(A)})_{kl}$$

Using the symmetry of the Killing form, we have

$$\sum_{kl}^{r} c_{ik}a_{jl}(B_{V(A)})_{kl} = \sum_{kl}^{r} c_{ik}a_{jl}(B_{V(A)})_{lk}$$

We know that $B_{V(A)} = \mathbf{a}^{-1}$. Thus

$$\sum_{kl}^{r} c_{ik}a_{jl}(B_{V(A)})_{lk} = \sum_{kl}^{r} c_{ik}a_{jl}(\mathbf{a}^{-1})_{lk} = \sum_{k}^{r} c_{ik}\delta_{jk} = c_{ij}$$

Thus we have

$$B_V([A_i, X], A_j') = c_{ij}$$

Continuing, we now work on the right side. We have already expressed the bracket $[A'_j, X]$ in the A'-basis. Thus

$$B_V(A_i, [X, A'_j]) = -B_V(A_i, [A'_j.X]) =$$

$$-B_V(A_i, \sum_k^r d_{jk} A'_k) = -\sum_k^r d_{jk} B_V(A_i, A'_k)$$

Now we write A'_k in the A-basis, using the matrix $\mathbf{a}$.

$$-\sum_k^r d_{jk} B_V(A_i, A'_k) = -\sum_k^r d_{jk} B_V(A_i, \sum_l^r a_{kl} A_l) =$$

$$-\sum_{kl}^r d_{jk} a_{kl} B_V(A_i, A_l) = -\sum_{kl}^r d_{jk} a_{kl} (B_{V(A)})_{il}$$

Using the symmetry of the Killing form, we have

$$-\sum_{kl}^r d_{jk} a_{kl} (B_{V(A)})_{il} = -\sum_{kl}^r d_{jk} a_{kl} (B_{V(A)})_{li} =$$

$$-\sum_{kl}^r d_{jk} a_{kl} (\mathbf{a}^{-1})_{li} = -\sum_k^r d_{jk} \delta_{ki} = -d_{ji}$$

Thus we have

$$B_V([A_i, X], A'_j) = -d_{ji}$$

and we can conclude that $c_{ij} = -d_{ji}$. Thus we have our desired result and we can affirm that $C_V(Xv) - X(C_V v) = 0$, or that C_V and X commute for all X in $\hat{g}$.

Finally, we show that the third property of the Casimir operator mentioned above, i.e., the trace the Casimir operator C_V is equal to the dimension of $\hat{g}$. Now we have

$$trace(C_V) = trace(\sum_{i=1}^r A_i \circ A'_i) = \sum_{i=1}^r (trace(A_i \circ A'_i))$$

and

$$trace(A_i \circ A_i') = B_V(A_i, A_i') = B_V(A_i', A_i) = (\mathcal{B}_V(A_i'))(A_i) = A_i^*(A_i) = \delta_{ii}$$

Thus

$$trace(C_V) = \sum_{i=1}^{r} (trace(A_i \circ A_i')) = \sum_{i=1}^{r} \delta_{ii} = r$$

We note that none of the above discussion of the Casimir operator depended on the field of scalars of the Lie algebra $\hat{g}$ except that it must be of characteristic 0. Thus the field of scalars of $\hat{g}$ can be either $\mathbb{R}$ or $\mathbb{C}$.

2.15.3 The Complete Reducibility of a Representation of a Semisimple Lie Algebra. We can now return to giving the proof of the complete reducibility of a representation of a semisimple Lie algebra. We recall the wording of the theorem.

> Let V be a representation of a semisimple Lie algebra $\hat{g}$ and let W be an invariant subspace of $\hat{g}$. Then there exists a subspace W' of V
>
> invariant by $\hat{g}$ which is complementary, i.e., $V = W \oplus W'$.

Here is the situation: we have a linear space V of dimension n, a semisimple Lie algebra $\hat{g}$, and a representation ρ of $\hat{g}$ on V. [This structure is frequently described by saying that V is a $\hat{g}$-module, and the $\hat{g}$-invariant subspace W is a $\hat{g}$-submodule. But we will continue to use the original terminology, i.e., we will call it a representation ρ of $\hat{g}$ on V The reader may run into this alternative terminology and thus we mention it here.]. This means that we have a Lie algebra homomorphism ρ of $\hat{g}$ into the Lie algebra $\widehat{gl}(V)$. Also since Lie algebra homomorphisms carry semisimple Lie algebras to semisimple Lie algebras, we know that $\rho(\hat{g})$ is also semisimple and of dimension $r < n^2$. [We remark that $r \neq n^2$ since $\widehat{gl}(V)$ always has a center, which means $\widehat{gl}(V)$ has a non-zero radical, and thus it is not semisimple.] Let W be a proper subspace of V which is invariant by $\rho(\hat{g})$, i.e., $(\rho(\hat{g}))(W) \subset W$, and $W \neq 0$ and $W \neq V$. [By the way this means that V must be at least two-dimensional.] It therefore seems obvious that we should be taking quotient spaces and using a proof by induction on the dimension n of V.

Thus we begin by forming the quotient space V/W of dimension less than n. First we show that for any invariant proper subspace, the representation ρ of $\hat{g}$ on V induces a representation ρ' of $\hat{g}$ on V/W. Thus for X in $\rho(\hat{g})$ and v in V we have by invariance

$$X(v + W) = X(v) + X(W) \subset X(v) + W$$

and we therefore define $X(v+W) := X(v)+W$. [We review quickly why this procedure is valid. Let v_1 and v_2 belong to the same coset. This means that $v_1 - v_2$ belongs to W. Now $X(v_1+W) \subset X(v_1)+W$ and $X(v_2+W) \subset X(v_2)+W$. Calculating, we have $(X(v_1)+W)-(X(v_2)+W) = (X(v_1)-X(v_2))+W = (X(v_1-v_2))+W = W$, since v_1-v_2 belongs to W and X leaves W invariant. Thus cosets go to cosets by X.]

Now X is linear since

$$X((v_1+W)+(v_2+W)) = X(v_1)+X(W)+X(v_2)+X(W) \subset \\ X(v_1)+W+X(v_2)+W = (X(v_1)+X(v_2))+W$$

and if c is a scalar then

$$X(c(v_1+W)) = c(X(v_1)+X(W)) \subset c(X(v_1)+W)$$

We conclude that $\rho'(\hat{g})$ is in $\widehat{gl}(V/W)$.

Now we must show that brackets go to brackets. For if $\rho(x) = X$ and $\rho(y) = Y$, we have

$$\rho([x,y])(v+W) = [\rho(x),\rho(y)](v+W) = [X,Y](v+W) = \\ X(Y(v+W)) - Y(X(v+W)) = \\ X(Y(v))+X(Y(W))-Y(X(v))-Y(X(W)) \subset X(Y(v))+W-Y(X(v))+W = \\ X(Y(v)) - Y(X(v)) + W = [X,Y](v)+W$$

and we have

$$\rho([x,y])(v+W) = \rho([x,y])(v) + \rho([x,y])(W) \subset \rho([x,y])(v)+W = \\ [\rho(x),\rho(y)](v)+W = [X,Y](v)+W$$

Hence we have a representation ρ' of $\hat{g}$ on the linear space V/W that is induced by ρ.

Since W is a proper subspace of V, we know that the dimension of V/W is less than the dimension of V. Let us say that the dimension of W is $m < n$. First we assume that W is itself reducible. (Now in this case we can set up the process of induction.)

This means that there is subspace W' in W, not equal to W or to 0, that is invariant by $\rho(\hat{g})$. Thus V must now be at least three-dimensional.

Here is where it would be good to look at an example to give us a feeling for what has to be done, Now we know that we have an eight-dimensional simple Lie algebra which we have met before in 2.15.2, namely $\hat{a}_2$. We know that $\hat{a}_2$ has a basis of eight elements $(h_1, h_2, e_1, e_2, e_3, f_1, f_2, f_3)$ whose brackets are given in 2.15.2. And we have a representation of $\hat{a}_2$ on $\mathbb{C}^3$, given by a homomorphism ρ of $\hat{a}_2$ into $\hat{gl}(\mathbb{C}^3)$, where $\rho(\hat{a}_2) = \hat{sl}(3,\mathbb{C})$, the $3x3$ complex trace zero matrices in $\hat{gl}(\mathbb{C}^3)$. We map this basis $(h_1, h_2, e_1, e_2, e_3, f_1, f_2, f_3)$ of $\hat{a}_2$ to eight *trace* 0 matrices in $\hat{gl}(\mathbb{C}^3)$, by the following correspondences:

$$\begin{aligned}
h_1 &\longmapsto \rho(h_1) = H_1 = E_{11} - E_{22} \\
h_2 &\longmapsto \rho(h_2) = H_2 = E_{22} - E_{33} \\
e_1 &\longmapsto \rho(e_1) = E_{12} \\
e_2 &\longmapsto \rho(e_2) = E_{13} \\
e_3 &\longmapsto \rho(e_3) = E_{23} \\
f_1 &\longmapsto \rho(f_1) = E_{21} \\
f_2 &\longmapsto \rho(f_2) = E_{31} \\
f_3 &\longmapsto \rho(f_3) = E_{32}
\end{aligned}$$

This map we know from 2.15.2 is a homomorphism of Lie algebras. A typical matrix in $\hat{gl}(\mathbb{C}^3)$ then is given by

$$c_1H_1 + c_2H_2 + c_3E_{12} + c_4E_{13} + c_5E_{23} + c_6E_{21} + c_7E_{31} + c_8E_{32} =$$

$$\begin{bmatrix} c_1 & c_3 & c_4 \\ c_6 & -c_1 + c_2 & c_5 \\ c_7 & c_8 & -c_2 \end{bmatrix}$$

where the c_i are in $\mathbb{C}$.

In this 3-dimensional example $V = \mathbb{C}^3$, and thus to start the induction we must show that the theorem holds in this case. Thus we must find a two-dimensional subspace W of V which is invariant by $\rho(\hat{a}_2) = \hat{sl}(3,\mathbb{C})$.

++
Here is also where we need a more general 2-dimensional invariant subspace to be our ground case. And we need a proof of the ground case. Thus the induction proof has a gap that we hope our readers can fill in. We ran out of time and inspiration at this point. Assuming that we have such a space and that we have shown that the theorem holds for this case we proceed below on the inductive steps.

We also think that the proof of the ground case, as we said, will strongly resemble the proof offered below for the inductive case(s) but we ran out of time and inspiration for carrying through our thoughts. The reader is invited to fill in the gaps that exist.
++

Assuming that the ground case has been proved, we now look at dimensions bigger than the ground case and we let W' be any invariant subspace of any invariant subspace W of V. We form the linear space V/W'. From above we know that ρ induces a representation ρ' on V/W'. [We remark that we are using the symbol ρ' for the induced representation on any quotient space and thus there should be no ambiguity.] We then proceed to show that W/W' is an invariant subspace of V/W', We start by letting X be in $\rho(\hat{g})$. We then have for w in W that

$$X(w + W') = X(w) + X(W') \subset X(w) + W'$$

which is in W/W' since $X(w)$ is in W. Thus, the induction assumption says that there exists a complementary subspace Z/W' invariant by $\rho'(\hat{g})$ such that $V/W' = W/W' \oplus Z/W'$. This says $V + W' = W + W' + Z + W'$ and thus $V+W' = W+Z+W'$ or $V = W+Z$. It also says that $W/W' \cap Z/W' = 0$ [as a coset] or $(W + W') \cap (Z + W') \subset W'$; and thus $(W \cap Z) + W' \subset W'$, which says that $(W \cap Z) \subset W'$. We seek now an estimate of the size of the dimension of Z. We know that

$$\begin{gathered} n - dimW' = m - dimW' + dimZ - dimW' \\ n = m - dimW' + dimZ \\ dimZ = n - (m - dimW') < n - m < n \end{gathered}$$

Since $W' \subset Z$ is invariant by $\rho(\hat{g})$, another induction argument gives a subspace L of Z invariant by $\rho(\hat{g})$ such that $Z = W' \oplus L$. We have

$$V = W + Z = W + W' + L = W + L$$

Now since $W \cap Z \subset W'$ and $Z = W' \oplus L$, we have that $W \cap (W' \oplus L) \subset W'$. and thus L is not part of W. Therefore $W \cap L = 0$ and this gives $V = W \oplus L$, which is our desired conclusion when the invariant subspace W of V is itself reducible.

However when W is irreducible, i.e., when $\rho(\hat{g})$ leaves no subspace of W invariant except 0 and W, then we do not have a W' nor can an inductive argument as that given above be set up. Moreover, how to proceed from here is not obvious. In fact, a whole new and surprising direction must be taken in this case.

We begin this new direction by first choosing our representation spaces for $\hat{g}$ to be *linear spaces of linear maps* involving V and W.

Since $W \subset V$, we have two related sets of maps of linear spaces involving V and W: $Hom(V, W)$ and $Hom(W, W) = End(W)$. We choose each of these as a representation space of $\hat{g}$, giving us linear spaces of linear maps of dimension nm and m^2 respectively. On each of these we seek to define a $\hat{g}$-representation: a $\rho_1(\hat{g})$ in $\widehat{gl}(Hom(V,W))$ and $\rho_2(\hat{g})$ in $\widehat{gl}(End(W))$.

By hypothesis we already have a $\hat{g}$-representation ρ on V.

$$\begin{aligned} \hat{g} &\longrightarrow \rho(\hat{g}) \subset \widehat{gl}(V) \\ x &\longmapsto \rho(x) = X \end{aligned}$$

Now we want to define a linear map of $\rho_1(\hat{g})$ into $\widehat{gl}(Hom(V,W))$ which preserves the bracket. To do this, we take advantage of the composition of linear maps.

$$\begin{aligned} \hat{g} &\longrightarrow \rho_1(\hat{g}) \in \widehat{gl}(Hom(V,W)) \\ x &\longmapsto \rho_1(x) = X\cdot \end{aligned}$$

$$\begin{aligned} x \longmapsto X\cdot : Hom(V,W) &\longrightarrow Hom(V,W) \\ \phi &\longmapsto X\cdot\phi := -\phi\circ X = -\phi X \end{aligned}$$

[Note that the necessity of the negative sign in this definition is certainly not obvious, but indeed it is needed below to prove that $[X,Y]\cdot = [X\cdot, Y\cdot]$ holds. Note, too, the special use of a dot to denote a special kind of function composition.]

Linearity is obvious. We have for X and Y in $\rho(\hat{g})$ and c a scalar that

$$(X_1 + X_2)\cdot\phi = -\phi(X_1+X_2) = -\phi(X_1) - \phi X_2) = (X_1)\cdot\phi + (X_2)\cdot\phi = (X_1\cdot + X_2\cdot)\cdot\phi$$
$$(cX)\cdot\phi = -\phi(cX) = c(-\phi X) = c(X\cdot\phi) = (c(X\cdot))\phi$$

Showing the bracket preservation is more delicate. There is no way that we can compose two functions in $Hom(V,W)$. But we can use the fact that the composition of functions does associate. We have for X and Y in $\rho(\hat{g})$:

$$\phi \longmapsto [X,Y]\cdot\phi = -\phi[X,Y] = -\phi(XY-YX) = -\phi(XY) + \phi(YX) = (-\phi X)Y + (\phi Y)X = -Y\cdot(-\phi X) - X\cdot(\phi Y) = -Y\cdot(X\cdot\phi) - X\cdot(-Y\cdot\phi) = -(Y\cdot X\cdot)\phi + (X\cdot Y\cdot)\phi = (-Y\cdot X\cdot + X\cdot Y\cdot)\phi = [X\cdot, Y\cdot]\phi$$

And thus we have the desired relation $[X,Y]\cdot = [X\cdot, Y\cdot]$.

In the same manner we have a representation of $\hat{g}$ on $Hom(W,W) = End(W)$. We define a linear map of $\rho_2(\hat{g})$ into $\widehat{gl}(End(W))$ by

$$\begin{aligned} \rho_2(\hat{g}) &\longrightarrow \widehat{gl}(End(W)) \\ X \longmapsto X\cdot : End(W) &\longrightarrow End(W) \\ \psi &\longmapsto X\cdot\psi := -\psi\circ X|_W = -\psi(X|_W) \end{aligned}$$

Note that this definition is valid since for all X in $\rho(\hat{g})$, X restricted to W leaves W invariant. (*It is here where we use the invariance of W by $\rho(\hat{g})$.*) Also, note that we will symbolize both representations ρ_1 and ρ_2 of $\hat{g}$ by ρ', and for X in $\rho(\hat{g})$ we will use, as we did above, the symbol $X\cdot$ for both representations.

Now we define the restriction map σ, a linear map which takes $Hom(V,W)$ into $Hom(W,W)$ and which respects the representation[in the sense that σ commutes with $X\cdot$: $(X\cdot)\sigma = \sigma(X\cdot)$. We define it as follows:

$$\begin{gathered} Hom(V,W) \xrightarrow{\sigma} Hom(W,W) \\ \phi \longmapsto \sigma(\phi) := \phi|_W \end{gathered}$$

Linearity is straightforward.

$$\begin{gathered} \sigma(\phi_1 + \phi_2) = (\phi_1 + \phi_2)|_W = \phi_1|_W + \phi_2|_W = \sigma(\phi_1) + \sigma(\phi_2) \\ \sigma(c\phi) = (c\phi)|_W = c(\phi|_W) = c\sigma(\phi) \end{gathered}$$

(where, of course, c is a scalar).

We now consider the following diagram for the representations:

$$\begin{array}{ccc} Hom(V,W) & \xrightarrow{\ \ \sigma\ \ } & Hom(W,W) \\ {\scriptstyle X\cdot}\downarrow & & \downarrow{\scriptstyle X\cdot} \\ Hom(V,W) & \xrightarrow{\ \ \sigma\ \ } & Hom(W,W) \end{array}$$

$$\begin{gathered} \phi \longmapsto \sigma(\phi) \longmapsto X \cdot \sigma(\phi) = X \cdot \phi|_W = -\phi|_W X|_W \\ \phi \longmapsto X \cdot \phi \longmapsto \sigma(X \cdot \phi) = \sigma(-\phi X) = (-\phi X)|_W = -\phi|_W X|_W \end{gathered}$$

and we see that it is commutative. Thus we can conclude that σ respects the representation, i.e.,

$$(X\cdot)\sigma = \sigma(X\cdot).$$

We now have arrived at the crucial steps in our attempt to show that $W \subset V$ – which is invariant and irreducible by $\rho(\hat{g})$, – has a complementary subset W' which is also invariant and irreducible.

We outline our approach.

Observation 1) Starting with the subspace W, we can form a new *invariant* and *irreducible* subspace of V of dimension $n-1$ which will have a complementary subspace of dimension one and thus will be irreducible, and invariant.

Since $W \subset V$, we have two related sets of maps of linear spaces involving V and W: $Hom(V, W)$ and $Hom(W, W) = End(W)$. We choose each of these as a representation space of $\hat{g}$, giving us linear spaces of linear maps of dimensions nm and m^2 respectively. On each of these we define a $\hat{g}$-representation: $\rho_1(\hat{g})$ in $\widehat{gl}(Hom(V, W))$ and $\rho_2(\hat{g})$ in $\widehat{gl}(End(W))$.

By hypothesis we already have a $\hat{g}$-representation ρ on V.

$$\begin{aligned} \hat{g} &\longrightarrow \rho(\hat{g}) \subset \widehat{gl}(V) \\ x &\longmapsto \rho(x) = X \end{aligned}$$

since we defined a linear map of $\rho_1(\hat{g})$ into $\widehat{gl}(Hom(V, W))$ which preserves the bracket:

$$\begin{aligned} \hat{g} &\longrightarrow \rho_1(\hat{g}) \in \widehat{gl}(Hom(V, W)) \\ x &\longmapsto \rho_1(x) = X\cdot \end{aligned}$$

$$\begin{aligned} x \longmapsto X\cdot : Hom(V, W) &\longrightarrow Hom(V, W) \\ \phi &\longmapsto X \cdot \phi := -\phi \circ X = -\phi X \end{aligned}$$

And thus we have the desired relation $[X, Y]\cdot = [X\cdot, Y\cdot]$. In the same manner we have a representation of $\hat{g}$ on $Hom(W, W) = End(W)$. We defined a linear map of $\rho_2(\hat{g})$ into $\widehat{gl}(End(W))$ by

$$\begin{aligned} \rho_2(\hat{g}) &\longrightarrow \widehat{gl}(End(W)) \\ X \longmapsto X\cdot : End(W) &\longrightarrow End(W) \\ \psi &\longmapsto X \cdot \psi := -\psi \circ X|_W = -\psi(X|_W) \end{aligned}$$

This definition is valid since for all X in $\rho(\hat{g})$, X restricted to W leaves W invariant. *It is here where we use the invariance of W by $\rho(\hat{g})$.* We will symbolize both representations ρ_1 and ρ_2 of $\hat{g}$ by ρ', and for X in $\rho(\hat{g})$ we will use, as above, the symbol $X\cdot$ for both representations.

Observation 2) To identify this new invariant and irreducible subspace of V of dimension $n-1$, we first we choose the identity map I_W in $Hom(W, W)$, which determines a one-dimensional set of transformations $\{cI_W : c \in \mathbb{C}\}$ in $Hom(W, W)$. Again the target space for the restriction map σ is $Hom(W, W)$ and thus $\sigma^{-1}\{(cI_W)\} \subset Hom(V, W)$, and has codimension one.

Now we have already defined the restriction map σ. It is a linear map which takes $Hom(V, W)$ into $Hom(W, W)$

$$\begin{gathered} Hom(V,W) \xrightarrow{\sigma} Hom(W,W) \\ \phi \longmapsto \sigma(\phi) := \phi|_W \end{gathered}$$

and which respects the representation, i.e.,

$$(X\cdot)\sigma = \sigma(X\cdot).$$

Now the target space for the restriction map σ is $Hom(W,W)$ and thus $\sigma^{-1}\{(cI_W)\} \subset Hom(V,W)$, and has codimension one, which gives

$$\sigma^{-1}(\{cI_W\}) = (ker(\sigma|_{\sigma^{-1}(\{cI_W\})})) \oplus \{cI_W\}$$

Observation 3) From all of this we will be able to conclude that the subspace $ker(\sigma|_{\sigma^{-1}(\{cI_W\})})$ of $\sigma^{-1}(\{cI_W\})$ is an invariant subspace.

We remark immediately that the action of $\hat{g}$ on $\{cI_W\}$ does <u>not</u> leave invariant the subspace $\{cI_W\}$:

$$X \cdot cI_W = -cI_W(X|_W) = -cX|_W$$

It still leaves $Hom(W,W)$ invariant but not $\{cI_W\}$. However, this set of endomorphisms $\{cI_W\}$ forms a one-dimensional space, and thus the set of transformations $\widehat{gl}(\{cI_W\})$ is one-dimensional, and brackets in such an algebra are zero. Thus we can conclude that the only representation of a semisimple Lie algebra on a one-dimensional linear space is the zero representation. This gives us the following diagram.

$$\begin{array}{ccc} \sigma^{-1}(\{cI_W\}) & \xrightarrow{\quad\sigma\quad} & \{cI_W\} \\ X\cdot \big\downarrow & & \big\downarrow (X')\cdot \\ \sigma^{-1}(\{cI_W\}) & \xrightarrow{\quad\sigma\quad} & \{0I_W\} \subset \{cI_W\} \end{array}$$

Since $\{cI_W\}$ is one-dimensional and the image of $\sigma^{-1}(\{cI_W\})$, we know that

$$\sigma^{-1}(\{cI_W\}) = ker(\sigma|_{\sigma^{-1}(\{cI_W\})}) \oplus \{cI_W\}$$

Moreover, we affirm that $X\cdot$ takes $\sigma^{-1}(\{cI_W\})$ onto itself. We also recall that the invariance relation, modified as follows, is valid:

$$((X')\cdot)\sigma = \sigma(X\cdot)$$

Observation 4) But knowing that W is irreducible we will be able to also conclude that $ker(\sigma_{\sigma^{-1}(\{cI_W\})})$ of co-dimension 1 *is irreducible.*

Now we move from our original invariant subset W of V to an irreducible representation in the $ker(\sigma_{\sigma^{-1}(\{cI_W\})})$ in the semisimple Lie algebra $\hat{g}$ in $Hom(V, W)$.

Here, too, since $\{cI_W\}$ is one-dimensional and is the image of $\sigma^{-1}(\{cI_W\})$, we know that

$$\sigma^{-1}(\{cI_W\}) = ker(\sigma|_{\sigma^{-1}(\{cI_W\})}) \oplus \{cI_W\}$$

We affirm , as above, that $X\cdot$ takes $\sigma^{-1}(\{cI_W\})$ onto itself since the invariance relation, modified as follows, is valid.

$$((X')\cdot)\sigma = \sigma(X\cdot)$$

This says that for any ϕ in $\sigma^{-1}(\{cI_W\})$ we have

$$\begin{aligned}(X')\cdot(\sigma(\phi)) &= \sigma((X\cdot)(\phi))\\ 0I_W &= \sigma((X\cdot)(\phi))\end{aligned}$$

and we can conclude that $(X\cdot)(\phi)$ is in $ker(\sigma|_{\sigma^{-1}(\{cI_W\})})$, giving us the fact that $\sigma^{-1}(\{cI_W\})$ is invariant by any X in $\rho(\hat{g})$, and indeed that the subspace $ker(\sigma|_{\sigma^{-1}(\{cI_W\})})$ of $\sigma^{-1}(\{cI_W\})$ is an invariant subspace also.

Thus, suppose that $ker(\sigma|_{\sigma^{-1}(\{cI_W\})})$ is reducible. Then there exists a subspace $\mathcal{A}$ of $ker(\sigma|_{\sigma^{-1}(\{cI_W\})})$ such that $\mathcal{A} \neq 0$ and $\mathcal{A} \neq ker(\sigma|_{\sigma^{-1}(\{cI_W\})})$, and $X\cdot(\mathcal{A}) \subset \mathcal{A}$ for all X in $\widehat{gl}(V)$. Now we know that

$$\dim(\sigma^{-1}(\{cI_W\})) = \dim(ker(\sigma|_{\sigma^{-1}(\{cI_W\})})) + \dim(im(\sigma|_{\sigma^{-1}(\{cI_W\})}))$$

Thus we see that the invariant subspace $ker(\sigma|_{\sigma^{-1}(\{cI_W\})})$ of $\sigma^{-1}(\{cI_W\})$ has codimension one. Now we take any linear function $\psi : V \longmapsto W$ in $\sigma^{-1}(\{cI_W\})$ such that it determines a one-dimensional subspace $\{c\psi\}$ complementary to the $(ker(\sigma|_{\sigma^{-1}(\{cI_W\})}))$, that is,

$$\sigma^{-1}(\{cI_W\}) = (ker(\sigma|_{\sigma^{-1}(\{cI_W\})})) \oplus \{c\psi\}$$

Putting all this information together, we can assert the following. We have $\psi(V) = W$ since $\psi|_W = \sigma(\psi) = c_0 I_W$. Now W is a subspace of V. Thus we know that $V = ker(\psi) \oplus W$. Now for all ϕ in $\mathcal{A}$ and all X in $\widehat{gl}(V)$ we know that -$X\cdot\phi = \phi X$ is also in $\mathcal{A}$ and for ϕ' in $ker(\sigma|_{\sigma^{-1}(\{cI_W\})})$ but not in $\mathcal{A}$, -$X\cdot\phi' = \phi' X$ is not in $\mathcal{A}$. This means that there exists a proper subspace W' in W such that $\phi|_{W'} = 0$ and $(\phi X)|_{W'} = 0$, and such that $\phi'|_{W'} = 0$ but such that $(\phi' X)|_{W'} \neq 0$. If not, then $(\phi' X)|_{W'} = 0$ and this would mean that W' would be the same as W, or that $\mathcal{A}$ would be equal to $ker(\sigma|_{\sigma^{-1}(\{cI_W\})})$. We conclude that $X|_{W'} \subset W'$ and thus W is reducible. But by hypothesis W is irreducible, and we can therefore conclude that $ker(\sigma|_{\sigma^{-1}(\{cI_W\})})$ is also irreducible and of co-dimension one.

Observation 5) However, what we want to assert is that it is also true that there exists a map ψ such that the one-dimensional subspace$\{c\psi\}$ is complementary to $ker(\sigma_{\sigma^{-1}(\{cI_W\})})$ and invariant. [Obviously, being one-dimensional, it is irreducible. And surprisingly, in this case we would have proven the theorem!]

Since $\{cI_W\}$ is one-dimensional and the image of $\sigma^{-1}(\{cI_W\})$, we know that

$$\sigma^{-1}(\{cI_W\}) = ker(\sigma|_{\sigma^{-1}(\{cI_W\})}) \oplus \{cI_W\}$$

We affirm that $X\cdot$ takes $\sigma^{-1}(\{cI_W\})$ onto itself. We still know that the invariance relation, modified as follows, is valid:

$$((X')\cdot)\sigma = \sigma(X\cdot)$$

This says that for any ϕ in $\sigma^{-1}(\{cI_W\})$ we have

$$\begin{aligned}(X')\cdot(\sigma(\phi)) &= \sigma((X\cdot)(\phi)) \\ 0I_W &= \sigma((X\cdot)(\phi))\end{aligned}$$

and we can conclude that $(X\cdot)(\phi)$ is in $ker(\sigma|_{\sigma^{-1}(\{cI_W\})})$, giving us the fact that $\sigma^{-1}(\{cI_W\})$ is invariant by any X in $\rho(\hat{g})$, and indeed that the subspace $ker(\sigma|_{\sigma^{-1}(\{cI_W\})})$ of $\sigma^{-1}(\{cI_W\})$ is an invariant subspace also. We would also like to affirm that it is irreducible.

To show this, we suppose that $ker(\sigma|_{\sigma^{-1}(\{cI_W\})})$ is reducible. Then there exists a subspace $\mathcal{A}$ of $ker(\sigma|_{\sigma^{-1}(\{cI_W\})})$ such that $\mathcal{A} \neq 0$ and $\mathcal{A} \neq ker(\sigma|_{\sigma^{-1}(\{cI_W\})})$, and $X \cdot (\mathcal{A}) \subset \mathcal{A}$ for all X in $\widehat{gl}(V)$. Now we know that

$$\dim(\sigma^{-1}(\{cI_W\})) = \dim(ker(\sigma|_{\sigma^{-1}(\{cI_W\})})) + \dim(im(\sigma|_{\sigma^{-1}(\{cI_W\})}))$$

Thus we see that the invariant subspace $ker(\sigma|_{\sigma^{-1}(\{cI_W\})})$ of $\sigma^{-1}(\{cI_W\})$ has codimension one. Now we take any linear function $\psi : V \longmapsto W$ in $\sigma^{-1}(\{cI_W\})$ such that it determines a one-dimensional subspace $\{c\psi\}$ complementary to the $(ker(\sigma|_{\sigma^{-1}(\{cI_W\})}))$, that is,

$$\sigma^{-1}(\{cI_W\}) = (ker(\sigma|_{\sigma^{-1}(\{cI_W\})})) \oplus \{c\psi\}$$

Once again, putting all this information together, as above, we can assert the following. We have $\psi(V) = W$ since $\psi|_W = \sigma(\psi) = c_0 I_W$ and W is a subspace of V. Thus we know that $V = ker(\psi) \oplus W$. Now for all ϕ in $\mathcal{A}$ and all X in $\widehat{gl}(V)$ we know that -$X \cdot \phi = \phi X$ is also in $\mathcal{A}$ and for ϕ' in $ker(\sigma|_{\sigma^{-1}(\{cI_W\})})$ but not in $\mathcal{A}$, $X \cdot \phi' = \phi' X$ is not in $\mathcal{A}$. This means that there exists a proper subspace W' in W such that $\phi|_{W'} = 0$ and $(\phi X)|_{W'} = 0$, and that $\phi'|_{W'} = 0$ but $(\phi' X)|_{W'} \neq 0$. If not, then $(\phi' X)|_{W'} = 0$, which means that W' would be the same as W, or that $\mathcal{A}$ would be equal to $ker(\sigma|_{\sigma^{-1}(\{cI_W\})})$. We conclude that $X|_{W'} \subset W'$ and thus W is reducible. But by hypothesis W is irreducible, and we can conclude that $ker(\sigma|_{\sigma^{-1}(\{cI_W\})})$ is also irreducible, and of co-dimension one.

Thus we have reduced our proof to the case where W is an invariant subspace of V and is also irreducible. This led us to the point where we produced a representation space for $\hat{g}$ of maps $\sigma^{-1}(cI_W)$ which has a decomposition

$$\sigma^{-1}(\{cI_W\}) = (ker(\sigma|_{\sigma^{-1}(\{cI_W\})})) \oplus \{c\psi\}$$

and in which $(ker(\sigma|_{\sigma^{-1}(\{cI_W\})}))$ was an invariant and irreducible subspace of codimension one of $\sigma^{-1}(\{cI_W\})$. Here the map ψ is arbitrary in the one-dimensional space $\{c\psi\}$. Now suppose that in this case we could find a ψ such that $\{c\psi\}$ is also invariant. [Obviously, being one-dimensional, it is irreducible.] This means we would have proven the theorem in this case.

But we shall see that proving the theorem in this case also proves the theorem for the case when the subset W is irreducible and invariant!

In the above proof we showed that $V = ker(\psi) \oplus W$. Let us rename $ker(\psi) = W'$. Now what does it mean for $\{c\psi\}$ to be an invariant subspace of $\sigma^{-1}(\{cI_W\})$? It says that for all X in $\rho(\hat{g})$

$$X \cdot \{c\psi\} \subset \{c\psi\}$$

or for some scalar c'

$$X \cdot \psi = c'\psi$$

This gives

$$(\psi X)(W') = \psi(X(W')) = c'\psi(W') = 0$$

since W' is the kernel of ψ. We can conclude the $X(W')$ is also in $ker(\psi) = W'$, and thus W' is invariant by $\rho(\hat{g})$ – the conclusion we have been seeking. Recall that the subspace W in this discussion was proper, invariant and also irreducible, but of any codimension.

Thus we are now reduced to proving the theorem in the case where we have an irreducible invariant subspace W of codimension one, that is, we need to prove that an irreducible invariant subspace W of codimension one has a complementary one-dimensional invariant subspace. It is here that we need the completely new tool to effect this proof, and that tool is the Casimir operator (which we treated in **2.15.2**).

++
Note that in the 3-dimensional base case cited above, if one has a two-dimensional invariant subspace then the proof given here would apply, it seems, to the base case as well, unless, of course, the base case is assumed in this part of the proof.
++

Observation 6) But proving the theorem in the case given in Observation 5 also proves the

THEOREM: A semisimple Lie algebra has a representation that is completely reducible

for the only case which remains, that is, when the subset W is irreducible and invariant. In this proof we show that $V = ker(\psi) \oplus W$, where $ker(\psi)$ is the invariant subspace of V that we are seeking. Let us rename $ker(\psi) = W'$.

Let us now make the following assumption. We return to our original expression of our theorem in this particular case. Thus if V is any finite dimensional linear space over a field $\mathbb{F}$ and W is any irreducible invariant subspace of codimension one, *let us then assume that under these conditions the theorem is valid, i.e., there exist an invariant subspace W' complementary to W.* In our situation the linear space is $\sigma^{-1}(\{cI_W\})$ and the irreducible invariant subspace of codimension one is $ker(\sigma|_{\sigma^{-1}(\{cI_W\})})$. Thus we can find a ψ such that $\{c\psi\}$ is an invariant subspace of $\sigma^{-1}(\{cI_W\})$ complementary to $ker(\sigma|_{\sigma^{-1}(\{cI_W\})})$.

Now under this assumption we can show that $V = ker(\psi) \oplus W$. Having called $ker(\psi) = W'$, we thus have the conclusion of our theorem: $V = W' \oplus W$.

Now what does it mean for $\{c\psi\}$ to be an invariant subspace of $\sigma^{-1}(\{cI_W\})$? It says that for all X in $\rho(\hat{g})$

$$X \cdot \{c\psi\} \subset \{c\psi\}$$

or that for some scalar c'

$$X \cdot \psi = c'\psi$$

This gives

$$(\psi X)(W') = \psi(X(W')) = c'\psi(W') = 0$$

We can conclude the $X(W')$ is in $ker(\psi) = W'$, and thus W' is invariant by $\rho(\hat{g})$, the conclusion we have been seeking. *Thus in this case we would have proved our Theorem.* Recall that the subspace W in this discussion was proper, invariant and also irreducible, but of any codimension.

But this conclusion is valid only on the assumption that for any irreducible invariant subspace W of *codimension one there exist an invariant subspace W' complementary to W.*

To prove this assumption we need the Casimir operator.

Since $\rho(\hat{g})$ is semisimple, we know that it has a nondegenerate Killing form on V, and thus we can define a Casimir operator C_V on V.

We now want to set up the situation so that we can apply Schur's Lemma [Appendix A.1.13]. We have only one linear space W. We also have $\rho(\hat{g})$ acting irreducibly on W. Now we need C_V to act invariantly on W. But $C_V = \sum_{i=1}^{r} A_i A'_i$ for an arbitrary basis (A_i) in $\rho(\hat{g})$. Since both A_i and A'_i are in $\rho(\hat{g})$, and W invariant by $\rho(\hat{g})$, then $C_V(W) \subset W$. [We remark that even though both A_i and A'_i are in $\rho(\hat{g})$, $A_i A'_i$ is not the bracket product and thus $A_i A'_i$ is not necessarily in $\rho(\hat{g})$.] By the commutativity of the Casimir operator, we know that for all w in W and for all X in $\rho(\hat{g})$, $X(C_V(w)) = C_V(X(w))$. Since $\rho(\hat{g})$ acts irreducibly on W, Schur's Lemma says that C_V is either an isomorphism or the zero map. But since the trace of C_V is equal to $r \neq 0$, C_V is certainly not the zero map. Thus we know that C_V is an isomorphism on W. And thus we can conclude that $im(C_V) = W$.

Since the co-dimension of W is one, if we can show that $ker(C_V) \neq 0$, we would know that it is one-dimensional, and we would have $V = W \oplus ker(C_V)$. We now go to the quotient space V/W. This is one-dimensional. Now from the invariance of W we have shown above that $\rho(\hat{g})$ induces a representation $\rho'(\hat{g})$ of $\hat{g}$ on V/W. Now since this representation is one-dimensional and since $\hat{g}$ is semisimple, we know that this must be the zero representation in V/W. Thus $\rho'(\hat{g}) = 0$. This means

$$C_V(v + W) = (\sum_{i}^{r} A_i A'_i)(v + W) \subset W$$

since A', A are both in $\rho(\hat{g})$, and the sum $\sum_i^r A_i A_i'$ is also in $\rho(\hat{g})$. Thus this quantity acting on any coset must go into the zero coset, which is W. We can conclude then that C_V acting on any element of V must have its image in W, and thus $ker(C_V)$ is non-empty and, of course, is one-dimensional. [We remark that in this argument we needed once again that $\rho(\hat{g})$ be semisimple, since we used the fact that $\rho'(\hat{g}) = \rho'(D(\hat{g})) = [\rho'(\hat{g}), \rho'(\hat{g})] = 0$. We also needed semisimplicity to conclude that the Killing form on V was nondegenerate. Now above we showed that $\widehat{gl}(V)$, even though it is not semisimple, did have a nondegenerate Killing form on V, and thus had a Casimir operator. But since $\widehat{gl}(V)$ is not semisimple, it is possible to have images in V/W which are not zero, but whose brackets would be zero, as required for the one-dimensional representation. Thus in this situation we could not conclude that the Casimir operator, even though defined for $\widehat{gl}(V)$, would have a nonzero kernel.]

We now have $V = W \oplus ker(C_V)$. To complete the argument we need to show that $ker(C_V)$ is invariant by $\rho(\hat{g})$. We let $v \neq 0$ in V be in the $ker(C_V)$. We show that $X(v)$ is also in the $ker(C_V)$ for any X in $\rho(\hat{g})$. This is true because of the commutativity of C_V and X.

$$C_V(X(v)) = X(C_V(v)) = X(0) = 0$$

And thus we can conclude that if W is an irreducible subspace of V of codimension one, there exist a complementary one-dimensional subspace W' also invariant by $\rho(\hat{g})$, namely, $W' = ker(C_V)$.

With this we have reached our conclusion that W as a proper invariant subset of whatever dimension of V has a complementary subset $W' = ker(\psi)$ which is invariant by $\hat{g}$.

2.16 Levi Decomposition Theorem

Recall that an arbitrary Lie algebra $\hat{g}$ possesses a radical, which we denote by $\hat{r}$. If $\hat{r} = 0$, then by definition $\hat{g}$ is semisimple. Let us assume now that $\hat{r} \neq 0$. Since this radical is an ideal, we can form the quotient Lie algebra $\hat{g}/\hat{r}$. This gives us the short exact sequence

$$0 \longrightarrow \hat{r} \longrightarrow \hat{g} \longrightarrow \hat{g}/\hat{r} \longrightarrow 0$$

We proved that this quotient Lie algebra is indeed semisimple. [See 2.4.] The question now is: does this short exact sequence split? There is a famous theorem of Levi that gives a positive answer to this question, namely that there exists a Levi factor $\hat{l}$ [a Lie algebra] such that $\hat{l}$ is isomorphic to $\hat{g}/\hat{r}$ [and thus $\hat{l}$ is semisimple] and such that

$$\hat{g} = \hat{l} \oplus \hat{r}$$

We note, however, that the above expression is just a linear space direct sum and not a direct sum of Lie algebras. The bracket of $\hat{l}$ with $\hat{r}$ does not have to be zero. Since $\hat{r}$ is an ideal, we know that $[\hat{l}, \hat{r}] \subset \hat{r}$ but it is not necessarily 0. [As we remarked above in 2.4, we call this a *semi-direct product of Lie algebras.*]

2.16.1 $\hat{\mathbf{g}} = \mathbf{D^1}\hat{\mathbf{g}}$ does not always imply that $\hat{\mathbf{g}}$ is semisimple. Before we present a proof of the Levi decomposition theorem, we first would like to recall that we have shown that for any semisimple Lie algebra $\hat{g}$, $D^1\hat{g} = \hat{g}$ [See 2.10.2.] At this moment we would like to show that this condition is not sufficient, that is, the condition $D^1\hat{g} = \hat{g}$ does not necessarily imply that $\hat{g}$ is semisimple.

Here is an example. Let the Lie algebra $\hat{g}$ be defined as follows. First we choose any semisimple Lie algebra $\hat{h}$, and let ρ be a representation of $\hat{h}$ on a linear space V over a field $\mathbb{F}$ such that no linear subspace of V is left invariant by $\rho(\hat{h})$ except the two improper subspaces 0 and V. (This, of course, says that the representation is irreducible.) We then define a Lie algebra $\hat{g} = \hat{h} \oplus V$, where the bracket in $\hat{h}$ is already defined because it is a Lie algebra. We make V into an abelian Lie algebra by defining its brackets to be 0. In addition we define a twisted bracket product between $\hat{h}$ and V:

$$[(x_1, v_1), (x_2, v_2)] := ([x_1, x_2], \rho(x_1) \cdot (v_2) - \rho(x_2) \cdot (v_1))$$

We now show that this set up does indeed define $\hat{g}$ to be a Lie algebra.

In the following we use the fact that ρ is a linear map from $\hat{h}$ to $End(V)$, i.e., for x_1 and x_2 in $\hat{h}$, we have $\rho(x_1 + x_2) = \rho(x_1) + \rho(x_2)$; and for x in $\hat{h}$ and c in $\mathbb{F}$, we have $\rho(cx) = c\rho(x)$. Likewise since $\rho(x)$ is a linear transformation on V, for v_1 and v_2 in V, we have $\rho(x) \cdot (v_1 + v_2) = \rho(x) \cdot (v_1) + \rho(x) \cdot (v_2)$; and for v in V and for c in $\mathbb{F}$, we have $\rho(x) \cdot (cv) = c(\rho(x) \cdot v)$.

We have distribution on the right for this bracket.

$$\begin{gathered}
[(x_1, v_1) + (x_2, v_2), (x_3, v_3)] = \\
[(x_1 + x_2, v_1 + v_2), (x_3, v_3)] = \\
([x_1 + x_2, x_3], \rho(x_1 + x_2) \cdot (v_3) - \rho(x_3) \cdot (v_1 + v_2)) = \\
([x_1, x_3] + [x_2, x_3], (\rho(x_1) + \rho(x_2)) \cdot (v_3) - \rho(x_3) \cdot (v_1 + v_2)) = \\
([x_1, x_3] + [x_2, x_3], \rho(x_1) \cdot (v_3) + \rho(x_2) \cdot (v_3) - \rho(x_3) \cdot (v_1) - \rho(x_3) \cdot (v_2)) = \\
([x_1, x_3], \rho(x_1) \cdot (v_3) - \rho(x_3) \cdot (v_1)) + ([x_2, x_3], \rho(x_2) \cdot (v_3) - \rho(x_3) \cdot (v_2)) = \\
[(x_1, v_1), (x_3, v_3)] + [(x_2, v_2), (x_3, v_3)]
\end{gathered}$$

Likewise we can show that we have distribution on the left. Also we have the distributivity property for scalars. For let c be in the field $\mathbb{F}$. Then

$$
\begin{gathered}
c[(x_1,v_1),(x_2,v_2)] = c([x_1,x_2],\rho(x_1)\cdot(v_2)-\rho(x_2)\cdot(v_1)) = \\
(c[x_1,x_2],c(\rho(x_1)\cdot(v_2)-\rho(x_2)\cdot(v_1))) = \\
([cx_1,x_2],(c(\rho(x_1)\cdot(v_2))-c(\rho(x_2)\cdot(v_1)))) = \\
([cx_1,x_2],(c\rho(x_1))\cdot(v_2)-\rho(x_2)\cdot(cv_1)) = \\
([cx_1,x_2],(\rho(cx_1))\cdot(v_2)-\rho(x_2)\cdot(cv_1)) = \\
[(cx_1,cv_1),(x_2,v_2)] = [c(x_1,v_1),(x_2,v_2)]
\end{gathered}
$$

and

$$
\begin{gathered}
c[(x_1,v_1),(x_2,v_2)] = c([x_1,x_2],\rho(x_1)\cdot(v_2)-\rho(x_2)\cdot(v_1)) = \\
(c[x_1,x_2],c(\rho(x_1)\cdot(v_2)-\rho(x_2)\cdot(v_1))) = \\
([x_1,cx_2],(c(\rho(x_1)\cdot(v_2))-c(\rho(x_2)\cdot(v_1)))) = \\
([x_1,cx_2],\rho(x_1)\cdot(cv_2)-c\rho(x_2)\cdot(v_1)) = \\
([x_1,cx_2],\rho(x_1)\cdot(cv_2)-(\rho(cx_2))\cdot(v_1)) = \\
[(x_1,v_1),(cx_2,cv_2)] = [(x_1,v_1),c(x_2,v_2)]
\end{gathered}
$$

Now we show that this bracket gives us the structure of a Lie algebra over $\mathbb{F}$. We need to show the anticommutativity property of the bracket:

$$
\begin{gathered}
[(x_2,v_2),(x_1,v_1)] = ([x_2,x_1],\rho(x_2)\cdot(v_1)-\rho(x_1)\cdot(v_2)) = \\
(-[x_1,x_2],-(\rho(x_1)\cdot(v_2)-\rho(x_2)\cdot(v_1))) = \\
-([x_1,x_2],\rho(x_1)\cdot(v_2)-\rho(x_2)\cdot(v_1)) = -[(x_1,v_1),(x_2,v_2)]
\end{gathered}
$$

Finally, we verify the Jacobian identity:

$$
\begin{gathered}
[[(x_1,v_1),(x_2,v_2)],(x_3,v_3)] = [([x_1,x_2],\rho(x_1)\cdot(v_2)-\rho(x_2)\cdot(v_1)),(x_3,v_3)] = \\
([[x_1,x_2],x_3],\rho([x_1,x_2])\cdot(v_3)-\rho(x_3)\cdot(\rho(x_1)\cdot(v_2)-\rho(x_2)\cdot(v_1))) = \\
([[x_1,x_2],x_3],[\rho(x_1),\rho(x_2)]\cdot(v_3)-\rho(x_3)\cdot(\rho(x_1)\cdot(v_2))+\rho(x_3)\cdot(\rho(x_2)\cdot(v_1))) = \\
([[x_1,x_2],x_3], \\
(\rho(x_1)\rho(x_2)-\rho(x_2)\rho(x_1))\cdot(v_3)-(\rho(x_3)\rho(x_1))\cdot(v_2)+(\rho(x_3)\rho(x_2))\cdot(v_1)) = \\
([[x_1,x_2],x_3], \\
(\rho(x_1)\rho(x_2))\cdot(v_3)-(\rho(x_2)\rho(x_1))\cdot(v_3)-(\rho(x_3)\rho(x_1))\cdot(v_2)+(\rho(x_3)\rho(x_2))\cdot(v_1))
\end{gathered}
$$

$$
\begin{gathered}
[[(x_3,v_3),(x_1,v_1)],(x_2,v_2)] = [([x_3,x_1],\rho(x_3)\cdot(v_1)-\rho(x_1)\cdot(v_3)),(x_2,v_2)] = \\
([[x_3,x_1],x_2],\rho([x_3,x_1])\cdot(v_2)-\rho(x_2)\cdot(\rho(x_3)\cdot(v_1)-\rho(x_1)\cdot(v_3))) = \\
([[x_3,x_1],x_2],[\rho(x_3),\rho(x_1)]\cdot(v_2)-\rho(x_2)\cdot(\rho(x_3)\cdot(v_1))+\rho(x_2)\cdot(\rho(x_1)\cdot(v_3))) = \\
([[x_3,x_1],x_2], \\
(\rho(x_3)\rho(x_1)-\rho(x_1)\rho(x_3))\cdot(v_2)-(\rho(x_2)\rho(x_3))\cdot(v_1)+(\rho(x_2)\rho(x_1))\cdot(v_3)) = \\
([[x_3,x_1],x_2], \\
(\rho(x_3)\rho(x_1))\cdot(v_2)-(\rho(x_1)\rho(x_3))\cdot(v_2)-(\rho(x_2)\rho(x_3))\cdot(v_1)+(\rho(x_2)\rho(x_1))\cdot(v_3))
\end{gathered}
$$

$$
\begin{gathered}
[[(x_2,v_2),(x_3,v_3)],(x_1,v_1)] = [([x_2,x_3],\rho(x_2)\cdot(v_3)-\rho(x_3)\cdot(v_2)),(x_1,v_1)] = \\
([[x_2,x_3],x_1],\rho([x_2,x_3])\cdot(v_1)-\rho(x_1)\cdot(\rho(x_2)\cdot(v_3)-\rho(x_3)\cdot(v_2))) = \\
([[x_2,x_3],x_1],[\rho(x_2),\rho(x_3)]\cdot(v_1)-\rho(x_1)\cdot(\rho(x_2)\cdot(v_3))+\rho(x_1)\cdot(\rho(x_3)\cdot(v_2))) =
\end{gathered}
$$

$([[x_2, x_3], x_1],$
$(\rho(x_2)\rho(x_3) - \rho(x_3)\rho(x_2)) \cdot (v_1) - (\rho(x_1)\rho(x_2)) \cdot (v_3) + (\rho(x_1)\rho(x_3)) \cdot (v_2)) =$
$([[x_2, x_3], x_1],$
$(\rho(x_2)\rho(x_3)) \cdot (v_1) - (\rho(x_3)\rho(x_2)) \cdot (v_1) - (\rho(x_1)\rho(x_2)) \cdot (v_3) + (\rho(x_1)\rho(x_3)) \cdot (v_2))$

Obviously the $\hat{h}$-component of these three expressions adds to zero since this is the Jacobi identity in $\hat{h}$. The V-component has 12 terms, 6 positive and 6 negative. On inspection we see that the 6 positive terms are balanced by the 6 negative terms, giving zero for the V-component also. Thus we have verified the Jacobi identity in $\hat{g} = \hat{h} \oplus V$, and we can conclude that $\hat{g} = \hat{h} \oplus V$ is a Lie algebra over $\mathbb{F}$.

We now show that V is a solvable ideal in $\hat{g}$. For an arbitrary (x, v) in $\hat{g}$ and an arbitrary $(0, u)$ in V, we have $[(x, v), (0, u)] = ([x, 0], \rho(x) \cdot u - \rho(0) \cdot v) = (0, \rho(x) \cdot u)$ which is certainly in V. Thus we know that V is an ideal in $\hat{g}$. Now for any two elements $(0, u_1)$ and $(0, u_2)$ in V, we have $[(0, u_1), (0, u_2)] = ([0, 0], \rho(0) \cdot u_2 - \rho(0) \cdot u_1) = (0, 0)$. Thus we see that V is actually an abelian ideal and thus is solvable.

However we can assert more. We can say that V is actually the radical of $\hat{g}$. Suppose W is a solvable ideal in $\hat{g}$. Let (x_1, v_1) and (x_2, v_2) be two elements in W. Then $[(x_1, v_1), (x_2, v_2)] = ([x_1, x_2], \rho(x_1) \cdot v_2 - \rho(x_2) \cdot v_1)$. Now the $\hat{h}$-component of this product, by iterations of the bracket product in $\hat{h}$, can never be pulled down to 0 since we know that $\hat{h}$ is semisimple, and thus $[\hat{h}, \hat{h}] = \hat{h}$. Thus the only way of having W solvable is that it have no $\hat{h}$-component. We conclude that W is contained in V, thus making V the radical of $\hat{g}$.

Finally, we calculate $D^1\hat{g} = [\hat{g}, \hat{g}] = [\hat{h} \oplus V, \hat{h} \oplus V]$. Now the $\hat{h}$-component of the bracket product gives $[\hat{h}, \hat{h}] = \hat{h}$, since $\hat{h}$ is semisimple. For the V-component of the bracket product, we take arbitrary (x_1, v_1) and (x_2, v_2) in $\hat{g}$ and calculate the V-component of the bracket product, $\rho(x_1) \cdot v_2 - \rho(x_2) \cdot v_1$. We wish to show that for any v in V, there is a pair $((x_1, v_1), (x_2, v_2))$ such that $v = \rho(x_1) \cdot v_2 - \rho(x_2) \cdot v_1$. Suppose there is no such pair. First let us fix one element of the pair, say, (x_1, v_1); and let us suppose we cannot reach v_0 in V by $v_0 = \rho(x_1) \cdot v_2 - \rho(x_2) \cdot v_1$ for any x_2 in $\hat{h}$ and any v_2 in V. We can write this as $v_0 \notin \rho(x_1) \cdot V - \rho(\hat{h}) \cdot v_1$. Suppose now that we choose the subset $\rho(\hat{h}) \cdot 0$ in $\rho(\hat{h}) \cdot V$. This gives $v_0 \notin \rho(\hat{h}) \cdot V - \rho(\hat{h}) \cdot 0$, or equivalently $v_0 \notin \rho(\hat{h}) \cdot V$. This says that for all x in $\hat{h}$ and all v in V, $\rho(x) \cdot (v) \neq v_0$. But this statement is equivalent to asserting that the map $\rho(\hat{h})$ operating on V is not surjective. Then $\rho(\hat{h}) \cdot V$ would give a proper subspace W of V. But $\rho(\hat{h}) \cdot W$ would again be in W. But this means that W is a proper invariant subspace of V by $\rho(\hat{h})$. However we know that ρ is an irreducible representation on V, and thus has no invariant subspaces except the two improper ones. We can therefore conclude that $\rho(\hat{h})$ operating on V is surjective, and thus we have $D^1\hat{g} = D^1(\hat{h} \oplus V) = \hat{h} \oplus V = \hat{g}$.

We have shown that $D^1\hat{g} = \hat{g}$ can be true even if $\hat{g}$ is not semisimple. For if $\hat{g} = \hat{h} \oplus V$, then we know that $D^1\hat{g} = \hat{g}$, but since $\hat{g}$ has a non-trivial radical V, it certainly is not semisimple. What is interesting about this example is that it illustrates the Levi Decomposition Theorem. $\hat{g}$ is decomposed as a direct sum of a semisimple part $\hat{h}$ and the radical V of $\hat{g}$. But in this example the radical is abelian. We will return to this example later. But at the present moment we want to prove another fact about Lie algebras.

2.16.2 rad($\hat{\mathbf{a}}$) = $\hat{\mathbf{a}} \cap \hat{\mathbf{r}}$. We would like to show that if $\hat{g}$ is a Lie algebra and $\hat{r}$ is its radical, then for any other ideal $\hat{a}$ of $\hat{g}$, the radical of $\hat{a}$, $rad(\hat{a})$, is determined by the following relation

$$rad(\hat{a}) = \hat{a} \cap \hat{r}$$

The fact that $\hat{a} \cap \hat{r} \subset rad(\hat{a})$ is straightforward. Since the intersection of two ideals in $\hat{g}$ is an ideal in $\hat{g}$, $\hat{a} \cap \hat{r} \subset \hat{a}$ is an ideal, and thus is also an ideal in $\hat{a}$. And since $\hat{a} \cap \hat{r} \subset \hat{r}$ is an ideal in $\hat{r}$ and $\hat{r}$ is solvable, then $\hat{a} \cap \hat{r}$ is a solvable ideal in $\hat{a}$. Thus $\hat{a} \cap \hat{r} \subset rad(\hat{a})$.

Now to prove $rad(\hat{a}) \subset \hat{a} \cap \hat{r}$, we obviously have $rad(\hat{a}) \subset \hat{a}$. Thus we need to prove that $rad(\hat{a})$ is also in $\hat{r}$, the radical of $\hat{g}$. To do this we go to quotient Lie algebras. Since $\hat{a} \cap \hat{r}$ is an ideal in $\hat{a}$, and $\hat{a}$ contains $rad(\hat{a})$, we have the quotient Lie algebra relation:

$$rad(\hat{a})/\hat{a} \cap \hat{r} \subset \hat{a}/\hat{a} \cap \hat{r}$$

Now if we can show that $\hat{a}/\hat{a} \cap \hat{r}$ is the zero coset, then we can conclude that $rad(\hat{a}) \subset \hat{a} \cap \hat{r}$. We use an isomorphism theorem of linear algebra to give the following isomorphism of Lie algebras:

$$\hat{a}/\hat{a} \cap \hat{r} \cong (\hat{a} + \hat{r})/\hat{r}$$

Then we show that $(\hat{a} + \hat{r})/\hat{r}$ is an ideal in $\hat{g}/\hat{r}$:

$$[(\hat{a} + \hat{r}) + \hat{r}, \hat{g} + \hat{r}] \subset [\hat{a} + \hat{r}, \hat{g}] \subset [\hat{a}, \hat{g}] + [\hat{r}, \hat{g}] \subset \hat{a} + \hat{r} = (\hat{a} + \hat{r}) + \hat{r}$$

We conclude that $(\hat{a} + \hat{r})/\hat{r}$ is an ideal in $\hat{g}/\hat{r}$. But we know that $\hat{g}/\hat{r}$ is semisimple. Now the only ideals that a semisimple Lie algebra has are semisimple ideals. Thus we can conclude that $(\hat{a} + \hat{r})/\hat{r}$ is semisimple. By the above isomorphism we can assert that $\hat{a}/\hat{a} \cap \hat{r}$ is also semisimple. Now we can show that $rad(\hat{a})/\hat{a} \cap \hat{r}$ is an ideal in $\hat{a}/\hat{a} \cap \hat{r}$ and is also solvable. It is an ideal because

$$[rad(\hat{a})+(\hat{a}\cap\hat{r}),\hat{a}+(\hat{a}\cap\hat{r})] \subset [rad(\hat{a}),\hat{a}]+[rad(\hat{a}),\hat{a}\cap\hat{r}]+[\hat{a}\cap\hat{r},\hat{a}]+[\hat{a}\cap\hat{r},\hat{a}\cap\hat{r}] \subset rad(\hat{a})+\hat{a}\cap\hat{r}+\hat{a}\cap\hat{r}+\hat{a}\cap\hat{r} = rad(\hat{a})+\hat{a}\cap\hat{r}$$

since $rad(\hat{a})$ and $\hat{a}\cap\hat{r}$ are ideals. Now $rad(\hat{a})/\hat{a}\cap\hat{r}$ is solvable since $rad(\hat{a})$ is solvable, and homomorphic images of solvable Lie algebras are also solvable. Thus we have a solvable ideal $rad(\hat{a})/\hat{a}\cap\hat{r}$ contained in a semisimple ideal $\hat{a}/\hat{a}\cap\hat{r}$. But the only solvable ideal that a semisimple Lie algebra can have is the trivial ideal 0. If we unwind the quotient, this means that $rad(\hat{a})$ is contained in $\hat{a}\cap\hat{r}$. We thus have our conclusion that $rad(\hat{a}) = \hat{a}\cap\hat{r}$.

2.16.3 Proof of the Levi Decomposition Theorem. We proceed now with the proof of the Levi Decomposition Theorem. It is obvious that we should use induction on the dimension of $\hat{g}$. Let us start, then. with the dimension of $\hat{g} = 1$. In this case $\hat{g}$ is abelian and thus the radical is equal to $\hat{g}$, and there is nothing to prove. Now we know that there are no two-dimensional simple Lie algebras (why is that so?) and therefore, once again, there is nothing to prove. If we have the dimension of $\hat{g} = 3$, then $sl_2(\mathbb{C})$, the set of $2x2$ trace zero matrices, forms a 3-dimensional simple Lie algebra, and once again there is nothing to prove. Thus the first dimension that can illustrate the theorem is where the dimension equal to 4. Here is such a Lie algebra:

$$\hat{g} = sl_2(\mathbb{C}) \oplus \hat{a}$$

where $\hat{a}$ is a one-dimensional abelian Lie algebra, the radical of $\hat{g}$.

Thus we assume that the Levi Decomposition Theorem is true for all Lie algebras whose dimension is less than the dimension of $\hat{g}$. First we want to assume that there is an ideal $\hat{a}$ of $\hat{g}$ that is properly contained in the radical $\hat{r}$, that is, $0 \subset \hat{a} \subset \hat{r}$ and $0 \neq \hat{a} \neq \hat{r}$. This means that $\hat{a}$ is also solvable. Since $\hat{a}$ is also properly contained in $\hat{g}$, the quotient Lie algebra $\hat{g}/\hat{a}$ has dimension less than $\hat{g}$. Thus assuming the Levi Decomposition Theorem for $\hat{g}/\hat{a}$, we have

$$\hat{g}/\hat{a} = \hat{r}/\hat{a} \oplus \hat{k}/\hat{a}$$

where the quotient algebra $\hat{k}/\hat{a}$ is semisimple. We would now like to show that $rad(\hat{g}/\hat{a})$ is actually equal to $\hat{r}/\hat{a}$, which we have assumed not to be equal to 0 nor to $\hat{r}$. Since $\hat{r}$ is solvable, then the homomorphic image of $\hat{r}$, $\hat{r}/\hat{a}$, is also solvable and thus is contained in the radical $rad(\hat{g}/\hat{a})$. Now let $\hat{b}/\hat{a}$ be any solvable subalgebra of $\hat{g}/\hat{a}$. Since $\hat{a}$ is also solvable, by the homomorphism theorem of solvable Lie algebras, we know that $\hat{b}$ is also solvable in $\hat{g}$. Thus we have $\hat{b}$ contained in the radical $\hat{r}$ of $\hat{g}$, giving $\hat{b}/\hat{a}$ contained in $\hat{r}/\hat{a}$. But the $rad(\hat{g}/\hat{a})$ is also a solvable subalgebra of $\hat{g}/\hat{a}$, and thus must be contained in $\hat{r}/\hat{a}$. We conclude that $\hat{r}/\hat{a} = rad(\hat{g}/\hat{a})$, giving

$$\hat{g}/\hat{a} = \hat{r}/\hat{a} \oplus \hat{k}/\hat{a}$$

Since $\hat{r}/\hat{a}$ is the radical of $\hat{g}/\hat{a}$, we know we have semisimple Lie algebras

$$(\hat{g}/\hat{a})/(\hat{r}/\hat{a}) \cong \hat{g}/\hat{r} \cong \hat{k}/\hat{a}$$

Recall that we are still seeking a semisimple Lie subalgebra $\hat{l}$ of $\hat{g}$ such that

$$\hat{g} = \hat{r} \oplus \hat{l}$$

that is, we need $\hat{l}$ to satisfy

$$\hat{k}/\hat{a} \cong \hat{l}$$

Once again we can use induction on the dimension of $\hat{g}$. We calculate the dimension of $\hat{k}$.

$$\begin{aligned} \dim(\hat{g}) \text{ - } \dim(\hat{a}) &= \dim(\hat{r}) \text{ - } \dim(\hat{a}) + \dim(\hat{k}) \text{ - } \dim(\hat{a}) \\ \dim(\hat{g}) \text{ - } \dim(\hat{r}) + \dim(\hat{a}) &= \dim(\hat{k}) \end{aligned}$$

But $\dim(\hat{r}) > \dim(\hat{a})$, which means that $\dim(\hat{g}) > \dim(\hat{k})$. Thus we can use induction on the Lie algebra $\hat{k}$. Applying the Levi Decomposition Theorem to $\hat{k}$, we have

$$\hat{k} = rad(\hat{k}) \oplus \hat{k}_2$$

where $\hat{k}_2$ is a semisimple Lie subalgebra of $\hat{g}$ isomorphic to $\hat{k}/rad(\hat{k})$. Now we show that the $rad(\hat{k}) = \hat{a}$. We know that $\hat{k}/rad(\hat{k})$ is a semisimple Lie algebra. Now $\hat{a}$ is solvable in $\hat{g}$, thus is solvable in $\hat{k}$, and thus we know that it is in $rad(\hat{k})$. Now let $\hat{d}$ be any solvable subalgebra of $\hat{k}$. Then $\hat{d}/\hat{a}$ is solvable since it is the homomorphic image of a solvable algebra $\hat{d}$. Since $\hat{d}$ is contained in $\hat{k}$, we have $\hat{d}/\hat{a}$ is contained in $\hat{k}/\hat{a}$. But $\hat{k}/\hat{a}$ is semisimple and thus has no nontrivial solvable algebras. Thus $\hat{d}/\hat{a} = 0$, which says that $\hat{d}$ is contained in $\hat{a}$. But since $rad(\hat{k})$ is a solvable subalgebra in $\hat{k}$, we know that $rad(\hat{k})$ is contained in $\hat{a}$. We conclude that $rad(\hat{k}) = \hat{a}$. This gives

$$\hat{k} = \hat{a} \oplus \hat{k}_2$$

with

$$\hat{a} \cap \hat{k}_2 = 0$$

We also know

$$\hat{k}/rad(\hat{k}) = \hat{k}/\hat{a} \cong \hat{g}/\hat{r} \cong \hat{k}_2$$

Thus we see that $\hat{k}_2$ is semisimple and isomorphic to $\hat{g}/\hat{r}$. Finally we have from $\hat{g}/\hat{a} = \hat{r}/\hat{a} \oplus \hat{k}/\hat{a}$ that

$$\hat{g} + \hat{a} = \hat{r} + \hat{a} + \hat{k} + \hat{a}$$
$$\hat{g} = \hat{r} + \hat{k} = \hat{r} + \hat{a} + \hat{k}_2 \quad \text{with} \quad \hat{a} \cap \hat{k}_2 = 0$$
$$\hat{g} = \hat{r} + \hat{k}_2 \quad \text{with} \quad \hat{a} \cap \hat{k}_2 = 0$$

But knowing also that $\hat{g}/\hat{a} = \hat{r}/\hat{a} \oplus \hat{k}/\hat{a}$, we claim that this says that $\hat{g} = \hat{r} \oplus \hat{k}_2$. For $\hat{g}/\hat{a} = \hat{r}/\hat{a} \oplus \hat{k}/\hat{a}$ says that $\hat{r} \cap \hat{k} \subset \hat{a}$. We know that $\hat{k}_2 \subset \hat{k}$. Since $\hat{g} = \hat{r} + \hat{k}_2$, we choose a d in $\hat{r} \cap \hat{k}_2$. But since $\hat{a} \cap \hat{k}_2 = 0$, we know that d is not in $\hat{a}$. But d in $\hat{r}$ and d in $\hat{k}_2$ and d not in $\hat{a}$ means that d is in $\hat{k}$. Thus d is in $\hat{r} \cap \hat{k}$, which is contained in $\hat{a}$. Therefore we reach a contradiction, and thus d is not in $\hat{k}_2$, and we can conclude that $\hat{g} = \hat{r} \oplus \hat{k}_2$. Thus we see that $\hat{k}_2$ is the $\hat{l}$ that we are seeking and this establishes the Levi Decomposition Theorem in the case where $\hat{g}$ has a solvable ideal $\hat{a}$ which is also contained properly in the radical $\hat{r}$ of $\hat{g}$.

But what is the situation if the above condition for $\hat{a}$ is not true? We make the following observations. Under certain conditions we know that the above ideal $\hat{a}$ always exists. Since $\hat{r}$ is a solvable Lie algebra, we know that for some k, $D^k\hat{r}$, which is an ideal in $\hat{r}$, is not equal to 0, but $D^{k+1}\hat{r} = 0$. Also we know that $\hat{r} \neq D^1\hat{r}$, for otherwise the process of taking successive brackets would never arrive at a $D^k\hat{r} = 0$, and having a k such that a $D^k\hat{r} = 0$ for some $k \geq 0$ is the definition of $\hat{r}$ being solvable. Thus we can conclude that in this case there is solvable ideal of $\hat{r}$ which is properly contained in $\hat{r}$. We would also like to affirm that it is also an ideal of $\hat{g}$, and thus also a solvable ideal of $\hat{g}$. Using the Jacobi identity, we have the following.

$$[\hat{g}, D^1\hat{r}] = [\hat{g}, [\hat{r}, \hat{r}]] \subset [[\hat{g}, \hat{r}], \hat{r}] + [\hat{r}, [\hat{g}, \hat{r}]] \subset [\hat{r}, \hat{r}] + [\hat{r}, \hat{r}] = [\hat{r}, \hat{r}] = D^1\hat{r}$$

Thus we see that $D^1\hat{r}$ is an ideal in $\hat{g}$. Continuing

$$[\hat{g}, D^2\hat{r}] = [\hat{g}, [D^1\hat{r}, D^1\hat{r}]] \subset [[\hat{g}, D^1\hat{r}], D^1\hat{r}] + [D^1\hat{r}, [\hat{g}, D^1\hat{r}]] \subset$$
$$[D^1\hat{r}, D^1\hat{r}] + [D^1\hat{r}, D^1\hat{r}] = [D^1\hat{r}, D^1\hat{r}] = D^2\hat{r}$$

Continuing in this way. we can conclude that $D^k\hat{r}$ is also an ideal in $\hat{g}$ for all $k > 0$. And thus we have a solvable ideal of $\hat{g}$ properly contained in the radical of $\hat{r}$, and we know that the Levi Decomposition theorem applies to this case.

But what happens when $D^1\hat{r} = 0$, i.e., when $\hat{r}$ is abelian and thus when there is no nontrivial ideal in $\hat{r}$ in the derived series of $\hat{r}$? Since $\hat{r}$ is an ideal in $\hat{g}$, we know that $[\hat{g}, \hat{r}] \subset \hat{r}$. Let us read this as $[\hat{g}, \hat{r}] = ad(\hat{g})(\hat{r}) \subset \hat{r}$; that is, we have the adjoint representation of $\hat{g}$ on $\hat{r}$. Now let us suppose that this representation is reducible. This means that there is a proper subspace $\hat{a}$ of $\hat{r}$ $[0 \neq \hat{a} \neq \hat{r}]$ which is invariant by $ad(\hat{g})$. This says that $ad(\hat{g})(\hat{a}) \subset \hat{a}$, or $[\hat{g}, \hat{a}] \subset \hat{a}$. [Thus, by the way, we see that the subspace $\hat{a}$ is actually an ideal of $\hat{g}$.] Unwinding all these facts, we can assert that, even when $\hat{r}$ is abelian and when the adjoint representation of $\hat{g}$ on $\hat{r}$ is reducible, we have an ideal $\hat{a}$ of $\hat{g}$ properly contained in $\hat{r}$, and again this is a hypothesis which enables us to prove the Levi Decomposition Theorem. [We note, as a side remark, that nowhere in the above proof of the theorem did we exclude the fact that $\hat{a}$ is abelian, which it is since it is contained in $\hat{r}$.]

In particular, then, we consider the case where the center $\hat{z}$ of $\hat{g}$ and $0 \neq \hat{z} \neq \hat{r}$. In this case the adjoint representation of $\hat{g}$ on $\hat{r}$ is reducible since the center, which is an ideal, gives $ad(\hat{g})(\hat{z}) = [\hat{g}, \hat{z}] = 0 \subset \hat{z}$. This says, of course, that the center is an ideal, and again the proof given above is valid. In particular we get $\hat{g} = \hat{r} \oplus \hat{k}_2$, where $\hat{k}_2$ is semisimple and $\hat{r}$ is the maximal abelian ideal, which is the radical, but also $\hat{k} = \hat{z} \oplus \hat{k}_2$, which shows how the center, which is a solvable abelian ideal, sits in $\hat{r}$ and thus in $\hat{g}$.

But suppose now that the center itself is the abelian radical, i.e., suppose that $r = z$. Then $ad(\hat{g})\hat{z} = [\hat{g}, \hat{z}] = 0 \subset \hat{z}$, and this says that $\hat{z}$ is irreducible. [Recall that a representation is irreducible if the only invariant subspaces are 0 and the representation space V itself. (See 2.8.1)] Now $\widehat{gl}(\hat{g})$ is our representation space and our representation is ad. We are now assuming that the center $\hat{z}$ of $\hat{g}$ is also the radical $\hat{r}$ of $\hat{g}$. We are asking, in this case, if this radical is reducible? But $ad(\hat{g})\hat{z} = [\hat{g}, \hat{z}] = 0$. Thus there is no proper subspace of $\hat{z}$ left invariant by $ad(\hat{g})$, and we can conclude that $\hat{z}$ is irreducible, and thus none of the above methods can be applied in this case. We therefore need another method of attack.

We first remark that if the center is the radical, then $D^1\hat{g}$ is semisimple. We know that $[\hat{g}, \hat{r}] = D^1\hat{g} \cap \hat{r}$ (see 2.4 but note that the proof there is not complete – the reader is challenged to complete it). Since $\hat{z}$ is the center, we have $[\hat{g}, \hat{z}] = 0 = D^1\hat{g} \cap \hat{z}$. On the other hand we know that $rad(D^1\hat{g}) = D^1\hat{g} \cap \hat{r}$ (see 2.16.2), and thus we have $rad(D^1\hat{g}) = D^1\hat{g} \cap \hat{z} = 0$, and we can conclude that $D^1\hat{g}$ is semisimple.

We also know that if the center is the radical, we have $\hat{g}/\hat{z}$ is isomorphic to a semisimple Lie algebra. We need then to show that this Lie algebra is exactly $D^1\hat{g}$, which we know is semisimple.

We return to the Killing form B of $\hat{g}$ and to the map $\mathcal{B}$ of $\hat{g}$ to its dual $\hat{g}^*$ that it determines.

$$\hat{g} \xrightarrow{\mathcal{B}} \hat{g}^*$$
$$x \longrightarrow \mathcal{B}(x) : \hat{g} \longrightarrow \mathbb{F}$$

$$y \longrightarrow \mathcal{B}(x)(y) := B(x,y)$$

If we restrict $\mathcal{B}$ to $D^1\hat{g}$, we know $\mathcal{B}$ is an isomorphism on this part of $\hat{g}$ since $D^1\hat{g}$ is semisimple. (Cf. the ends of 2.14.2 and 2.14.3.) Thus, the only element in $D^1\hat{g}$ that goes to 0 in $\hat{g}^*$ is the 0 element in $D^1\hat{g} \subset \hat{g}$. Let $\hat{k}$ be the kernel of the $\mathcal{B}$ map. Then $\hat{k} \cap D^1\hat{g} = 0$ and we have $\hat{k} \oplus D^1\hat{g} \subset \hat{g}$. We see that $\hat{z} \subset \hat{k}$. We restrict $\mathcal{B}$ to $\hat{z}$.

$$\begin{aligned}
&\mathcal{B}(\hat{z}) : \hat{g} \longrightarrow \mathbb{F} \\
&\mathcal{B}(t) : \hat{g} \longrightarrow \mathbb{F} \\
&x \longmapsto \mathcal{B}(t)(x) = B(t,x) = tr(ad(t) \circ ad(x)),
\end{aligned}$$

where t is in $\hat{z}$. Since t is in the center of $\hat{g}$, we know that $[t, \hat{g}] = 0 = ad(t)(\hat{g})$. Thus $ad(t)$ is the zero map and therefore $tr(ad(t) \circ ad(x)) = 0$. Thus for t in $\hat{z}$, $\mathcal{B}(t)$ is the zero map in $\hat{g}^*$, and we can conclude that $\hat{z}$ is in $\hat{k}$. Finally we want to show that $\hat{k}$ is an ideal in $\hat{g}$, and thus a Lie subalgebra of $\hat{g}$. Thus we must show that $[\hat{k}, \hat{g}] \subset \hat{k}$. To do this we use the associativity of the Killing form B. Let x and y be in $\hat{g}$ and t and t_1 be in $\hat{k}$. Then we have

$$0 = \mathcal{B}(t)([t_1, x]) = B(t, [t_1, x]) = -B(t, [x, t_1]) = -B([t, x], t_1)$$

and

$$0 = \mathcal{B}(t)([y, x]) = B(t, [y, x]) = -B(t, [x, y]) = -B([t, x], y)$$

and thus $\mathcal{B}$ restricted to $[\hat{k}, \hat{g}]$ acting on $\hat{g}$ is the zero map, and we can conclude that $[\hat{k}, \hat{g}]$ is in $\hat{k}$, which says that $\hat{k}$ is an ideal of $\hat{g}$. We now let $\hat{s}$ be any subalgebra of $\hat{g}$ such that $\hat{s} \oplus \hat{k} \oplus D^1\hat{g} = \hat{g}$. Now if we show that $\hat{s}$ must be 0, this will give us what we are seeking, the Levi decomposition theorem. Since $D^1\hat{g}$ is an ideal, we know

$$\hat{g}/D^1\hat{g} \cong \hat{s} \oplus \hat{k}$$

But we also know that $\hat{g}/D^1\hat{g}$ abelianizes $\hat{g}$, and thus $\hat{s} \oplus \hat{k}$ is an abelian Lie algebra. Therefore

$$\begin{aligned}
0 = [\hat{s} \oplus \hat{k}, \hat{s} \oplus \hat{k}] = [\hat{s}, \hat{s}] \oplus [\hat{s}, \hat{k}] \oplus [\hat{k}, \hat{k}] = \\
[\hat{s}, \hat{s}] \oplus [\hat{k}, \hat{k}]
\end{aligned}$$

which says that $[\hat{s}, \hat{s}] = 0$ and $[\hat{k}, \hat{k}] = 0$. Thus $\hat{k}$ is an abelian ideal and must be contained in the radical $\hat{z}$, the center of $\hat{g}$. But we also know that $\hat{z} \subset \hat{k}$, and thus we can conclude that $\hat{k} = \hat{z}$. This also says that $\hat{s}$ must be 0 and we have $\hat{z} \oplus D^1\hat{g} = \hat{g}$, which, of course, is a Levi decomposition of $\hat{g}$ which we have been seeking.

Thus we now are reduced to proving the Levi Decomposition Theorem in the case when $\hat{r} \neq 0$, is abelian and not equal to the center, and $ad(\hat{g})$ acts irreducibly on $\hat{r}$, i.e., $ad(\hat{g})$ leaves no subspace [ideal] of $\hat{r}$ invariant except 0 and $\hat{r}$ itself. [We remark here that this implies that the center $\hat{z}$ of $\hat{g}$ is 0. Here is the proof. Our hypotheses are that the radical $\hat{r}$ is abelian — $ad(\hat{r})(\hat{r}) = [\hat{r}, \hat{r}] = 0$; and that $ad(\hat{g})$ acting on $\hat{r}$ is irreducible. This means that if $\hat{a}$ is contained in $\hat{r}$ and $ad(\hat{g})(\hat{a})$ is contained in $\hat{a}$, then $\hat{a} = \hat{r}$ or $\hat{a} = 0$. Now let the center be $\hat{z}$. Since $\hat{z}$ is solvable, we know that it is contained in the radical $\hat{r}$. Now $ad(\hat{z})(\hat{g}) = 0 \subset \hat{z}$. Thus $\hat{z} = \hat{r}$ or $\hat{z} = 0$. But $ad(\hat{g})(\hat{r}) = \hat{r} \neq 0$ and $\hat{r} \neq \hat{z}$. Thus $\hat{z} = 0$.]

We also remark that when we proved above that $\hat{g}$ could be $D^1\hat{g}$ even if $\hat{g}$ is not semisimple (see 2.16.1) we exhibited a Lie algebra $\hat{g} = \hat{h} \oplus V$ with an abelian radical V and a semisimple subalgebra $\hat{h}$. And in this case we had a representation of $\hat{h}$ on V which was irreducible. We also had

$$[\hat{g}, V] = [\hat{h} + V, V] = [\hat{h}, V] + [V, V] = [\hat{h}, V] + 0 = [\hat{h}, V] = V$$

But this also says that $ad(\hat{g})(V) = V$, i.e., that $ad(\hat{g})$ acts irreducibly on the radical V, and this is exactly the situation in which we now find ourselves. In this example we started with a semisimple Lie algebra $\hat{h}$ which acted irreducibly on an abelian Lie algebra V and produced a Lie algebra $\hat{g}$ in which the Levi Decomposition was true. However we are now given the Lie algebra $\hat{g}$ with an abelian radical on which it acts irreducibly [by the adjoint representation] and we need to find a semisimple Lie subalgebra $\hat{k}$ which is complementary to this abelian radical.

But first let us again return to our example above. We assumed that we had a linear space $\hat{g}$ which contained a semisimple Lie algebra $\hat{h}$ and a complementary linear subspace V, i.e., we assumed that $\hat{g} = \hat{h} \oplus V$. Also we assumed that we had an irreducible representation ρ of $\hat{h}$ in V. Then we defined a bracket product in $\hat{g}$ as

$$[(x_1, v_1), (x_2, v_2)] := ([x_1, x_2], \rho(x_1) \cdot v_2 - \rho(x_2) \cdot v_1)$$

making $\hat{g}$ into a Lie algebra, with $\hat{h}$ being a Lie subalgebra of $\hat{g}$. Finally we also showed that V was abelian and the radical of $\hat{g}$. Under these conditions we proved that $D^1\hat{g} = [\hat{g}, \hat{g}] = \hat{g}$. But we also remarked that this was an example of the Levi Decomposition Theorem, where indeed the radical was abelian and $[\hat{g}, V] = V$.

But we wish to say more. We are in the context where we have a Lie algebra $\hat{g}$ and its radical $\hat{r}$, which is abelian. We take any linear subspace $\hat{k}$ of $\hat{g}$ complementary to $\hat{r}$ in $\hat{g}$, i.e., $\hat{g} = \hat{r} \oplus \hat{k}$. Now for this choice of such an arbitrary linear subspace, there is no reason why we can assert that it is a Lie subalgebra of $\hat{g}$. But the Levi Decomposition Theorem allows us to say that there are certain linear subspaces with this property, and indeed these subspaces are semisimple Lie subalgebras. Now how does our example above fit into this scheme? We know one more fact about the algebra $\hat{g}$. We know that the adjoint representation of $\hat{g}$ acting on the linear space $\hat{r}$, the radical, is irreducible and not equal to 0. This means the $ad(\hat{g}) \cdot \hat{r} = [\hat{g}, \hat{r}] = \hat{r}$. Now let us assume the Levi Decomposition Theorem in this case: $\hat{g} = \hat{l} \oplus \hat{r}$, where $\hat{l}$ is a semisimple Lie subalgebra of $\hat{g}$. Then

$$\hat{r} = [\hat{g}, \hat{r}] = [\hat{l} \oplus \hat{r}, \hat{r}] = [\hat{l}, \hat{r}] + [\hat{r}, \hat{r}] = [\hat{l}, \hat{r}] + 0$$

Here we see that $[\hat{l}, \hat{r}] = \hat{r}$. Thus we see that $[\hat{l}, \hat{r}] = ad(\hat{l}) \cdot \hat{r}$, and this says that $ad(\hat{l})$ acts irreducibly on $\hat{r}$, the radical of $\hat{g}$. We now calculate the bracket in $\hat{g}$.

$$\begin{gathered}[\hat{g}, \hat{g}] = [\hat{l} \oplus \hat{r}, \hat{l} \oplus \hat{r}] = \\ [\hat{l}, \hat{l}] + [\hat{l}, \hat{r}] + [\hat{r}, \hat{l}] + [\hat{r}, \hat{r}] = \\ [\hat{l}, \hat{l}] + [\hat{l}, \hat{r}] + [\hat{r}, \hat{l}] + 0\end{gathered}$$

Choosing u_1 and u_2 in $\hat{l}$; and s_1 and s_2 in $\hat{r}$, we have

$$\begin{gathered}[(u_1 + s_1), (u_2 + s_2)] = [u_1, u_2] + [u_1, s_2] + [s_1, u_2] + [s_1, s_2] = \\ [u_1, u_2] + [u_1, s_2] - [u_2, s_1] + [s_1, s_2] = \\ [u_1, u_2] + ad(u_1)(s_2) - ad(u_2)(s_1) + 0\end{gathered}$$

Thus we see that the bracket in $\hat{g}$ follows exactly the bracket in the above example where the irreducible representation on $\hat{r}$ is now the adjoint representation. We certainly can now conclude that under our hypotheses in $\hat{g}$ we have $D^1\hat{g} = \hat{g}$.

We proceed now to the proof of the Levi Decomposition theorem in the case where we are assuming that the radical $\hat{r}$ of $\hat{g}$ is abelian, and that $ad(\hat{g})$ acts irreducibly on $\hat{r}$, and is not equal to 0, i.e., $ad(\hat{g}) \cdot \hat{r} = [\hat{g}, \hat{r}] = \hat{r}$. We recall the proof of the theorem of the complete reducibility of a semisimple representation. (See 2.15.3.) The final step in that proof was a reduction in which we introduced linear spaces of maps as the representation spaces, and then chose certain special maps which would lead us to the conclusion we were seeking. Likewise we now do the same in our present proof. Our representation space will now be the set of endomorphisms of $\hat{g}$ into $\hat{g}$: $End(\hat{g}) = \widehat{gl}(\hat{g})$. Thus our representation ρ will be a map from $\hat{g}$ to the Lie algebra of brackets $\widehat{gl}(\widehat{gl}(\hat{g}))$ which is linear and takes brackets to brackets. Thus [using the symbol $X\cdot$ for $\rho(x)$] we have

$$\begin{gathered}\hat{g} \xrightarrow{\rho} \widehat{gl}(\widehat{gl}(\hat{g})) \\ x \longmapsto \rho(x) = X\cdot : \widehat{gl}(\hat{g}) \longrightarrow \widehat{gl}(\hat{g}) \\ \phi \longmapsto X \cdot \phi\end{gathered}$$

However we do have a natural representation in this case, the adjoint representation of $\widehat{gl}(\hat{g})$ on $\widehat{gl}(\hat{g})$. And since for x in $\hat{g}$, $ad(x)$ is in $\widehat{gl}(\hat{g})$, we take the adjoint representation of $ad(x)$ on $\widehat{gl}(\hat{g})$: $X\cdot := ad(ad(x))$.

$$\begin{gathered}\hat{g} \xrightarrow{\rho} \widehat{gl}(End(\hat{g})) \\ x \longmapsto \rho(x) = X\cdot := ad(ad(x)) : \widehat{gl}(\hat{g}) \longrightarrow \widehat{gl}(\hat{g}) \\ \phi \longmapsto X \cdot \phi = ad(ad(x))(\phi) = [ad(x), \phi] = ad(x)\phi - \phi(ad(x))\end{gathered}$$

If we let $X \cdot \phi$ act on y in $\hat{g}$, we have

$$(X \cdot \phi)(y) = [ad(x), \phi](y) = (ad(x)\phi - \phi(ad(x)))(y) = [x, \phi(y)] - \phi[x, y]$$

We would now like to make explicit this representation. First we show that $X\cdot$ is in $\widehat{gl}(\widehat{gl}(\hat{g}))$, that is, $X\cdot$ acts linearly on $\widehat{gl}(\hat{g})$.

$$\begin{gathered}X \cdot (\phi_1 + \phi_2) = ad(x)(\phi_1 + \phi_2) - (\phi_1 + \phi_2)(ad(x)) = \\ ad(x)(\phi_1) + ad(x)(\phi_2) - \phi_1(ad(x)) - \phi_2(ad(x)) = \\ ad(x)(\phi_1) - \phi_1(ad(x)) + ad(x)(\phi_2) - \phi_2(ad(x)) = \\ X \cdot \phi_1 + X \cdot \phi_2\end{gathered}$$

For c a scalar in $\mathbb{F}$, we have

$$\begin{gathered}X \cdot (c\phi) = ad(x)(c\phi) - (c\phi)(ad(x)) = c(ad(x)\phi) - c(\phi(ad(x))) = \\ c(ad(x)\phi - \phi(ad(x))) = c(X \cdot \phi)\end{gathered}$$

We conclude that $X\cdot$ does indeed belong to $\widehat{gl}(\widehat{gl}(\hat{g}))$

We now show that this map is linear, i.e.,

$$\begin{gathered}\rho(x_1 + x_2) = \rho(x_1) + \rho(x_2) \\ \rho(cx) = c\rho(x)\end{gathered}$$

(where c is in the scalar field $\mathbb{F}$) and that the brackets are preserved, i.e.,

$$\rho[x_1, x_2] = [\rho(x_1), \rho(x_2)]$$

We have

$$\begin{gathered}\rho(x_1 + x_2)(\phi) = [ad(x_1 + x_2), \phi] = [ad(x_1) + ad(x_2), \phi] = [ad(x_1), \phi] + \\ [ad(x_2), \phi] = X_1 \cdot \phi + X_2 \cdot \phi = \rho(x_1)(\phi) + \rho(x_2)(\phi) = (\rho(x_1) + \rho(x_2))(\phi)\end{gathered}$$

$$\rho(cx)(\phi) = [ad(cx), \phi] = [cad(x), \phi] = c[ad(x), \phi] = \\ c(X \cdot \phi) = c(\rho(x)\phi) = (c\rho(x))(\phi)$$

$$\rho[x_1, x_2](\phi) = [ad[x_1, x_2], \phi] = [[ad(x_1), ad(x_2)], \phi]$$

As would be expected, we now use the Jacobi identity. [In fact we have also used the Jacobi identity by writing $ad[x_1, x_2] = [ad(x_1), ad(x_2)]$, thus making the second time we have used this identity in this proof.]

$$[[ad(x_1), ad(x_2)], \phi] = [ad(x_1), [ad(x_2), \phi]] - [ad(x_2), [ad(x_1), \phi]] = \\ X_1 \cdot [ad(x_2), \phi] - X_2 \cdot [ad(x_1), \phi] = X_1 \cdot (X_2 \cdot \phi) - X_2 \cdot (X_1 \cdot \phi) = \\ ((X_1\cdot)(X_2\cdot))\phi - ((X_2\cdot)(X_1\cdot))\phi = [X_1\cdot, X_2\cdot]\phi = [\rho(x_1), \rho(x_2)]\phi$$

At this point we would like to remark that the subalgebra $\hat{r}$ also has a representation on $\widehat{gl}(\hat{g})$. Obviously we define it as follows.

$$\hat{r} \xrightarrow{\rho} \widehat{gl}(\widehat{gl}(\hat{g})) \\ d \longmapsto \rho(d) = D\cdot : \widehat{gl}(\hat{g}) \longrightarrow \widehat{gl}(\hat{g}) \\ \phi \longmapsto D \cdot \phi := [ad(d), \phi] = ad(d)\phi - \phi(ad(d))$$

If we let $D \cdot \phi$ act on y in $\hat{g}$, we have

$$(D \cdot \phi)(y) = [ad(d), \phi](y) = (ad(d)\phi - \phi(ad(d)))(y) = [d, \phi(y)] - \phi[d, y]$$

We remark that this representation is nothing but the representation ρ above restricted to the subalgebra $\hat{r}$. Since $\hat{r} \subset \hat{g}$, we know $ad(ad(\hat{r})) \subset ad(ad(\hat{g})) \subset \widehat{gl}(\widehat{gl}(\hat{g}))$, and thus $ad(ad(d))$ is in $\widehat{gl}(\widehat{gl}(\hat{g}))$ for all d in $\hat{r}$.

Now, as before, we define three special subspaces in $\widehat{gl}(\hat{g})$ and our arguments using these subspaces are very much the same as before. We think that going over the arguments serves here as a good method of consolidating understanding Moreover, once again in the following, the scalars, denoted by c, c_i's and the like, are in $\mathbb{F}$. Here are the subspaces:

$$C := \{\phi \in \widehat{gl}(\hat{g}) | \phi(\hat{g}) \subset \hat{r}; \phi|_{\hat{r}} = cI_{\hat{r}}\} \\ B := \{\phi \in \widehat{gl}(\hat{g}) | \phi(\hat{g}) \subset \hat{r}; \phi|_{\hat{r}} = 0\} \\ A := \{ad(x) \in \widehat{gl}(\hat{g}) | x \in \hat{r}\}$$

[We remark immediately that A is the image of $\hat{r}$ by ad:

$$\hat{r} \xrightarrow{ad} ad(\hat{r}) \subset (\widehat{gl}(\hat{g}))$$

and since $\hat{g}$ has no center, the map is an injection.]

The fact that these sets are subspaces is immediate. For C: we have for ϕ_1 and ϕ_2 in C and c and c_i scalars,

$$(\phi_1 + \phi_2)(\hat{g}) = (\phi_1)(\hat{g}) + (\phi_2)(\hat{g}) \subset (\hat{r} + \hat{r}) \subset \hat{r}$$
$$(c\phi)(\hat{g}) = c(\phi(\hat{g})) \subset (c\hat{r}) \subset \hat{r}$$

$$(\phi_1 + \phi_2)(\hat{r}) = (\phi_1)(\hat{r}) + (\phi_2)(\hat{r}) = (c_1 I_{\hat{r}})(\hat{r}) + (c_2 I_{\hat{r}})(\hat{r}) = ((c_1 + c_2) I_{\hat{r}})(\hat{r})$$
$$(c\phi)(\hat{r}) = c(\phi(\hat{r})) = (c(c_1 I_{\hat{r}}))(\hat{r}) = ((cc_1) I_{\hat{r}})(\hat{r})$$

Thus we conclude that C is a subspace of $\widehat{gl}(\hat{g})$. For B: we observe that we have the same calculations except that now all the scalar multiples, when restricted to $\hat{r}$, are mapped to 0, which fact obviously gives us the desired conclusion that B is a subspace of $\widehat{gl}(\hat{g})$. Finally for A: as stated in a comment above, this is indeed equal to the image of $\hat{r}$ in $\widehat{gl}(\hat{g})$ by the adjoint representation, and we know that such an image is a linear space.

We see immediately that B is contained in C [set $c = 0$]; and that A is contained in B since $(ad(t))x = [t, x] \in \hat{r}$, for we know that for t in $\hat{r}$ and x in $\hat{g}$ we have $[t, x] \in [\hat{r}, \hat{g}] = \hat{r}$; and $(ad(t))|_{\hat{r}})(s)$ for t in $\hat{r}$ and s in $\hat{r}$ means $[t, s]$ is in $[\hat{r}, \hat{r}] = 0$. [We remark that here, when treating A, we are using the full force of our hypotheses.] The set A indeed is equal to the image of $\hat{r}$ in $\widehat{gl}(\hat{g})$ by the adjoint representation under the given conditions, namely, $[\hat{r}, \hat{g}] = \hat{r}$ and $[\hat{r}, \hat{r}] = 0$. The set B is the set of all ϕ in $\widehat{gl}(\hat{g})$ which satisfies this condition, except now ϕ need not be an adjoint map $ad(t)$ for some t in $\hat{r}$. The set C is set of all ϕ in $\widehat{gl}(\hat{g})$ such the $\phi(\hat{g})$ is contained in $\hat{r}$ and is a scalar multiple when restricted to $\hat{r}$.

We know that $\hat{g}$ acts by ρ, the representation defined above, on $\widehat{gl}(\hat{g})$ and we will now show that each of these three subspaces of $\widehat{gl}(\hat{g})$ is invariant by $\rho(\hat{g})$.

First we want to show that $\rho(\hat{g})(C) \subset C$. If $X\cdot = \rho(x)$ for x in $\hat{g}$, and ϕ is in C, then we have for y in $\hat{g}$

$$(X \cdot \phi)(y) = [x, \phi(y)] - \phi[x, y]$$

But $\phi(y)$ is in $\hat{r}$, since ϕ is in C; and thus $[x, \phi(y)]$ is in $[\hat{g}, \hat{r}] = \hat{r}$. Also $\phi[x, y]$ is in $\hat{r}$ since $[x, y]$ is in $\hat{g}$ and $\phi(\hat{g})$ is contained in $\hat{r}$. Thus $[x, \phi(y)] - \phi[x, y]$ is in $\hat{r}$ and we conclude that $(X \cdot \phi)(\hat{g})$ is contained in $\hat{r}$ for all x in $\hat{g}$. We now restrict $\rho(\hat{g})(C)$ to $\hat{r}$. Letting s be in $\hat{r}$, we have

$$(X \cdot \phi)(s) = [x, \phi(s)] - \phi[x, s]$$

Since s is in $\hat{r}$, $\phi(s) = (cI|_{\hat{r}})(s) = cs$. Thus $[x, \phi(s)] = [x, cs] = c[x, s]$. Also since $[x, s]$ is in $[\hat{g}, \hat{r}] = \hat{r}$, $\phi[x, s] = (cI|_{\hat{r}})[x, s] = c[x, s]$. We conclude that

$$(X \cdot \phi)(s) = [x, \phi(s)] - \phi[x, s] = c[x, s] - c[x, s] = 0$$

and thus $(X \cdot \phi)(s) = 0 = (0I|_{\hat{r}})(s)$ for all s in $\hat{r}$. We can conclude that $\rho(\hat{g})(C) \subset C$. In fact we see that $\rho(\hat{g})(C) \subset B$.

Next we want to show that $\rho(\hat{g})(B) \subset B$. However, since B is contained in C, we conclude that $\rho(\hat{g})(B) \subset \rho(\hat{g})(C) \subset B$. If we analyze this calculation, as done above, we see that the first part — analyzing $(X \cdot \phi)(y)$ — just repeats itself; and in second part — $(X \cdot \phi)(s)$ — we observe all the elements are mapped to zero. Thus we can again immediately conclude that $\rho(\hat{g})(B) \subset B$.

Finally, we show that $\rho(\hat{g})(A) \subset A$. If we take $X\cdot = \rho(x)$ for all x in $\hat{g}$ and $ad(t)$ in A for all t in $\hat{r}$, then for y in $\hat{g}$ we have

$$(X \cdot ad(t))(y) = [x, ad(t)(y)] - ad(t)([x, y]) = [x, [t, y]] - [t, [x, y]]$$

Then using the Jacobi identity in $\hat{g}$.

$$[x, [t, y]] - [t, [x, y]] = [x, [t, y]] + [t, [y, x]] = -[y, [x, t]] = [[x, t], y] = ad([x, t])(y)$$

we thus have

$$X \cdot ad(t) = ad[x, t] \in A$$

since t is in $\hat{r}$ and thus $[x, t]$ is in $\hat{r}$. We can conclude that $ad([x, t])$ is in A, and this says that $\rho(\hat{g})(A) \subset A$. [Again we remark that in treating the set A, we are using the full structure of the Lie algebra since we are here using the Jacobi identity.] Also we note here that the properties of this set A will be vital to our proof of the final phase of the Levi decomposition theorem.

Next we show that each of the three subspaces C, B and A of $\widehat{gl}(\hat{g})$ is invariant by $\rho(\hat{r})$. But we know that since $\hat{r} \subset \hat{g}$, we have immediately $\rho(\hat{r}) \subset \rho(\hat{g})$. Thus we have $\rho(\hat{r})(C) \subset \rho(\hat{g})(C) \subset C$. Likewise we can conclude that $\rho(\hat{g})(B) \subset B$ and $\rho(\hat{g})(A) \subset A$.

However, how $\rho(\hat{r})$ acts on these three subsets gives more information than just keeping these three subsets of $\widehat{gl}(\hat{g})$ invariant. In fact how $\rho(\hat{r})$ acts on C will be an important piece of information as we conclude this proof later on. Thus, we calculate $D\cdot = \rho(d)$ for d in $\hat{r}$, and ϕ is in C. We have for y in $\hat{g}$

$$(D \cdot \phi)(y) = [d, \phi(y)] - \phi[d, y]$$

But $\phi(y)$ is in $\hat{r}$, since ϕ is in C; and thus $[d, \phi(y)]$ is in $[\hat{r}, \hat{r}] = 0$. Also $\phi[d, y]$ is in $\hat{r}$ since $[d, y]$ is in $\hat{r}$ and $\phi(\hat{r})$ is $(cI|_{\hat{r}})(\hat{r}) = c\hat{r}$. Thus

$$(D \cdot \phi)(y) = [d, \phi(y)] - \phi[d, y] = 0 - c[d, y] = -c(ad(d)y)$$

giving $D \cdot \phi = -c(ad(d))$ with c depending on the function ϕ. And we can conclude that $\rho(d)$ for d in $\hat{r}$ acting on C actually gives us an element in A since $\rho(\hat{r})(C) \subset A$. Since we already know that C contains B and B contains A, obviously we then have $\rho(\hat{r})(B) \subset A$ and $\rho(\hat{r})(A) \subset A$.

However on analyzing the fact that $\rho(\hat{r})(A) \subset A$, we meet something vital for our proof. If we take $D \cdot = \rho(d)$ and $ad(t)$ in A for all d and t in $\hat{r}$, then for y in $\hat{g}$ we have

$$(D \cdot ad(t))(y) = [d, ad(t)(y)] - ad(t)([d, y]) = [d, [t, y]] - [t, [d, y]]$$

Now since d and t are in $\hat{r}$ and y is in $\hat{g}$, we have $[t, y]$ in $[\hat{r}, \hat{g}] = \hat{r}$, and thus $[d, [t, y]]$ is in $[\hat{r}, \hat{r}] = 0$. Also $[d, y]$ is in $\hat{r}$, and thus $[t, [d, y]] = 0$, giving $(D \cdot ad(t))(y) = 0$. Since we want $D \cdot ad(t)$ to be some $ad(s)$ for some s in $\hat{r}$, this says that we can take any s in $\hat{r}$ such that $ad(s) = 0$. But we know that the map ad restricted to $\hat{r}$ is injective. [The center is 0, and thus ad is injective.] Thus for each y in $\hat{g}$ there is one and only one element s in $\hat{r}$ which maps to zero, that is, $s = 0$, and we know that $D \cdot ad(t) = ad(0) = 0$.

Let us see what happens if we use the Jacobi identity in $\hat{g}$.

$$[d, [t, y]] - [t, [d, y]] = [d, [t, y]] + [t, [y, d]] = -[y, [d, t]] = [[d, t], y] = ad([d, t])(y)$$

We thus have $D \cdot ad(t) = ad([d, t]) \in A$. But again we observe that $[t, y]$ is in $\hat{r}$, and $[d, [t, y]]$ is 0; and $[y, d]$ is in $\hat{r}$ and $[t, [y, d]]$ is 0. Also $[d, t]$ is 0 and thus $ad([d, t]) = 0$. This calculation immediately gives the fact that $d = 0$ is the only element in $\hat{r}$ which satisfies $D \cdot ad(t) = ad([d, t]) \in A$.

We thus have $\rho(\hat{r})$ acting on the three subsets C, B and A of $\widehat{gl}(\hat{g})$ in an invariant manner. We now form the quotient spaces C/A and C/B. Since

$$C := \{\phi \in \widehat{gl}(\hat{g}) | \phi(\hat{g}) \subset \hat{r}; \phi|_{\hat{r}} = cI_{\hat{r}}\}$$

and

$$B := \{\phi \in \widehat{gl}(\hat{g}) | \phi(\hat{g}) \subset \hat{r}; \phi|_{\hat{r}} = 0\}$$

we see immediately that $C = \{cI_{\hat{r}}\} + B$, and we conclude that C/B is one-dimensional.

We now claim that C/A and C/B are representation spaces for $\hat{g}$. This is reasonable since C, B, and A are invariant by $\rho(\hat{g})$ and thus are representation spaces for $\hat{g}$. We define a representation of $\hat{g}$ on C/A. If we let ρ' be this representation map, then $\rho'(\hat{g})$ is in $\widehat{gl}(C/A)$. [Below, again as before, we use the symbol $X' \cdot$ for $\rho'(x)$]. Thus

$$\begin{aligned} \hat{g} &\longrightarrow \rho'(\hat{g}) \\ x \longmapsto \rho'(x) &= X'\cdot : C/A \longrightarrow C/A \\ &\qquad \phi + A \longmapsto X' \cdot (\phi + A) := (X \cdot \phi) + A \end{aligned}$$

We observe that the symbol $X' \cdot (\phi + A)$ can also be read as $X \cdot (\phi + A) = (X \cdot \phi) + (X \cdot A)$ since $\phi + A$ is just a subset of C, and we know how $X\cdot$ acts on C. Thus we have $X' \cdot (\phi + A) = (X \cdot \phi) + (X \cdot A) = (X \cdot \phi) + A$, since we know that $X \cdot A$ is in A. Now, again by 2.7.2, we know that we have a valid definition. Likewise if we let ρ' be this representation map for $\hat{g}$ in C/B, then $\rho'(\hat{g})$ is in $\widehat{gl}(C/B)$. Again

$$\begin{aligned} \hat{g} &\longrightarrow \rho'(\hat{g}) \\ x \longmapsto \rho'(x) &= X'\cdot : C/B \longrightarrow C/B \\ &\qquad \phi + B \longmapsto X' \cdot (\phi + B) := (X \cdot \phi) + B \end{aligned}$$

Again we observe that $X' \cdot (\phi + B)$ can also be read as $X \cdot (\phi + B) = (X \cdot \phi) + (X \cdot B)$ since $\phi + B$ is just a subset of C, and we know how $X\cdot$ acts on C. Thus $X' \cdot (\phi + B) = (X \cdot \phi) + (X \cdot B) = (X \cdot \phi) + B$, since we know that $X \cdot B$ is in B. Again by 2.7.2, we know that we have a valid definition.

Now from the above definitions we know that these maps are in $\widehat{gl}(C/A)$ and $\widehat{gl}(C/B)$. Next we show how ρ' takes addition in $\hat{g}$ to addition in $\widehat{gl}(C/A)$:

$$\rho'(x_1 + x_2) = \rho'(x_1) + \rho'(x_2)$$

We remark that we are just using cosets in C. Thus we have by the definition of ρ'

$$\rho'((x_1 + x_2)(\phi + A) = \rho(x_1 + x_2)(\phi) + A$$

Continuing, we have

$$\begin{aligned} \rho(x_1 + x_2)(\phi) + A &= (\rho(x_1) + \rho(x_2))(\phi) + A \\ &= \rho(x_1)(\phi) + \rho(x_2)(\phi) + A \\ &= \rho(x_1)(\phi) + A + \rho(x_2)(\phi) + A \\ &= \rho'(x_1)(\phi + A) + \rho'(x_2)(\phi + A) \\ &= (\rho'(x_1) + \rho'(x_2))(\phi + A) \end{aligned}$$

and thus giving us the linearity of ρ' with respect to addition.

$$\rho'(x_1 + x_2) = \rho'(x_1) + \rho'(x_2)$$

Also, with c in $\mathbb{F}$

$$
\begin{aligned}
\rho'(c(x_1))(\phi + A) &= \rho(c(x_1))(\phi) + A \\
&= c\rho(x_1)(\phi) + cA \\
&= c(\rho(x_1)(\phi) + A)) \\
&= c(\rho'(x_1)(\phi + A) \\
&= c(\rho'(x_1))(\phi + A)
\end{aligned}
$$

Hence we can conclude that we have the linearity of ρ' with respect to scalar multiplication.

Finally we show that brackets go to brackets:

$$\rho'[x_1, x_2] = [\rho'(x_1), \rho'(x_2)]$$

$$
\begin{aligned}
(\rho'[x_1, x_2])(\phi + A) &= (\rho[x_1, x_2])(\phi) + A \\
&= ([\rho(x_1), \rho(x_2)])(\phi) + A \\
&= (\rho(x_1)\rho(x_2) - \rho(x_2)\rho(x_1))(\phi) + A \\
&= \rho(x_1)(\rho(x_2)(\phi) - \rho(x_2)(\rho(x_1)(\phi) + A \\
&= \rho(x_1)(\rho(x_2)(\phi) + A - \rho(x_2)(\rho(x_1)(\phi) + A \\
&= \rho'(x_1)((\rho(x_2)(\phi) + A) - \rho'(x_2)((\rho(x_1)(\phi) + A) \\
&= \rho'(x_1)\rho'(x_2)(\phi + A) - \rho'(x_2)\rho'(x_1)(\phi + A) \\
&= (\rho'(x_1)\rho'(x_2) - \rho'(x_2)\rho'(x_1))(\phi + A) \\
&= ([\rho'(x_1), (\rho'(x_2)])(\phi + A)
\end{aligned}
$$

giving

$$\rho'[x_1, x_2] = [\rho'(x_1), \rho'(x_2)]$$

It is evident that we can make the same calculations for C/B and show that the map ρ' from $\hat{g}$ to $\widehat{gl}(C/B)$ takes addition in $\hat{g}$ to addition in $\widehat{gl}(C/B)$; takes scalar multiplication in $\hat{g}$ to scalar multiplication in $\widehat{gl}(C/B)$; and takes brackets in $\hat{g}$ to brackets in $\widehat{gl}(C/B)$.

But what we are seeking is a representation of the Lie algebra $\hat{g}/\hat{r}$ in C/A and C/B. Again since $\hat{r}$ is a subalgebra of $\hat{g}$, we know immediately from the above that $\hat{r}$ has a representation in C/A and C/B since $\hat{r}$ also leaves invariant the spaces C, B and A. Now by using cosets in $\hat{g}$ and in C, we can show that $\hat{g}/\hat{r}$ has indeed a representation in C/A and in C/B.

Thus we define a representation of $\hat{g}/\hat{r}$ on C/A. We again let ρ' be this representation map, and thus $\rho'(\hat{g}/\hat{r})$ is in $\widehat{gl}(C/A)$. We have

$$
\begin{aligned}
\hat{g}/\hat{r} &\longrightarrow \rho'(\hat{g}/\hat{r}) \\
x + \hat{r} &\longmapsto \rho'(x + \hat{r}) : C/A \longrightarrow C/A \\
&\qquad\qquad \phi + A \longmapsto \rho'(x + \hat{r})(\phi + A) := \rho(x + \hat{r})(\phi) + A
\end{aligned}
$$

We observe that the symbol $\rho'(x+\hat{r})(\phi+A)$ can also be read as $\rho(x+\hat{r})(\phi)+\rho(x+\hat{r})(A)$. Now since $\rho(x)$ and $\rho(\hat{r})$ leave A invariant, we do indeed get $\rho(x+\hat{r})(\phi)+A$ as claimed, and we see that we have a valid definition. We need to show how ρ' takes addition in $\hat{g}/\hat{r}$ to addition in $\widehat{gl}(C/A)$:

$$\rho'((x_1+\hat{r})+(x_2+\hat{r})) = \rho'(x_1+\hat{r})+\rho'(x_2+\hat{r})$$

Thus we have by the definition of ρ'

$$\rho'((x_1+\hat{r})+(x_2+\hat{r}))(\phi+A) = \rho((x_1+\hat{r})+(x_2+\hat{r}))(\phi)+A$$

Continuing, we have

$$\begin{aligned}\rho((x_1+\hat{r})+(x_2+\hat{r}))(\phi)+A &= \rho(x_1+\hat{r})(\phi)+\rho(x_2+\hat{r})(\phi)+A\\ &= \rho(x_1+\hat{r})(\phi)+A+\rho(x_2+\hat{r})(\phi)+A\\ &= \rho'(x_1+\hat{r})(\phi+A)+\rho'(x_2+\hat{r})(\phi+A)\\ &= (\rho'(x_1+\hat{r})+\rho'(x_2+\hat{r}))(\phi+A)\end{aligned}$$

giving

$$\rho'((x_1+\hat{r})+(x_2+\hat{r}))(\phi+A) = (\rho'(x_1+\hat{r})+\rho'(x_2+\hat{r}))(\phi+A)$$

and we can conclude that we have linearity of ρ' with respect to addition.

Also, with c in $\mathbb{F}$

$$\begin{aligned}\rho'(c(x_1+\hat{r}))(\phi+A) &= \rho(c(x_1+\hat{r}))(\phi)+A\\ &= (c\rho(x_1+\hat{r}))(\phi)+cA\\ &= c(\rho(x_1+\hat{r})(\phi)+A))\\ &= c(\rho'(x_1+\hat{r}))(\phi+A)\\ &= ((c\rho')(x_1+\hat{r}))(\phi+A)\end{aligned}$$

and we can conclude that we have linearity of ρ' with respect to scalar multiplication.

Finally, we show that brackets go into brackets.

$$\begin{aligned}(\rho'[x_1+\hat{r},x_2+\hat{r}])(\phi+A) &= (\rho[x_1+\hat{r},x_2+\hat{r}])(\phi)+A\\ &= ([\rho(x_1+\hat{r}),\rho(x_2+\hat{r}])(\phi)+A\\ &= (\rho(x_1+\hat{r})\rho(x_2+\hat{r})-\rho(x_2+\hat{r})\rho(x_1+\hat{r}))(\phi)+A\\ &= (\rho(x_1+\hat{r})\rho(x_2+\hat{r}))(\phi)-(\rho(x_2+\hat{r})\rho(x_1+\hat{r}))(\phi)+A\\ &= (\rho(x_1+\hat{r})\rho(x_2+\hat{r}))(\phi)+A-(\rho(x_2+\hat{r})\rho(x_1+\hat{r}))(\phi)+A\\ &= \rho'(x_1+\hat{r})(\rho(x_2+\hat{r})(\phi)+A)-\rho'(x_2+\hat{r})(\rho(x_1+\hat{r})(\phi)+A)\\ &= \rho'(x_1+\hat{r})(\rho'(x_2+\hat{r})(\phi+A))-\rho'(x_2+\hat{r})(\rho'(x_1+\hat{r})(\phi+A))\\ &= (\rho'(x_1+\hat{r})(\rho'(x_2+\hat{r}))(\phi+A)-(\rho'(x_2+\hat{r})(\rho'(x_1+\hat{r}))(\phi+A)\\ &= (\rho'(x_1+\hat{r})(\rho'(x_2+\hat{r})-\rho'(x_2+\hat{r})(\rho'(x_1+\hat{r}))(\phi+A)\\ &= ([\rho'(x_1+\hat{r}),\rho'(x_2+\hat{r}])(\phi+A)\end{aligned}$$

and thus we have

$$\rho'[x_1 + \hat{r}, x_2 + \hat{r}] = [\rho'(x_1 + \hat{r}), \rho'(x_2 + \hat{r})]$$

It is evident that we can make the same calculations for C/B. The map ρ' from $\hat{g}$ to $\widehat{gl}(C/B)$

$$\begin{aligned} \hat{g}/\hat{r} &\longrightarrow \rho'(\hat{g}/\hat{r}) \\ x + \hat{r} &\longmapsto \rho'(x + \hat{r}) : C/B \longrightarrow C/B \\ &\qquad \phi + B \longmapsto \rho'(x + \hat{r})(\phi + B) := \rho(x + \hat{r})(\phi) + B \end{aligned}$$

takes addition in $\hat{g}$ to addition in $\widehat{gl}(C/B)$; takes scalar multiplication in $\hat{g}$ to scalar multiplication in $\widehat{gl}(C/B)$; and takes brackets in $\hat{g}$ to brackets in $\widehat{gl}(C/B)$.

With respect to the representation of $\hat{g}/\hat{r}$ on C/B, we can make one more remark. Since C/B is one-dimensional and $\hat{g}/\hat{r}$ is semisimple, the only possible representation of $\hat{g}/\hat{r}$ on C/B is the zero representation.

$$\begin{aligned} \hat{g}/\hat{r} &\longrightarrow \rho'(\hat{g}/\hat{r}) \\ x + \hat{r} &\longmapsto \rho'(x + \hat{r}) : C/B \longrightarrow C/B \\ &\qquad \phi + B \longmapsto \rho'(x + \hat{r})(\phi + B) = 0(\phi + B) = B \end{aligned}$$

since the zero coset of C/B is B.

Having established representations of $\hat{g}/\hat{r}$ on C/A and C/B, we now seek a linear map σ of C/A to C/B which preserves these representations in the sense that σ commutes with these representations. If such a map can be defined, we know that the image and kernel of this map will be invariant subspaces respectively of C/B and C/A by the representation of $\hat{g}/\hat{r}$. We define σ as the natural coset map

$$\begin{aligned} C/A &\xrightarrow{\sigma} C/B \\ \phi + A &\longmapsto \sigma(\phi + A) := \phi + B \end{aligned}$$

Since A is contained in B, the map is well defined. Let ϕ_1 and ϕ_2 be two elements of C which are in the same coset, which means that $\phi_1 - \phi_2$ is in A. Now $\sigma(\phi_1 + A) = \phi_1 + B$ and $\sigma(\phi_2 + A) = \phi_2 + B$. We want to prove that ϕ_1 and ϕ_2 are in the same coset of C/B. We take their difference: $\phi_1 - \phi_2$, which is in A. But A is contained in B. Thus $\phi_1 - \phi_2$ is contained in B, which says that ϕ_1 and ϕ_2 are in the same coset of C/B.

The map is obviously linear and it preserves sums and scalar multiples, as can be seen from:

$$\sigma((\phi_1 + A) + (\phi_2 + A)) = \sigma((\phi_1 + \phi_2) + A) = (\phi_1 + \phi_2) + B =$$
$$(\phi_1 + B) + (\phi_2 + B) = \sigma(\phi_1 + A) + \sigma(\phi_2 + A)$$

and

$$\sigma(c(\phi + A)) = \sigma(c\phi + A) = c\phi + B = c\phi + cB = c(\phi + B) = c(\sigma(\phi + A))$$

for any ϕ_1 and ϕ_2 in C and any c in $\mathbb{F}$.

And finally, we see that the σ map respects all of the above representations in the sense that it commutes with them. We have representations, denoted by ρ', of $\hat{g}$ on C/A and C/B as well as of $\hat{r}$ and of $\hat{g}/\hat{r}$. Thus we assert that

$$(\rho'(\hat{g}))(\sigma) = \sigma(\rho'(\hat{g}))$$

since for any x in $\hat{g}$ and ϕ in C, we have

$$\rho'(\hat{x})(\sigma(\phi + A))) = X' \cdot (\sigma(\phi + A)) = X' \cdot (\phi + B) = X \cdot \phi + B$$

and

$$\sigma(\rho'(x)(\phi + A)) = \sigma(X' \cdot (\phi + A)) = \sigma((X \cdot \phi) + A) = X \cdot \phi + B$$

We conclude that the map σ commutes with the two representations of $\hat{g}$.

We also observe that for x in $\hat{g}$

$$\rho'(x)(\phi + B) = X' \cdot (\phi + B) = X \cdot \phi + B \subset B$$

since $X \cdot \phi \subset B$.

We also assert that the σ map respects the representation of $\hat{r}$ on C/A and C/B in the sense that it commutes with them:

$$(\rho'(\hat{r}))(\sigma) = \sigma(\rho'(\hat{r}))$$

Here is proof. For any d in $\hat{r}$ and ϕ in C,

$$\rho'(d)(\sigma(\phi + A)) = D' \cdot (\sigma(\phi + A)) = D' \cdot (\phi + B) = D \cdot \phi + B$$

and

$$\sigma(\rho'(d)(\phi + A)) = \sigma(D' \cdot (\phi + A)) = \sigma((D \cdot \phi) + A) = D \cdot \phi + B$$

We conclude that the map σ commutes with the two representations of $\hat{r}$.

We also observe that for d in $\hat{r}$

$$\rho'(d)(\phi + B) = D' \cdot (\phi + B) = D \cdot \phi + B \subset A + B \subset B$$

since $D \cdot C \subset A$.

Thus it is evident that the σ map also respects the representation of $\hat{g}/\hat{r}$ on C/A and C/B, i.e.,

$$\rho'(\hat{g}/\hat{r})(\sigma) = \sigma(\rho'(\hat{g}/\hat{r}))$$

For any $x + \hat{r}$ in $\hat{g}/\hat{r}$ and $\phi + A$ in C/A, we have

$$\rho'(\hat{x} + \hat{r})(\sigma(\phi + A)) = \rho'(\hat{x} + \hat{r})(\phi + B) = \rho(x + \hat{r})(\phi) + B$$

and

$$\sigma(\rho'(x + \hat{r})(\phi + A)) = \sigma(\rho(x + \hat{r})(\phi) + A) = \rho(x + \hat{r})(\phi) + B$$

We conclude that the map σ commutes with the two representations of $\hat{g}/\hat{r}$. And we again remark that

$$\rho'(\hat{x} + \hat{r})(\phi + B) = \rho(x + \hat{r})(\phi) + B \subset B$$

We return now to the map σ.

$$\begin{gathered} C/A \xrightarrow{\sigma} C/B \\ \phi + A \longmapsto \sigma(\phi + A) := \phi + B \end{gathered}$$

From the definitions of C, B, and A, we see immediately that σ is surjective, and indeed surjects onto a one-dimensional space C/B. We see that C surjects onto B: for any ϕ in C, there is a ϕ' in B with the same action; and acting on $\hat{r}$, any ϕ in C is a scalar map, while the only scalar map in B is the 0-scalar map. To help see this, recall that

$$\begin{gathered} C := \{\phi \in \widehat{gl}(\hat{g}) | \phi(\hat{g}) \subset \hat{r}; \phi|_{\hat{r}} = cI_{\hat{r}}\} \\ B := \{\phi \in \widehat{gl}(\hat{g}) | \phi(\hat{g}) \subset \hat{r}; \phi|_{\hat{r}} = 0\} \end{gathered}$$

Thus, as we remarked above, the image set of σ is C/B and is one-dimensional. Now we know that the representation of $\hat{g}/\hat{r}$ leaves invariant this image space. From the above calculations we know that

$$\rho'(\hat{x} + \hat{r})(\phi + B) = \rho(x + \hat{r})(\phi) + B \subset B$$

and this says that $\hat{g}/\hat{r}$ takes C/B to the zero coset B which is contained in C/B, and thus leaves invariant the image space C/B. But it also says that the representation of $\hat{g}/\hat{r}$ on C/B is the zero representation. We remark that we have already arrived at this conclusion since we know that $\hat{g}/\hat{r}$ is a semisimple Lie algebra and that is one-dimensional.

We also know that any map which commutes with the representations leaves invariant the kernel of the map. Let us confirm this statement. The kernel of the σ map is the set of all the cosets that map to B, which obviously are the cosets B/A of C/A. Now $\hat{g}/\hat{r}$ acting on B/A gives, for ϕ in B,

$$\rho'(\hat{x}+\hat{r})(\phi+A) = \rho(x+\hat{r})(\phi)+A =$$
$$[ad(x+\hat{r}),\phi]+A = [ad(x),\phi]+[ad(\hat{r}),\phi]+A$$

Now for any y in $\hat{g}$, we have $[ad(x),\phi](y) = ad(x)(\phi(y)) - \phi(ad(x)(y))$. We know that $\phi(y)$ in $\hat{r}$ and $ad(x)(\hat{r})$ is in $\hat{r}$. Also $ad(x)(y)$ is in $\hat{g}$ and ϕ on $\hat{g}$ is in $\hat{r}$. Now for any s in $\hat{r}$, we have $[ad(x),\phi](s) = ad(x)(\phi(s)) - \phi(ad(x)(s))$. Since $\phi(s) = 0$, $ad(x)(\phi(s)) = 0$. Since $[x,s]$ is in $\hat{r}$, $\phi(ad(x)(s)) = 0$. Thus $[ad(x),\phi](s) = 0-0 = 0$, and we can conclude that $[ad(x),\phi]$ is in B.

Again for any y in $\hat{g}$, we have $[ad(\hat{r})),\phi](y) = ad(\hat{r})(\phi(y)) - \phi(ad(\hat{r})(y))$. We know $\phi(y)$ is in $\hat{r}$ and $ad(\hat{r})(\hat{r})$ is 0. Also $ad(\hat{r})(y)$ is in $\hat{r}$ and ϕ on $\hat{r}$ is 0. Finally for any s in $\hat{r}$ we have $[ad(\hat{r}),\phi](s) = ad(\hat{r})(\phi(s)) - \phi(ad(\hat{r})(s))$. Since $\phi(s)$ is 0, we have that $ad(\hat{r})(0) = 0$. Since $ad(\hat{r})(s) = 0$, we have that $\phi(0) = 0$ and we conclude that $[ad(\hat{r})),\phi]$ is in B. Thus $\rho'(B/A) \subset B/A$, and we confirm that the representation $\hat{g}/\hat{r}$ leaves invariant the kernel of the map σ.

We are now at the following point. We have a subspace of (C/A), the kernel of σ, and we know that $(C/A) = ker(\sigma) \oplus S$ where S is any subspace of C/A of dimension equal to one and not contained in $ker(\sigma)$, i.e., S is $\{c\phi_0\}+A$ and $\sigma(S) = C/B$ for some ϕ_0 in C. We also know that $ker(\sigma)$ is an invariant subset by the action of the semisimple Lie algebra $\hat{g}/\hat{r}$. Now of all the one-dimensional subspaces S, we are obviously seeking those that respect the action of $\hat{g}/\hat{r}$ on S, that is, that are invariant by this action.

Now we know from 2.15.3 that representations of semisimple Lie algebras have the complete reducibility property. Thus we know that an S exists such that $\hat{g}/\hat{r}$ acting on S leaves S invariant. [There is no reason, by the way, to assert that there is only one unique S.] This says the $\rho'(\hat{g}/\hat{r})(S) \subset S$. Now we know that S is also one-dimensional and that the only representation of a semisimple Lie algebra on a one-dimensional space is the zero representation. Thus we seek a ϕ_0 in C such that $\rho(x+\hat{r})((c\phi_0)+A) \subset A$. [Obviously we can omit the scalar in ϕ_0 and write $\rho(x+\hat{r})(\phi_0+A) \subset A$, where ϕ_0 is in C and not in B, and where A is the zero coset of C/A.] This, of course, is the condition that identifies the kind of one-dimensional subspace S that we seek.

We can reduce this information as follows. For x in $\hat{g}$

$$\begin{gathered}
\rho(x+\hat{r})(\phi_0+A)\subset A\\
(\rho(x)+\rho(\hat{r}))(\phi_0+A)\subset A\\
\rho(x)(\phi_0)+\rho(x)(A)+(\rho(\hat{r}))(\phi_0)+(\rho(\hat{r}))(A)\subset A\\
X\cdot\phi_0+D\cdot\phi_0+A\subset A
\end{gathered}$$

since $\rho(\hat{g})$ leaves A invariant . Now let x be an arbitrary element of $\hat{g}$. Then we can just write for all x in $\hat{g}$ that $X\cdot\phi_0+A\subset A$ for some ϕ_0 in C but not in B. This says that $[ad(x),\phi_0]=ad(a)$ for some a in $\hat{r}$. We observe that $[ad(x),\phi_0]=X\cdot\phi_0$ is in B, since $\rho(\hat{g})(C)\subset B$. But since $A\subset B$, we seek a ϕ_0 in C but not in B for which, for x in $\hat{g}$, $\rho(x)(\phi_0)$ is in $A\subset B$. We remark that we already know that for all d in $\hat{r}$, $\rho(d)(\phi_0)$ is in $A\subset B$, since this is true for all ϕ in C, i.e., $(\rho(\hat{r}))(C)\subset A$. Therefore what is new is that we want to chose a ϕ_0 in C such that this is also true for all x in $\hat{g}$.

Now we can identify at least one ϕ_0. Since ϕ_0 is in C and not in B, we let ϕ_0 restricted to $\hat{r}$ be the identity transformation on $\hat{r}$.

$$\begin{aligned}
\phi_0:\hat{r}&\longrightarrow(\phi_0)(\hat{r})\\
t&\longmapsto(\phi_0)(t):=I(t)=t
\end{aligned}$$

To define ϕ_0 on $\hat{g}$ but not in $\hat{r}$ we choose an arbitrary d_0 in $\hat{r}$.

$$\begin{aligned}
\phi_0:\hat{g}\backslash\hat{r}&\longrightarrow(\phi_0)(\hat{g}\backslash\hat{r})\\
x&\longmapsto(\phi_0)(x):=ad(d_0)\cdot x=[d_0,x]\in\hat{r}
\end{aligned}$$

In matrix notation, for any basis in $\hat{r}$ and any basis in $\hat{g}\backslash\hat{r}$, we would obtain

$$\phi_0=\begin{bmatrix} I & ad(d_0)\\ 0 & 0\end{bmatrix}$$

where the square matrix I is independent of the basis chosen in $\hat{r}$, while the rectangular matrix $ad(d_0)$ would depend on the basis chosen in $\hat{g}\backslash\hat{r}$. Now for any x in $\hat{g}\backslash\hat{r}$ we calculate

$$\rho(x)(\phi_0)=[ad(x),\phi_0]=ad(x)\circ\phi_0-\phi_0\circ ad(x)$$

Operating $\rho(x)(\phi_0)$ on t in $\hat{r}$, we have

$$\begin{gathered}
(ad(x)\circ\phi_0-\phi_0\circ ad(x))(t)=(ad(x)\circ\phi_0)(t)-(\phi_0\circ ad(x)(t)=\\
(ad(x)\circ I)(t)-(I\circ ad(x)(t)=[x,t]-[x,t]=0
\end{gathered}$$

since $ad(x)(t) = [x,t]$ is in $\hat{r}$. We conclude that $\rho(x)(\phi_0)$ is in B. Operating $\rho(x)(\phi_0)$ on y in $\hat{g}\backslash\hat{r}$, we have

$$\rho(x)(ad(d_0)) = [ad(x), ad(d_0)] = ad[x, d_0] \in ad(\hat{r})$$

which says that $\rho(\hat{g})(\phi_0)$ is in A, which is the conclusion which we have been seeking.

Thus we have identified at least one ϕ_0 in C not in B such that $\rho(x+\hat{r})(\phi_0 + A)$ is in A for all x in $\hat{g}$. But now we ask how does this give us information about $\hat{g}$? Well, $\rho(x+\hat{r})(\phi_0 + A)$ being in A for all x in $\hat{g}$ says that, choosing an x in $\hat{g}$, we have $[ad(x), \phi_0] = ad(a)$ for some a in $\hat{r}$. Now we know that for any ϕ in C, $[ad(r), \phi]$ is in A since $\rho(\hat{r})(C) \subset A$. First let us take a d in $\hat{r}$ and see what element a in $\hat{r}$ corresponds to it. But we already made this calculation when we were showing that $\hat{r}$ left invariant the subspace C. If $D\cdot = \rho(d)$ for d in $\hat{r}$, and ϕ_0 is in C, we have for y in $\hat{g}$

$$(D \cdot \phi_0)(y) = [ad(d), \phi_0](y)] = (ad(d) \circ \phi_0 - \phi_0 \circ ad(d))(y) = \\ [d, \phi_0(y)] - \phi_0([d,y])$$

But $\phi_0(y)$ is in $\hat{r}$, since ϕ_0 is in C; and thus $[d, \phi_0(y)]$ is in $[\hat{r}, \hat{r}] = 0$. Also $\phi_0([d,y]) = I_{\hat{r}}([d,y])$ since $[d,y]$ is in $\hat{r}$ and $\phi_0|\hat{r}$ is the identity on $\hat{r}$. And thus we have our conclusion that $D \cdot \phi_0 = -ad(d) = ad(-d)$. This is a satisfying result for it says that a d in $\hat{r}$, by this construction, determines $ad(-d)$ in $ad(\hat{r})$, and thus this construction injects $\hat{r}$ into $ad(\hat{r})$. [Note that we already know this from the fact that $\hat{g}$ has no center and thus ad is an injection of $\hat{g}$ into $ad(\hat{g})$.] But this is true for any ϕ in C. What are we then adding when we choose our particular ϕ_0? We are asking when for x in $\hat{g}\backslash\hat{r}$ do we have $[ad(x), \phi_0]$ in $ad(\hat{r})$? Since ad is injective, the only element x in $\hat{g}\backslash\hat{r}$ that injects into $ad(\hat{r})$ is the 0 element. This is saying that we have a map

$$\hat{g} \longrightarrow \hat{r}$$

such that $\hat{r}$ injects onto $\hat{r}$ and $\hat{g}\backslash\hat{r}$ maps to 0.

Now we can show that this map is a Lie algebra morphism! Since $\hat{r}$ injects onto $\hat{r}$, our conclusion that we have a Lie algebra morphism on $\hat{r}$ is immediate. We define an $\hat{l}$ in $\hat{g}\backslash\hat{r}$ such that the representation map $\rho(\hat{g})$ restricted to $\hat{l}$ acting on ϕ_0 is 0:

$$\hat{l} := \{p \in \hat{g} \mid P \cdot \phi_0 = 0\}$$

or equivalently

$$\hat{l} := \{p \in \hat{g} \mid [ad(p), \phi_0] = 0\}$$

Here we are using our conclusion from above that $[ad(p), \phi_0] = ad(a) = 0$ means that a in $\hat{r}$ is itself 0. Now obviously $\hat{l}$ is a linear subspace of $\hat{g}$. For c in $\mathbb{F}$ and P, P_1 and P_2 in $\hat{l}$ we have

$$(P_1 + P_2) \cdot \phi_0 = (P_1 \cdot + P_2 \cdot)\phi_0 = P_1 \cdot \phi_0 + P_2 \cdot \phi_0 = 0 + 0 = 0$$

$$(cP) \cdot \phi_0 = c(P \cdot \phi_0) = c0 = 0$$

We show also that brackets close. For p and q in $\hat{l}$, we have

$$\rho([p,q])\phi_0 = [\rho(p), \rho(q)]\phi_0 = [P\cdot, Q\cdot]\phi_0 =$$
$$(P \cdot Q \cdot - Q \cdot P \cdot)\phi_0 =$$
$$(P \cdot Q \cdot)\phi_0 - (Q \cdot P \cdot)\phi_0 = P \cdot (Q \cdot \phi_0) - Q \cdot (P \cdot \phi_0) = P \cdot (0) - Q \cdot (0) = 0 - 0 = 0$$

We can also see this fact by using the Jacobi identity [twice!] in the following manner.

$$[ad[p,q], \phi_0] = [[ad(p), ad(q)], \phi_0] = [[ad(p), \phi_0], ad(q)] + [ad(q), [ad(p), \phi_0]] =$$
$$[0, ad(q)] + [ad(q), 0] = 0$$

Thus we have found a subalgebra $\hat{l}$ of $\hat{g}$ of the kind that we have been seeking! Here is the proof that this is so. First we show that $\hat{l} \cap \hat{r} = 0$. We can conclude this from our analysis above. However making the argument explicit, we suppose $a \neq 0$ is in $\hat{l} \cap \hat{r}$. Since a is in $\hat{l}$, we know that $[ad(a), \phi_0] = 0$. This says that $ad(a)(\phi_0) = \phi_0(ad(a))$. Now suppose $a \neq 0$ is also in $\hat{r}$. Then for any x in $\hat{g}$, we know that $(\phi_0)(x)$ is in $\hat{r}$, and thus $ad(a)((\phi_0)(x)) = 0$. Thus $\phi_0(ad(a)) = 0$. Now for all x in $\hat{g}$

$$\phi_0(ad(a)(x)) = \phi_0([a,x]) = I_{\hat{r}}([a,x]) = [a,x]$$

since $[a,x]$ is in $\hat{r}$. And thus we arrive at the point where we are asking when $[a,x] = 0$, for a in $\hat{r}$ and for all x is in $\hat{g}$. But this says that a is then in the center of $\hat{g}$. But as we have said above, the center of $\hat{g}$ must be 0. Thus we can conclude that $a = 0$, and we have our conclusion that $\hat{l} \cap \hat{r} = 0$.

We next show that $\hat{g} = \hat{l} + \hat{r}$. Let x be any element in $\hat{g}$. We know that $X \cdot \phi_0$ is in A. Thus there is a a in $\hat{r}$ such that $ad(a) = X \cdot \phi_0$. But we also know that this same element $ad(a)$ comes from a in $\hat{r}$ by this same construction, i.e., $ad(a) = -A \cdot \phi_0$. Thus we have $X \cdot \phi_0 + A \cdot \phi_0 = 0$ or $(X + A) \cdot \phi_0 = [ad(x+a), \phi_0] = 0$. This says that $x + a$ is in $\hat{l}$. Now for each x in $\hat{g}$ we have $x = (x+a) - a$ and thus we have $\hat{g} = \hat{l} + \hat{r}$, which fact gives us our conclusion that $\hat{g} = \hat{l} \oplus \hat{r}$.

We remark immediately that the subalgebra $\hat{l}$ is semisimple. Using 2.16.2, we know that $rad(\hat{l}) = \hat{l} \cap \hat{r}$. But we have just shown that $\hat{l} \cap \hat{r} = 0$. Thus $rad(\hat{l}) = 0$ and this makes $\hat{l}$ semisimple.

To complete this proof, we need to show that $\hat{g}/\hat{r}$ is isomorphic to $\hat{l}$. The fact that $\hat{g}/\hat{r}$ is isomorphic to $\hat{l}$ in the sense of linear spaces is a direct consequence of an isomorphism theorem of Linear Algebra.

$$\hat{g}/\hat{r} = (\hat{l} + \hat{r})/\hat{r} \cong \hat{l}/(\hat{l} \cap \hat{r}) = \hat{l}/0 = \hat{l}$$

Thus we only need to define a Lie algebra map from $\hat{g}/\hat{r}$ to $\hat{l}$. For x in $\hat{g}$, we let $x = p + s$ for p in $\hat{l}$ and s in $\hat{r}$, and we define the map

$$\begin{gathered} \hat{g}/\hat{r} \longrightarrow \hat{l} \\ x + \hat{r} = p + s + \hat{r} = p + \hat{r} \longmapsto p \end{gathered}$$

Now for $y = q + t$, y in $\hat{g}$, q in $\hat{l}$ and t in $\hat{r}$, we have

$$\begin{gathered} [x + \hat{r}, y + \hat{r}] = [p + s + \hat{r}, q + t + \hat{r}] = \\ [p, q] + [p, t] + [p, \hat{r}] + [s, q] + [s, t] + [s, \hat{r}] + [\hat{r}, q] + [\hat{r}, t] + [\hat{r}, \hat{r}] = \\ [p, q] + \hat{r} + \hat{r} + \hat{r} + 0 + 0 + \hat{r} + 0 + 0 = \\ [p, q] + \hat{r} \longmapsto [p, q] \end{gathered}$$

and thus brackets do go into brackets, and we have a Lie algebra isomorphism between $\hat{g}/\hat{r}$ and $\hat{l}$. And we conclude that $\hat{g} = \hat{l} \oplus \hat{r}$ where $\hat{l}$ is semisimple and isomorphic to $\hat{g}/\hat{r}$ and $\hat{r}$ is the radical of $\hat{g}$.

And this is the Levi Decomposition Theorem!

2.17 Ado's Theorem

Ado's Theorem states that every finite dimensional Lie algebra [over $\mathbb{R}$ or $\mathbb{C}$] is essentially <u>linear</u>, that is, it has a faithful finite dimensional representation in some matrix algebra $\widehat{gl}(V)$ [over $\mathbb{R}$ or $\mathbb{C}$]. This means the following: if ρ is a representation of $\hat{g}$ in V:

$$\begin{aligned} \rho : \hat{g} &\longrightarrow \widehat{gl}(V) \\ c &\longmapsto \rho(c) : V \longrightarrow V \\ & \qquad\qquad\quad x \longmapsto \rho(c)(x) \end{aligned}$$

and $\rho[c, d] = [\rho(c), \rho(d)]$ then $\rho(c)(x)$ being <u>faithful</u> means that $\rho(c) = 0$ implies $c = 0$, i.e., the kernel of $\rho = 0$. [This is just another language for saying that we have a homomorphism from an abstract Lie algebra into the Lie algebra of the set of matrices of some dimension n; and "faithful" means we have a isomorphism.]

Now if we let ρ be the adjoint representation of $\hat{g}$ and let $\hat{z}$ be the center of $\hat{g}$, then we have the following.

$$\begin{aligned} ad : \hat{g} &\longrightarrow \widehat{gl}(\hat{g}) \\ c &\longmapsto ad(c) : \hat{g} \longrightarrow \hat{g} \\ &\qquad d \longmapsto ad(c)(d) = [c, d] \end{aligned}$$

[Recall that the adjoint representation is a representation since we know from the Jacobi identity that $ad[c, d] = [ad(c), ad(d)]$.] Now for all a in $\hat{z}$, $[a, d] = 0$ for all d in $\hat{g}$. This then says that $ad(a)$ is the zero map, and thus the center of the adjoint representation is in the kernel of ad. However if $ad(a)(d) = 0$ for all d in $\hat{g}$, then a is in the center of $\hat{g}$, and thus the center is the kernel of the adjoint representation. Therefore, when the center of a Lie algebra is zero, the adjoint representation is faithful, and we have Ado's Theorem fulfilled.

Now suppose the center $\hat{z} \neq 0$. Then we seek a linear space W and a representation τ of $\hat{g}$ in W which is faithful when restricted to the center $\hat{z}$, that is, we want

$$\begin{aligned} \tau : \hat{g} &\longrightarrow \widehat{gl}(W) \\ c &\longmapsto \tau(c) : W \longrightarrow W \\ &\qquad x \longmapsto \tau(c)(x) \end{aligned}$$

and for a in the center $\hat{z}$, if $\tau(a)(x) = 0$ for x in W, then $a = 0$.

Now if we combine these two representations.

$$\begin{gathered} \rho = ad \oplus \tau : \hat{g} \longrightarrow \rho(\hat{g}) = ad(\hat{g}) \oplus \tau(\hat{g}) \subseteq \widehat{gl}(\hat{g}) \oplus \widehat{gl}(W) \\ x \longmapsto \rho(x) = (ad \oplus \tau)(x) = \\ ad(x) \oplus \tau(x) : (y, w) \longmapsto ad(x)(y) + \tau(x)(w) \end{gathered}$$

where $\widehat{gl}(\hat{g}) \oplus \widehat{gl}(W)$ consists of matrices of the form

$$\begin{bmatrix} \widehat{gl}(\hat{g}) & 0 \\ 0 & \widehat{gl}(W) \end{bmatrix}$$

and then we restrict τ to the center $\hat{z}$ of $\hat{g}$, then for a in the center $\hat{z}$ we have

$$\begin{bmatrix} ad(a) & 0 \\ 0 & \tau(a) \end{bmatrix} \begin{bmatrix} y \\ w \end{bmatrix} = \begin{bmatrix} ad(a)(y) \\ \tau(a)(w) \end{bmatrix}$$

and we ask for all y in $\hat{g}$ and w in W such that

$$\begin{bmatrix} ad(a)(y) \\ \tau(a)(w) \end{bmatrix} = 0$$

Since τ is faithful on the center $\hat{z}$, for a in $\hat{z}$, we have $\tau(a) = 0$ only if $a = 0$. And if $ad(0) = 0$, then we get the 0 vector only when the matrix

$$\begin{bmatrix} ad(a) & 0 \\ 0 & \tau(a) \end{bmatrix} = 0$$

Now we restrict ρ to the set $\hat{g} \setminus \hat{z}$. Thus for x in $\hat{g} \setminus \hat{z}$ we have

$$\begin{bmatrix} ad(x) & 0 \\ 0 & \tau(x) \end{bmatrix} \begin{bmatrix} y \\ w \end{bmatrix} = \begin{bmatrix} ad(x)(y) \\ \tau(x)(w) \end{bmatrix}$$

and we ask for all y in $\hat{g}$ and w in W when

$$\begin{bmatrix} ad(x)(y) \\ \tau(x)(w) \end{bmatrix} = 0$$

Since ad is faithful on the set $\hat{g} \setminus \hat{z}$, for x in $\hat{g} \setminus \hat{z}$ we have $ad(x) = 0$ only if $x = 0$. And since $\tau(0) = 0$, this says we get the 0 vector only when the matrix

$$\begin{bmatrix} ad(x) & 0 \\ 0 & \tau(x) \end{bmatrix} = 0$$

and in both of these cases the matrix gives a faithful representation. Thus, if τ and W exist, we can conclude that every finite dimensional Lie algebra $\hat{g}$ over $\mathbb{C}$ has a *faithful* finite dimensional representation, which is what what Ado's Theorem states.

And thus we are on the hunt for W and the representation τ of $\hat{g}$ in W which is faithful on $\hat{z}$ and from this point on we will be working with *complex* Lie algebras. Now if the center $\hat{z}$ is one-dimensional, we can produce a faithful representation of $\hat{z}$ in the following manner. The center $\hat{z}$, being one-dimensional, means that it is isomorphic to the scalars $\mathbb{C}$ which can be considered a complex Lie algebra in the following manner. The field structure of $\mathbb{C}$ immediately gives us that $\mathbb{C}$ is a linear space over $\mathbb{C}$ and it also has the structure of an associative algebra over $\mathbb{C}$. Now every associative algebra becomes a Lie algebra if we define the bracket $[u, v] = u \cdot v - v \cdot u$. If u and v are in $\mathbb{C}$, then $uv - vu = 0$ since multiplication in $\mathbb{C}$ is commutative. Thus $\mathbb{C}$ as a Lie algebra is an abelian Lie algebra. And the center $\hat{z}$ of $\mathbb{C}$ is $\mathbb{C}$ itself.

We now can build a representation of $\mathbb{C}$ in the 2x2 nilpotent matrices over $\mathbb{C}$. These matrices are of the form

$$\begin{bmatrix} 0 & u \\ 0 & 0 \end{bmatrix}$$

They comprise a nilpotent Lie subalgebra of the Lie algebra of all 2x2 matrices over $\mathbb{C}$. It is evident that the addition of two nilpotent matrices and a scalar times a nilpotent matrix give nilpotent matrices, and thus the nilpotent matrices form a linear subspace of $\widehat{gl}(2,\mathbb{C})$. Also by multiplying two nilpotent matrices together, we see that the nilpotent matrices are a subalgebra of the matrix algebra $\widehat{gl}(2,\mathbb{C})$, and, in fact, the product of any two of these nilpotent matrices is always the zero matrix.

$$\begin{bmatrix} 0 & u_1 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & u_2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

But this also says that these nilpotent matrices are a Lie subalgebra of $\widehat{gl}(2,\mathbb{C})$, since the Lie bracket of these nilpotent matrices is always the zero matrix. And thus these nilpotent matrices form a commutative Lie subalgebra.

Now we have a faithful representation σ of $\hat{z} = \mathbb{C}$ in the nilpotent matrices of $\widehat{gl}(2,\mathbb{C})$, as can be seen here.

$$\begin{aligned} \sigma : \mathbb{C} &\longrightarrow \left\{ \begin{bmatrix} 0 & u \\ 0 & 0 \end{bmatrix} \right\} \\ u &\longmapsto \sigma(u) = \begin{bmatrix} 0 & u \\ 0 & 0 \end{bmatrix} : \mathbb{C}^2 \longrightarrow \mathbb{C}^2 \\ &\qquad\qquad \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \longmapsto \begin{bmatrix} 0 & u \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} u \cdot v_2 \\ 0 \end{bmatrix} \end{aligned}$$

We do have a representation. Obviously σ is linear. Now if u, v are in $\hat{g}$, then $[u, v] = 0$; and $[\sigma(u), \sigma(v)] = 0$ since the bracket of any two nilpotent matrices in $\widehat{gl}(2,\mathbb{C})$ is also equal to 0. Thus we do have a representation. It is also faithful. Since the center $\hat{z}$ of $\mathbb{C}$ is $\mathbb{C}$ itself, we take any u in $\mathbb{C}$. Now suppose $\sigma(u) = 0$:

$$\sigma(u) = \begin{bmatrix} 0 & u \\ 0 & 0 \end{bmatrix} = 0$$

We see immediately that implies that $u = 0$, and thus we have a faithful representation of $\hat{z}$ when $\hat{z}$ is one-dimensional. In fact this representation is a nilpotent representation in the sense that $\sigma(u)$ is a <u>nilpotent matrix</u> for all u in $\hat{z}$.

[Side comment: as noted several times already, we will call a representation ρ of a Lie algebra $\hat{g}$ a *nilrepresentation* if $\rho(x)$ is a nilpotent matrix for each x in the nilradical $\hat{n}$ of $\hat{g}$.]

Now if the center $\hat{z}$, which is abelian, has dimension $k > 1$, we can produce a faithful representation by taking the direct sum of k copies of the above one-dimensional faithful representation. We will lay out the case for $k = 2$ and it will be obvious how one could proceed for larger dimensions.

Let us then consider the case when the center $\hat{z}$ of $\hat{g}$, which is abelian, has dimension $k = 2$. Choosing a basis $\{c_1, c_2\}$ for $\hat{z}$, we write for any a in $\hat{z}$

$$a = u_1 c_1 + u_2 c_2$$

for some u_1 and u_2 in $\mathbb{C}$.

We now examine the following matrices over $\mathbb{C}$.

$$\begin{bmatrix} 0 & u_1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & u_2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

We see immediately that these matrices are nilpotent since

$$\begin{bmatrix} 0 & u_1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & u_2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & u_1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & u_2 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

They form a Lie subalgbra of the Lie algebra of all 4x4 matrices over $\mathbb{C}$. It is evident that the addition of two nilpotent matrices of this type and a scalar times a nilpotent matrix of this type again give nilpotent matrices of this type, and thus the nilpotent matrices of this type form a linear subspace of $\widehat{gl}(4,\mathbb{C})$. Also we see that the nilpotent matrices are a subalgebra of the matrix algebra $\widehat{gl}(4,\mathbb{C})$ since the product of two nilpotent matrices is always the zero matrix. And thus in fact these nilpotent matrices form a commutative Lie subalgebra.

Now we have a faithful representation σ of $\hat{z}$ in the nilpotent matrices of $\widehat{gl}(4,\mathbb{C})$.

$$\sigma : \hat{z} \longrightarrow \left\{ \begin{bmatrix} 0 & u_1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & u_2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right\}$$

$$a = u_1 c_1 + u_2 c_2 \longmapsto \begin{bmatrix} 0 & u_1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & u_2 \\ 0 & 0 & 0 & 0 \end{bmatrix} : \mathbb{C}^4 \longrightarrow \mathbb{C}^4$$

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} \longmapsto \begin{bmatrix} 0 & u_1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & u_2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = \begin{bmatrix} u_1 \cdot v_2 \\ 0 \\ u_2 \cdot v_4 \\ 0 \end{bmatrix}$$

We do have a representation. Obviously σ is linear. Now if a_1 and a_2 are in $\hat{z}$, then $[a_1, a_2] = 0$ since $\hat{g}$ is abelian; and $[\sigma(a_1), \sigma(a_2)] = 0$ since the bracket of any two nilpotent matrices in $\widehat{gl}(4,\mathbb{C})$ is also equal to 0. Thus we do have a representation. It is also faithful. Since the center is $\hat{z}$, we take any $a = u_1c_1 + u_2c_2$ in $\hat{z}$. Now suppose $\sigma(a) = 0$:

$$\sigma(a) = \begin{bmatrix} 0 & u_1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & u_2 \\ 0 & 0 & 0 & 0 \end{bmatrix} = 0$$

We see immediately that this implies that u_1 and u_2 are equal to 0, and thus $a = 0$ and we have a faithful representation of $\hat{z}$ when $\hat{z}$ is two-dimensional. In fact this representation is a <u>nilrepresentation</u> since these matrices are nilpotent.

Thus we see that we can have a <u>faithful</u> representation of the center for any dimension k in the nilpotent matrices of $\widehat{gl}(2k,\mathbb{C})$. Now we are seeking a subalgebra $\hat{k}$ of $\hat{g}$ such that $\hat{g} = \hat{z} \oplus \hat{k}$. Obviously if $\hat{z}$ is the radical $\hat{r}$ of $\hat{g}$, then the Levi Decomposition Theorem, which identifies a semisimple Lie subalgebra $\hat{k}$, says that $\hat{z} \oplus \hat{k} = \hat{g}$. And now we have a faithful representation on $\hat{z}$ [the representation σ we constructed above] and a faithful representation on $\hat{k}$ [the adjoint representation, which is always faithful on a semisimple Lie algebra] and these facts give Ado's theorem in this case.

But suppose that $\hat{z}$ is not the radical. Of course, $\hat{z}$ is always contained in the radical. [Since the center is an abelian ideal, it is nilpotent, and thus solvable and this implies that it is contained in a maximal solvable ideal, the radical.] Our task now is to get a representation τ of the radical $\hat{r}$ which preserves the above mentioned faithful representation of the center. Once we have accomplished this, we can reason as follows. We complement $\hat{r}$ with a semisimple Lie subalgebra $\hat{k}$ [by Levi's Theorem], so that $\hat{g} = \hat{r} \oplus \hat{k}$. We now define a representation ρ on $\hat{g}$ such that ρ restricted to the radical is the representation τ [which is faithful on $\hat{z}$], and restricted to $\hat{k}$ is the adjoint representation. Since the adjoint representation is faithful on $\hat{k}$, we can conclude that the representation ρ is faithful on $\hat{g}$, which is what Ado's Theorem asserts, namely that every Lie algebra has a faithful representation. In fact we can make a stronger assertion: Every Lie algebra has a faithful nilrepresentation because a *nilrepresentation* is a representation ρ in which for every

x in $\hat{g}$ which is in the nilradical $\hat{n}$ of $\hat{g}$ [the maximal nilpotent ideal $\hat{n}$ in $\hat{g}$], $\rho(x)$ is a nilpotent endomorphism.

Fortunately every finite dimensional Lie algebra does possess two subalgebras other then the radical — the maximal solvable subalgebra and the maximal nilpotent subalgebra, i.e., the nilradical $\hat{n}$ (see 2.5.2). And for each of these special subalgebras we can construct a complete flag. [See Appendix to 2.17 for details.] [Recall that a flag for a linear space V is a sequence of subspaces

$$0 = V_0 \subset\neq V_1 \subset\neq ... \subset\neq V_l \subset\neq V$$

and that a complete flag is a flag where the subspaces grow one dimension in each step.] Since the nilradical $\hat{n}$ is a nilpotent Lie algebra, it determines a nilpotent complete flag

$$\begin{gathered} \hat{n} = \hat{a}_n \supset \hat{a}_{n-1} \supset \hat{a}_{n-2} \supset ... \supset \hat{a}_1 \supset \hat{a}_0 = \hat{z} \\ [\hat{a}_k, \hat{a}_k] \subset \hat{a}_k, \text{ a nilpotent subalgebra} \\ \hat{a}_{k+1} = \hat{a}_k \oplus \hat{h}_{k+1}, \text{ dim } \hat{h}_{k+1} = 1 \\ [\hat{n}, \hat{a}_{k+1}] \subset \hat{a}_k \subset \hat{a}_{k+1}\text{: thus } \hat{a}_{k+1} \text{ is an ideal in } \hat{n} \end{gathered}$$

Thus, starting with the center $\hat{z}$, which, of course, is nilpotent, we can add nilpotent subalgebras one dimension greater than the previous one until we reach the nilradical $\hat{n}$.

Now since the radical $\hat{r}$ is a solvable Lie algebra, it determines a solvable complete flag

$$\begin{gathered} \hat{r} = \hat{a}_r \supset \hat{a}_{r-1} \supset \hat{a}_{r-2} \supset ... \supset \hat{a}_{n-1} \supset \hat{a}_n = \hat{n} \\ [\hat{a}_l, \hat{a}_l] \subset \hat{a}_l, \text{ a solvable subalgebra} \\ \hat{a}_{l+1} = \hat{a}_l \oplus \hat{h}_{l+1}, \text{ dim } \hat{h}_{l+1} = 1 \\ [\hat{r}, \hat{a}_l] \subset \hat{a}_l\text{: thus } \hat{a}_l \text{ is an ideal in } \hat{r} \end{gathered}$$

Thus, starting with the nilradical $\hat{n}$ (which, of course, is solvable), we can add solvable subalgebras one dimension greater than the previous one until we reach the radical $\hat{r}$. This produces a complete flag of subalgebras, where the center $\hat{z} = \hat{a}_0$ sits at the bottom:

$$\hat{z} = \hat{a}_0 \subset \hat{a}_1 \subset ... \subset \hat{a}_k \subset ... \subset \hat{a}_n = \hat{n} \subset ... \subset \hat{a}_l \subset ... \subset \hat{a}_r = \hat{r}$$

and where the $\hat{a}_k'$'s are nilpotent, for $k <= n$, and the $\hat{a}_l$'s are solvable but not nilpotent. In constructing this complete flag for a Lie algebra $\hat{r}$, we have that each a_i, for $i >= n$, is a subalgebra of $\hat{r}$, and moving from a_i to a_{i+1}, a one-dimensional subspace $\hat{h}_{i+1}$ of $\hat{r}$ complementary to a_i was chosen. This means that $\hat{h}_{i+1}$ is a 1-dimensional Lie subalgebra, and $a_{i+1} = a_i \oplus h_{i+1}$ and indeed a_i is an ideal in $\hat{r}$: $[\hat{r}, \hat{a}_i] \subset \hat{a}_i$. [In the case of a nilpotent complete flag, we have another condition, namely $[\hat{n}, \hat{a}_i] \subset \hat{a}_{i-1} \subset \hat{a}_i$.]

Now we are seeking a representation on the radical $\hat{r}$ which, when restricted to the center $\hat{z} = \hat{a}_0$ of dimension n, is the faithful representation of this center in the nilpotent matrices of $\widehat{gl}(2n,\mathbb{C})$. We do this by moving up our flag one dimension at a time.

Now there is a Proposition which asserts that we can do this in a more general situation:

Proposition:

Let $\hat{g}$ be a Lie algebra which is a direct sum of a solvable ideal $\hat{a}$ and a subalgebra $\hat{h}$. Let μ be a representation of $\hat{a}$. Then there is a representation τ of $\hat{g}$ such that

$$\hat{a} \cap ker(\tau) \subset ker(\mu)$$

It also states that if the nilradical $\hat{n}$ of $\hat{g}$ is the nilradical of $\hat{a}$, or if the nilradical of $\hat{g}$ is itself $\hat{g}$, then τ may be taken to be a nilrepresentation.

[I might remark that the last mentioned condition — $\hat{a} \cap ker(\tau) \subset ker(\mu)$ — insures that the representation τ remains faithful on the center $\hat{z}$ as we move up the flag.]

However at this point in our exposition we will not prove this Proposition in its generality. Rather we circumvent this proof by explicitly defining a representation τ_i on a_i as we move up the flag such that, when restricted to the center $\hat{z}$, this representation is faithful. And we will show that the condition

$$\hat{a}_{i-1} \cap ker(\tau_i) \subset ker(\tau_{i-1})$$

is indeed satisfied at each step.

Thus, starting from the center $\hat{z} = \hat{a}_0$, we have the following:

For $n = 1$, we have $\hat{a}_1 = \hat{z} \oplus \hat{h}_1$. Now $\hat{z}$ is a solvable ideal in $\hat{a}_1$ [since $\hat{z}$ is the center of $\hat{g}$] and $\hat{h}_1$ is a one-dimensional subalgebra of $\hat{n}$ not contained in $\hat{z}$. We know that σ is the faithful representation of the center $\hat{z}$ in the nilpotent matrices of $\widehat{gl}(2k,\mathbb{C})$ as constructed above. We define τ_0 to be this representation σ and define τ_1 to be a representation of $\hat{a}_1$

$$\tau_1 : \hat{a}_1 = \hat{z} \oplus \hat{h}_1 \longrightarrow \widehat{gl}(2k,\mathbb{C})$$

which on the center $\hat{z}$ is the representation σ found above, and on $\hat{h}_1$ is 0. Thus $ker(\tau_1)$ is $\hat{h}_1$, and since σ is faithful, its kernel is 0 and $\tau_1(\hat{h}_1) = 0$. Thus we observe that the condition

$$\hat{a}_0 \cap ker(\tau_1) \subset ker(\tau_0)$$

is satisfied and $\hat{a}_0 \cap ker(\tau_1) = \hat{z} \cap ker(\sigma) = 0$ and $ker(\tau_0) = ker(\sigma) = 0$. Finally, since $\hat{a}_1$ is nilpotent, its nilradical is itself and the image of τ_1 is the image of τ_0 which is the image of σ, which is in the nilpotent matrices of $\widehat{gl}(2k,\mathbb{C})$, and thus we have a nilrepresentation.

What we do, then, is move up the nilpotent complete flag one dimension at a time from the center $\hat{z} = \hat{a}_0$ to the nilradical $\hat{n} = \hat{a}_n$, and then up the solvable complete flag to the radical $\hat{r} = \hat{a}_r$, and at each step we define a representation τ_i of the Lie subalgebra $\hat{a}_i$ which is faithful on the center $\hat{z}$, and whose image is a set of nilpotent matrices in $\widehat{gl}(2k,\mathbb{C})$. And we remark that indeed all these representations satisfy the condition:

$$\hat{a}_{i-1} \cap ker(\tau_i) \subset ker(\tau_{i-1})$$

Thus, in the end, we will have a nilrepresentation of the radical $\hat{r}$ which is faithful on the center $\hat{z}$.

The following calculation is informative. We now have a nilpotent Lie algebra $\hat{a}_1$ in our complete nilpotent flag of the nilradical $\hat{n}$. We calculate

$$[\hat{a}_1, \hat{a}_1] = [\hat{a}_0 \oplus \hat{h}_1, \hat{a}_0 \oplus \hat{h}_1] = [\hat{a}_0, \hat{a}_0] + [\hat{a}_0, \hat{h}_1] + [\hat{h}_1, \hat{a}_0] + [\hat{h}_1, \hat{h}_1]$$

Now knowing that $\hat{a}_0 = \hat{z}$, we have $[\hat{a}_0, \hat{a}_0] = 0$. Now $[\hat{a}_0, \hat{h}_1]$ and $[\hat{h}_1, \hat{a}_0]$ are the same subspace [since in any Lie algebra, $[x, y] = -[y, x]$], and since $\hat{a}_0$ is the center of $\hat{g}$, this gives $[\hat{a}_0, \hat{h}_1] = 0$. And finally, since $\hat{h}_1$ is one-dimensional, $[\hat{h}_1, \hat{h}_1] = 0$. Thus $[\hat{a}_1, \hat{a}_1] = 0$, and $\hat{a}_1$ is abelian but it is not in the center of $\hat{g}$. We can, however, conclude that $\hat{a}_1$ is a nilpotent Lie subalgebra of $\hat{g}$. Also, we note that $[\hat{a}_1, \hat{a}_0] = 0 \subset \hat{a}_0$, and thus $\hat{a}_0$ is an ideal of $\hat{a}_1$ which fact is a consequence of being part of the complete nilpotent flag.

Thus we move up our nilpotent flag one dimension: $\hat{a}_2 = \hat{a}_1 \oplus \hat{h}_2$, where h_2 is a one-dimensional linear space in $\hat{n}$ complementary to $\hat{a}_1$.

From $\hat{a}_1$ to $\hat{a}_2$:

For $n = 2$, we have $\hat{a}_2 = \hat{a}_1 \oplus \hat{h}_2$. Since $\hat{a}_1$ is an ideal in $\hat{n}$, we have $[\hat{n}, \hat{a}_1] \subset \hat{a}_1$, and thus $[\hat{a}_2, \hat{a}_1] \subset \hat{a}_1$, giving us $\hat{a}_1$ as an ideal in $\hat{a}_2$. Now we define τ_2 to be a representation of $\hat{a}_2$

$$\tau_2 : \hat{a}_2 = \hat{a}_1 \oplus \hat{h}_2 = \hat{z} \oplus \hat{h}_1 \oplus \hat{h}_2 \longrightarrow \widehat{gl}(2k,\mathbb{C})$$

which on $\hat{a}_1$ is the representation τ_1 found above, and on $\hat{h}_2$ is 0. Thus $ker(\tau_2)$ is $\hat{h}_1 \oplus \hat{h}_2$: since σ is faithful and $\tau_2(\hat{h}_1 \oplus \hat{h}_2) = 0$. Thus we observe that the condition

$$\hat{a}_1 \cap ker(\tau_2) \subset ker(\tau_1)$$

is satisfied since $\hat{a}_1 \cap ker(\tau_2) = (\hat{z} \oplus \hat{h}_1) \cap (\hat{h}_1 \oplus \hat{h}_2) = \hat{h}_1$ and $ker(\tau_1) = \hat{h}_1$. Finally since $\hat{a}_2$ is nilpotent, its nilradical is itself and the image of τ_2 is the image of τ_1 which is the image of σ, which is in the set of nilpotent matrices of $\widehat{gl}(2k,\mathbb{C})$, and thus we have a nilrepresentation.

We again make the following calculation. We now have a nilpotent Lie algebra $\hat{a}_2$ in our complete nilpotent flag of the nilradical $\hat{n}$. We calculate

$$[\hat{a}_2, \hat{a}_2] = [\hat{a}_1 \oplus \hat{h}_2, \hat{a}_1 \oplus \hat{h}_2] = [\hat{a}_1, \hat{a}_1] + [\hat{a}_1, \hat{h}_2] + [\hat{h}_2, \hat{a}_1] + [\hat{h}_2, \hat{h}_2]$$

Now knowing that $\hat{a}_1$ is abelian, we have $[\hat{a}_1, \hat{a}_1] = 0$. Also $[\hat{a}_1, \hat{h}_2]$ and $[\hat{h}_2, \hat{a}_1]$ are the same subspace. Finally, $\hat{h}_2$ is one- dimensional and thus $[\hat{h}_2, \hat{h}_2] = 0$. Thus we have $[\hat{a}_2, \hat{a}_2] = [\hat{a}_1, \hat{h}_2]$. Also we note that $[\hat{a}_2, \hat{a}_1] \subset \hat{a}_0 \subset \hat{a}_1$, because we have the properties of a nilpotent complete flag and a_1 is an ideal in a_2.

Let us now move up one more step in our complete nilpotent flag, if possible. We have $\hat{a}_3 = \hat{a}_2 \oplus \hat{h}_3$, where $\hat{h}_3$ is a one-dimensional subspace complementary to $\hat{n}$.

From $\hat{a}_2$ to $\hat{a}_3$:

For $n = 3$, we have $\hat{a}_3 = \hat{a}_2 \oplus \hat{h}_3$. Since $\hat{a}_2$ is an ideal in $\hat{n}$, we have $[\hat{n}, \hat{a}_2] \subset \hat{a}_2$, and thus $[\hat{a}_3, \hat{a}_2] \subset \hat{a}_2$, giving us $\hat{a}_2$ an ideal in $\hat{a}_3$. Now we define τ_3 to be a representation of $\hat{a}_3$

$$\tau_3 : \hat{a}_3 = \hat{a}_2 \oplus \hat{h}_3 = \hat{z} \oplus \hat{h}_1 \oplus \hat{h}_2 \oplus \hat{h}_3 \longrightarrow \widehat{gl}(2k,\mathbb{C})$$

which on $\hat{a}_2$ is the representation τ_2 found above, and on $\hat{h}_3$ is 0. Thus $ker(\tau_3)$ is $\hat{h}_1 \oplus \hat{h}_2 \oplus \hat{h}_3$: since σ is faithful — thus its kernel is 0 — and $\tau_3(\hat{h}_1 \oplus \hat{h}_2 \oplus \hat{h}_3) = 0$. Thus we observe that the condition

$$\hat{a}_2 \cap ker(\tau_3) \subset ker(\tau_2)$$

is satisfied, since $\hat{a}_2 \cap ker(\tau_3) = (\hat{z} \oplus \hat{h}_1 \oplus \hat{h}_2) \cap (\hat{h}_1 \oplus \hat{h}_2 \oplus \hat{h}_3) = \hat{h}_1 \oplus \hat{h}_2$ and $ker(\tau_2) = \hat{h}_1 \oplus \hat{h}_2$. Finally, since $\hat{a}_3$ is nilpotent, its nilradical is itself; and the image of τ_3 is the image of τ_2 which is the image of σ, which is in the set of nilpotent matrices of $\widehat{gl}(2k,\mathbb{C})$, and thus we have a nilrepresentation.

We again make the following calculation. We now have a nilpotent Lie algebra $\hat{a}_3$ in our complete nilpotent flag of the nilradical $\hat{n}$. We calculate

$$[\hat{a}_3, \hat{a}_3] = [\hat{a}_2 \oplus \hat{h}_3, \hat{a}_2 \oplus \hat{h}_3] = [\hat{a}_2, \hat{a}_2] + [\hat{a}_2, \hat{h}_3] + [\hat{h}_3, \hat{a}_2] + [\hat{h}_3, \hat{h}_3]$$

Now $[\hat{a}_2, \hat{h}_3]$ and $[\hat{h}_3, \hat{a}_2]$ are the same subspace. Also, $\hat{h}_3$ is one-dimensional and thus $[\hat{h}_3, \hat{h}_3] = 0$. Thus $[\hat{a}_3, \hat{a}_3] = [\hat{a}_2, \hat{a}_2] + [\hat{a}_2, \hat{h}_3] = [\hat{a}_3, \hat{a}_2]$. Also we remark that $[\hat{a}_3, \hat{a}_2] \subset \hat{a}_1 \subset \hat{a}_2$, by the properties of a nilpotent complete flag, and also a_2 is an ideal in a_3.

At this point let us recall the flag that we are building:

$$\hat{z} = \hat{a}_0 \subset \hat{a}_1 \subset ... \subset \hat{a}_k \subset ... \subset \hat{a}_n = \hat{n} \subset ... \subset \hat{a}_l \subset ... \subset \hat{a}_r = \hat{r}$$

where $\hat{n}$ is the nilradical and $\hat{r}$ is the radical of our algebra $\hat{g}$. Continuing as above, we will eventually reach the nilradical $\hat{n}$ (since our Lie algebra is finite-dimensional).

Let us now explore what happens when we begin to add on the solvable Lie subalgebras. Thus we are now at this point in the flag:

$$... \subset \hat{a}_{n-1} \subset \hat{a}_n = \hat{n} \subset \hat{a}_{n+1} \subset ...$$

and we have $\hat{a}_{n+1} = \hat{a}_n \oplus \hat{h}_{n+1}$, where $\hat{h}_{n+1}$ is a one-dimensional subalgebra of the radical $\hat{r}$ not contained in $\hat{n}$. Now we are seeking a representation τ_{n+1} of $\hat{a}_{n+1}$ which is faithful on $\hat{a}_0$. Using the properties of our solvable complete flag, we know that $\hat{a}_n$ is a nilpotent and thus solvable, ideal in $\hat{a}_{n+1}$, and $\hat{h}_{n+1}$, being one-dimensional, is a Lie subalgebra of $\hat{a}_{n+1}$.

From $\hat{a}_n$ to $\hat{a}_{n+1}$:

We have $\hat{a}_{n+1} = \hat{a}_n \oplus \hat{h}_{n+1}$. Since $\hat{a}_n$ is an ideal in $\hat{r}$, we have $[\hat{r}, \hat{a}_n] \subset \hat{a}_n$, and thus $[\hat{a}_{n+1}, \hat{a}_n] \subset \hat{a}_n$, giving us that $\hat{a}_n$ is an ideal in $\hat{a}_{n+1}$. Now define τ_{n+1} to be a representation of $\hat{a}_{n+1}$

$$\tau_{n+1} : \hat{a}_{n+1} = \hat{a}_n \oplus \hat{h}_{n+1} = \hat{z} \oplus \hat{h}_1 \oplus \hat{h}_2 \oplus ... \oplus \hat{h}_{n-2} \oplus \hat{h}_{n-1} \oplus \hat{h}_n \oplus \hat{h}_{n+1} \longrightarrow \widehat{gl}(2k,\mathbb{C})$$

which on $\hat{a}_n$ is the representation τ_n found above, and on $\hat{h}_{n+1}$ is 0. Thus $ker(\tau_{n+1})$ is $\hat{h}_{n+1} \oplus \hat{h}_1 \oplus \hat{h}_2 \oplus ... \oplus \hat{h}_{n-2} \oplus \hat{h}_{n-1} \oplus \hat{h}_n$: since σ is faithful and $\tau_{n+1}(\hat{h}_{n+1}) = 0$. Thus we observe that the condition

$$\hat{a}_n \cap ker(\tau_{n+1}) \subset ker(\tau_n)$$

is satisfied and $\hat{a}_n \cap ker(\tau_{n+1}) = (\hat{z} \oplus \hat{h}_1 \oplus ... \hat{h}_n) \cap (\hat{h}_1 \oplus \hat{h}_2 \oplus ... \oplus \hat{h}_{n+1}) = \hat{h}_1 \oplus ... \oplus \hat{h}_n$; and $ker(\tau_n) = \hat{h}_1 \oplus ... \oplus \hat{h}_n$. Finally, since the nilradical of $\hat{a}_{n+1}$ is $a_n = \hat{n}$, the image of τ_{n+1} restricted to $\hat{a}_n$ is again the image of σ, which is in the nilpotent matrices of $\widehat{gl}(2k,\mathbb{C})$, and thus we have a nilrepresentation.

Thus we see that we can move from the nilpotent subalgebras to the solvable subalgebras and that in fact nothing in the series of calculations changes.

We again make the following calculation. We now have a solvable Lie algebra $\hat{a}_{n+1}$ in our complete solvable flag of the radical $\hat{r}$. We calculate

$$[\hat{a}_{n+1}, \hat{a}_{n+1}] = [\hat{a}_n \oplus \hat{h}_{n+1}, \hat{a}_n \oplus \hat{h}_{n+1}] =$$
$$[\hat{a}_n, \hat{a}_n] + [\hat{a}_n, \hat{h}_{n+1}] + [\hat{h}_{n+1}, \hat{a}_n] + [\hat{h}_{n+1}, \hat{h}_{n+1}]$$

Now $[\hat{a}_n, \hat{h}_{n+1}]$ and $[\hat{h}_{n+1}, \hat{a}_n]$ are the same subspace. Also, $\hat{h}_{n+1}$ is one- dimensional, and thus $[\hat{h}_{n+1}, \hat{h}_{n+1}] = 0$. Thus $[\hat{a}_{n+1}, \hat{a}_{n+1}] = [\hat{a}_n, \hat{a}_n] + [\hat{a}_n, \hat{h}_{n+1}] = [\hat{a}_{n+1}, \hat{a}_n]$. Also we remark that $[\hat{a}_{n+1}, \hat{a}_n] \subset \hat{a}_n$, which is the property of a solvable complete flag, and thus a_n is an ideal in a_{n+1}.

And, finally, we arrive at the radical $\hat{r}$ of our Lie algebra $\hat{g}$ where we have a representation τ_r of $\hat{r}$ which is faithful on the center $\hat{z}$. Now we use Levi's Theorem to get a decomposition of $\hat{g} = \hat{r} \oplus \hat{k}$ for some semisimple subalgebra $\hat{k}$. Finally using any representation τ on $\hat{g}$ which when restricted to the radical $\hat{r}$ is the representation τ_r, we define a representation $\rho = ad \oplus \tau$, which gives us a faithful finite dimensional representation of $\hat{g}$. [We remark that this representation was over the field $\mathbb{C}$.]

And thus we have Ado's Theorem: Every Lie algebra over $\mathbb{C}$ has a faithful finite dimensional nilrepresentation in some matrix algebra $\widehat{gl}(V,\mathbb{C})$.

****

Now in order to get a better grasp of what we are doing let us work through the following example.

Let us examine, therefore, the following set of matrices:

$$A_6 = \left\{ \begin{bmatrix} 0 & u_1 & a_1 & c_1 \\ 0 & 0 & b_1 & 0 \\ 0 & 0 & d_2 & 0 \\ 0 & 0 & 0 & d_1 \end{bmatrix} \right\}$$

First we show that these matrices form a Lie algebra. Clearly we need only compute the bracket product:

$$\begin{bmatrix} 0 & u_1 & a_1 & c_1 \\ 0 & 0 & b_1 & 0 \\ 0 & 0 & d_2 & 0 \\ 0 & 0 & 0 & d_1 \end{bmatrix} \begin{bmatrix} 0 & u_3 & a_3 & c_3 \\ 0 & 0 & b_3 & 0 \\ 0 & 0 & d_4 & 0 \\ 0 & 0 & 0 & d_3 \end{bmatrix} - \begin{bmatrix} 0 & u_3 & a_3 & c_3 \\ 0 & 0 & b_3 & 0 \\ 0 & 0 & d_4 & 0 \\ 0 & 0 & 0 & d_3 \end{bmatrix} \begin{bmatrix} 0 & u_1 & a_1 & c_1 \\ 0 & 0 & b_1 & 0 \\ 0 & 0 & d_2 & 0 \\ 0 & 0 & 0 & d_1 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & 0 & -a_3d_2 + a_1d_4 + b_3u_1 - b_1u_3 & -c_3d_1 + c_1d_3 \\ 0 & 0 & -b_3d_2 + b_1d_4 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus we do have a Lie algebra of 6 dimensions.

Next we show that this algebra is solvable.

$D^1A_6 = [A_6, A_6] =$

$$\left\{ \begin{bmatrix} 0 & 0 & -a_3d_2 + a_1d_4 + b_3u_1 - b_1u_3 & -c_3d_1 + c_1d_3 \\ 0 & 0 & -b_3d_2 + b_1d_4 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right\}$$

$D^2A_6 = [D^1A_6, D^1A_6] =$

$$\begin{bmatrix} 0 & 0 & -a_3d_2 + a_1d_4 + b_3u_1 - b_1u_3 & -c_3d_1 + c_1d_3 \\ 0 & 0 & -b_3d_2 + b_1d_4 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & -a_7d_6 + a_5d_8 + b_7u_5 - b_5u_7 & -c_7d_5 + c_5d_7 \\ 0 & 0 & -b_7d_6 + b_5d_8 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} -$$

$$\begin{bmatrix} 0 & 0 & -a_7d_6 + a_5d_8 + b_7u_5 - b_5u_7 & -c_7d_5 + c_5d_7 \\ 0 & 0 & -b_7d_6 + b_5d_8 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & -a_3d_2 + a_1d_4 + b_3u_1 - b_1u_3 & -c_3d_1 + c_1d_3 \\ 0 & 0 & -b_3d_2 + b_1d_4 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

And thus we do have a solvable Lie algebra.

We now take the brackets of the six basis matrices.

$$U_1 = \begin{bmatrix} 0 & u_1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}; A_1 = \begin{bmatrix} 0 & 0 & a_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}; B_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & b_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$C_1 = \begin{bmatrix} 0 & 0 & 0 & c_1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}; D_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & d_1 \end{bmatrix}; D_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & d_2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$[U_1, A_1] = 0; [U_1, B_1] = \begin{bmatrix} 0 & 0 & u_1b_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}; [U_1, C_1] = 0; [U_1, D_1] = 0; [U_1, D_2] = 0$$

$$[A_1, B_1] = 0; [A_1, C_1] = 0; [A_1, D_1] = 0; [A_1, D_2] = \begin{bmatrix} 0 & 0 & a_1d_2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$[B_1, C_1] = 0; [B_1, D_1] = 0; [B_1, D_2] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & b_1d_2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$[C_1, D_1] = \begin{bmatrix} 0 & 0 & 0 & c_1d_1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}; [C_1, D_2] = 0; [D_1, D_2] = 0$$

From these brackets we see that A_6 does not have a trivial center.

We now want to lay out a solvable complete flag corresponding to the solvable Lie algebra A_6. We have

$$A_6 = \left\{ \begin{bmatrix} 0 & u_1 & a_1 & c_1 \\ 0 & 0 & b_1 & 0 \\ 0 & 0 & d_2 & 0 \\ 0 & 0 & 0 & d_1 \end{bmatrix} \right\} \supset D^1 A_6 = \left\{ \begin{bmatrix} 0 & 0 & a_3 & c_3 \\ 0 & 0 & b_3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right\} \supset D^2 A_6 = 0$$

Thus we see a complete flag is

$$\left\{\begin{bmatrix} 0 & u_1 & a_1 & c_1 \\ 0 & 0 & b_1 & 0 \\ 0 & 0 & d_2 & 0 \\ 0 & 0 & 0 & d_1 \end{bmatrix}\right\} \supset \left\{\begin{bmatrix} 0 & u_1 & a_1 & c_1 \\ 0 & 0 & b_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & d_1 \end{bmatrix}\right\} \supset \left\{\begin{bmatrix} 0 & u_1 & a_1 & c_1 \\ 0 & 0 & b_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right\} \supset$$

$$\left\{\begin{bmatrix} 0 & 0 & a_1 & c_1 \\ 0 & 0 & b_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right\} \supset \left\{\begin{bmatrix} 0 & 0 & a_1 & 0 \\ 0 & 0 & b_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right\} \supset \left\{\begin{bmatrix} 0 & 0 & a_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right\} \supset 0$$

Just examining the matrices we see that the matrices

$$\left\{\begin{bmatrix} 0 & u_1 & a_1 & c_1 \\ 0 & 0 & b_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right\}; \left\{\begin{bmatrix} 0 & 0 & a_1 & c_1 \\ 0 & 0 & b_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right\}$$

$$\left\{\begin{bmatrix} 0 & 0 & a_1 & 0 \\ 0 & 0 & b_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right\}; \left\{\begin{bmatrix} 0 & 0 & a_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right\}$$

are nilpotent, while the matrices

$$\left\{\begin{bmatrix} 0 & u_1 & a_1 & c_1 \\ 0 & 0 & b_1 & 0 \\ 0 & 0 & d_2 & 0 \\ 0 & 0 & 0 & d_1 \end{bmatrix}\right\} \supset \left\{\begin{bmatrix} 0 & u_1 & a_1 & c_1 \\ 0 & 0 & b_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & d_1 \end{bmatrix}\right\}$$

are solvable. Thus we see that the 6-dimensional set of matrices A_6 is its own radical $\hat{r}$, which contains the 4-dimensional nilradical $\hat{n} = A_4$ comprised of the matrices

$$\left\{\begin{bmatrix} 0 & u_1 & a_1 & c_1 \\ 0 & 0 & b_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right\}$$

and that it does not have a zero center.

Now we first climb up the Nilpotent Complete Flag to the nilradical A_4. To do this, we first calculate the lower central series for A_4:

$C^1 A_4 = [A_4, A_4] =$

$$\left[\left\{\begin{bmatrix} 0 & u_1 & a_1 & c_1 \\ 0 & 0 & b_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right\}, \left\{\begin{bmatrix} 0 & u_3 & a_3 & c_3 \\ 0 & 0 & b_3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right\}\right] = \left\{\begin{bmatrix} 0 & 0 & b3u1 - b1u3 & c_1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right\}$$

$C^2A_4 = [A_4, C^1A_4] =$

$$\left[\left\{\begin{bmatrix} 0 & u_1 & a_1 & c_1 \\ 0 & 0 & b_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right\}, \left\{\begin{bmatrix} 0 & 0 & b3u1 - b1u3 & c_1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right\}\right] = 0$$

Thus we choose the first set of matrices in the Complete Nilpotent Flag to be

$$A_1 = \left\{\begin{bmatrix} 0 & 0 & a_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right\}.$$

It is one-dimensional, and thus is an abelian subalgebra of A_4. Indeed from our knowledge of the Complete Nilpotent Flag, we know that A_1 is an ideal in A_4. From the brackets of the basis vectors we have immediately that this bracket $[A_1, A_4] = 0$ and this affirms that the matrices A_1 is an ideal in A_4.

We choose the next set of matrices in the Complete Nilpotent Flag to be

$$A_2 = \left\{\begin{bmatrix} 0 & 0 & a_1 & c_1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right\}$$

which is equal to A_1 plus the 1-dimensional subspace C_1

$$C_1 = \left\{\begin{bmatrix} 0 & 0 & 0 & c_1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right\}$$

We see from the bracket of the basis vectors that

$$[A_2, A_4] = \left\{ \begin{bmatrix} 0 & 0 & -b_1u_3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right\}$$

and thus A_2 is an ideal in A_4, but also we see that $[A_2, A_4] \subset A_1 \subset A_2$, which is as it should be from our knowledge of the Complete Nilpotent Flag. We remark also that $C^1A_4 = A_2$.

We choose the next set of matrices in the Complete Nilpotent Flag to be

$$A_3 = \left\{ \begin{bmatrix} 0 & 0 & a_1 & c_1 \\ 0 & 0 & b_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right\}$$

which is equal to A_2 plus the 1-dimensional subspace B_1, where

$$B_1 = \left\{ \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & b_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right\}$$

We see from the bracket of the basis vectors that

$$[A_3, A_4] = \left\{ \begin{bmatrix} 0 & 0 & -b_1u_3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right\}$$

and thus A_3 is an ideal in A_4, but also we see that $[A_4, A_5] \subset A_2 \subset A_3$, which is as it should be from our knowledge of the Complete Nilpotent Flag.

Now the next set of matrices in the Complete Nilpotent Flag is

$$A_4 = \left\{ \begin{bmatrix} 0 & u_1 & a_1 & c_1 \\ 0 & 0 & b_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right\}$$

which, of course, is the nilradical of A_6. We observe that A_4 is equal to A_3 plus the 1-dimensional subspace U_1

$$U_1 = \left\{ \begin{bmatrix} 0 & u_1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right\},$$

is indeed a subalgebra of our Lie algebra A_6, since it is generated by a basis vector for A_6.

We continue up our flag. We are now in the Solvable Complete flag part of our Lie algebra A_6. Everything repeats except that we no longer have the relation

$$[\hat{n}, \hat{a}_{i+1}] \subset \hat{a}_i \subset \hat{a}_{i+1}$$

(which was a characteristic of nilpotent Lie algebras and not of solvable Lie algebras) but only the relation

$$[\hat{r}, \hat{a}_{i+1}] \subset \hat{a}_{i+1}$$

We choose the next set of matrices in the Complete Solvable Flag to be

$$A_5 = \left\{ \begin{bmatrix} 0 & u_1 & a_1 & c_1 \\ 0 & 0 & b_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & d_1 \end{bmatrix} \right\}$$

which is equal to A_4 plus the 1-dimensional subspace D_1

$$D_1 = \left\{ \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & d_1 \end{bmatrix} \right\}$$

We see from the bracket of the basis vectors that

$$[A_4, A_5] = \left\{ \begin{bmatrix} 0 & 0 & b_2u_1 - b_1u_2 & c_1d_2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right\}$$

and thus A_4 is an ideal in A_5.

Finally, the next set of matrices in the Complete Solvable Flag is the radical $\hat{r} = A_6$

$$A_6 = \left\{ \begin{bmatrix} 0 & u_1 & a_1 & c_1 \\ 0 & 0 & b_1 & 0 \\ 0 & 0 & d_2 & 0 \\ 0 & 0 & 0 & d_1 \end{bmatrix} \right\}$$

which is equal to A_5 plus the 1-dimensional subspace D_2

$$D_2 = \left\{ \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & d_2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right\}$$

Again we see from the bracket of the basis vectors that A_5 is an ideal in A_6 since

$$[A_5, A_6] = \left\{ \begin{bmatrix} 0 & 0 & a_1 d_4 + b_3 u_1 - b_1 u_3 & -c_3 d_1 + c_1 d_3 \\ 0 & 0 & b_1 d_4 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right\}$$

****

****

Appendix for 2.17: A Review of Some Terminology:

Nilpotent Lie Algebra:

We have

$$\begin{gathered} C^0\hat{g} := \hat{g} \\ C^1\hat{g} := [\hat{g}, C^0\hat{g}] = [\hat{g}, \hat{g}] \subset \hat{g} \\ C^2\hat{g} := [\hat{g}, C^1\hat{g}] \subset [\hat{g}, \hat{g}] = C^1\hat{g} \\ C^3\hat{g} := [\hat{g}, C^2\hat{g}] \subset [\hat{g}, C^1\hat{g}] = C^2\hat{g} \\ . \\ . \\ . \\ C^{k-1}\hat{g} := [\hat{g}, C^{k-2}\hat{g}] \subset [\hat{g}, C^{k-3}\hat{g}] = C^{k-2}\hat{g} \\ C^k\hat{g} := [\hat{g}, C^{k-1}\hat{g}] = 0 \subset C^{k-1}\hat{g} \\ C^{k-1}\hat{g} \subset \text{center}(\hat{g}) \neq 0 \\ \text{Each } C^i\hat{g} \text{ is an ideal of } \hat{g} \end{gathered}$$

****

Nilpotent Complete Flag

We have

$$C^0\hat{g} = \hat{g} = \hat{a}_0 \supset \hat{a}_1 \supset \ldots \supset a_s = C^l\hat{g} \supset \ldots \supset \hat{a}_t = C^{l+1}\hat{g} \supset \ldots \supset a_n = C^k\hat{g} = 0$$

where $\dim \hat{g} = n = \dim a_0 = \dim a_1 + 1; \dim a_2 = \dim a_1 + 1; \ldots . \dim a_n = \dim a_{n-1} + 1$.

Also we have $\hat{g} \supset C^1\hat{g} \supset C^2\hat{g} \supset C^3\hat{g} \supset ... \supset C^{k-1}\hat{g} \supset C^k\hat{g} = 0$ where

$$\text{dim } \hat{g} = \text{n} > \text{dim } C^1\hat{g} > \text{dim } C^2\hat{g} > \text{dim } C^3\hat{g} > \ . \ . \ . > \text{dim } C^{k-1}\hat{g} > \text{dim } C^k\hat{g} = 0$$

****

Nilpotent Lie Algebra implies a Nilpotent Complete Flag:

We have

$\hat{g} \supset C^1\hat{g} \supset C^2\hat{g} \supset C^3\hat{g} \supset ... \supset C^{k-1}\hat{g} \supset C^k\hat{g} = 0$ where

$$\text{dim } \hat{g} = \text{n} > \text{dim } C^1\hat{g} > \text{dim } C^2\hat{g} > \text{dim } C^3\hat{g} > \ . \ . \ .> \text{dim } C^{k-1}\hat{g} > \text{dim } C^k\hat{g} = 0$$

Now having chosen $\hat{a}_0 = \hat{g} = C^0\hat{g}$, we then choose $\hat{a}_0 = \hat{a}_1 \oplus \hat{h}_0$ where $\hat{a}_1 \supset C^1\hat{g}$ is of dimension $n-1$ and $\hat{h}_0$ is an arbitrary complementary subspace of dimension = 1.

If $\hat{a}_1 \neq C^1\hat{g}$, we continue with $\hat{a}_1 = \hat{a}_2 \oplus \hat{h}_1$, where $\hat{a}_2 \subset C^1\hat{g}$ and is of dimension $n-2$ and $\hat{h}_1$ is an arbitrary complementary subspace of $\hat{a}_1$ and is of dimension = 1.

Continuing in this manner we obtain a complete flag

$$\hat{g} = \hat{a}_0 \supset \hat{a}_1 \supset \hat{a}_2 \supset \ldots \supset \hat{a}_{n-1} \supset \hat{a}_n = 0.$$

where each $C^i\hat{g}$ appears for some $\hat{a}_j$.

Thus we have

$$C^0\hat{g} = \hat{a}_0 \supset \hat{a}_1 \supset \ldots \supset a_s = C^l\hat{g} \supset \ldots \supset \hat{a}_t = C^{l+1}\hat{g} \supset \ldots \supset a_n = 0$$

Now choose an $\hat{a}_i$ such that

$$\ldots \supset \hat{a}_s = C^l\hat{g} \supseteq \ldots \supseteq \hat{a}_i \supset \hat{a}_{i+1} \supseteq \ \ldots \supseteq \hat{a}_t = C^{l+1}\hat{g} \supset \ldots$$

Then we have

$$[\hat{g}, \hat{a}_i] = [\hat{a}_0, \hat{a}_i] \subseteq [C^0\hat{g}, C^l\hat{g}] = C^{l+1}\hat{g} = \hat{a}_t \subseteq \hat{a}_{i+1} \subset \hat{a}_i$$

and thus $\hat{a}_i$ is an ideal in $\hat{g}$; and indeed this means $[\hat{a}_i, \hat{a}_i] \subset \hat{a}_i$, and thus $\hat{a}_i$ is a subalgebra. [We remark that nilpotency of the Lie algebra determines the critical inclusion $\hat{a}_{i+1} \subset \hat{a}_i$; and also this inclusion means

$$[\hat{g}/\hat{a}_{i+1}, \hat{a}_i/\hat{a}_{i+1}] \subseteq \hat{a}_{i+1}/\hat{a}_{i+1} = 0$$

and this relation says that $\hat{a}_i/\hat{a}_{i+1}$ is in the center of $\hat{g}/\hat{a}_{i+1}$.]

****

Nilpotent Complete Flag implies Nilpotent Lie Algebra:

We are given the nilpotent complete flag

$$\hat{g} = \hat{a}_0 \supset \hat{a}_1 \supset \hat{a}_2 \supset ... \supset \hat{a}_{n-1} \supset \hat{a}_n = 0$$

We use induction to prove that $C^i\hat{g} \subset \hat{a}_i$; that is, $C^i = [\hat{g}, C^{i-1}\hat{g}] \subset \hat{a}_i$.

For $i = 0$, we have $C^0\hat{g} = \hat{g} = \hat{a}_0 = \hat{g}$.

Assume that the claim is true for i. This gives $C^i\hat{g} \subset \hat{a}_i$.

We prove that the claim is true for $i + 1$, i.e., $C^{i+1}\hat{g} \subset \hat{a}_{i+1}$. Now since $\hat{a}_i = \hat{a}_{i+1} \oplus \hat{h}_i$, we have

$$\begin{aligned} C^{i+1}\hat{g} = [\hat{g}, C^i\hat{g}] &\subset [\hat{g}, \hat{a}_i] \\ &= [\hat{g}, \hat{a}_{i+1} \oplus \hat{h}_i] \\ &= [\hat{g}, \hat{a}_{i+1}] + [\hat{g}, \hat{h}_i] \\ &\subset \hat{a}_{i+1} + \hat{a}_{i+1} \\ &\subset \hat{a}_{i+1} \end{aligned}$$

since $\hat{a}_{i+1}$ is an ideal in $\hat{g}$ and $[\hat{g}, \hat{h}_i] \subset [\hat{g}, \hat{a}_i] \subset \hat{a}_{i+1}$. Thus for some n, $C^{n-1}\hat{g} \subset \hat{a}_{n-1}$, $C^n\hat{g} \subset \hat{a}_n = 0$ and thus $\hat{g}$ is a nilpotent Lie algebra.

****

Solvable Lie Algebra:

In a solvable Lie Algebra, we have:

$$D^0\hat{g} := \hat{g}$$
$$D^1\hat{g} := [D^0\hat{g}, D^0\hat{g}] = [\hat{g}, \hat{g}]; D^1\hat{g} = C^1\hat{g}$$
$$D^2\hat{g} := [D^1\hat{g}, D^1\hat{g}]$$
$$D^3\hat{g} := [D^2\hat{g}, D^2\hat{g}]$$
$$.$$
$$.$$
$$.$$
$$D^{k-1}\hat{g} := [D^{k-2}\hat{g}, D^{k-2}\hat{g}]$$
$$D^k\hat{g} := [D^{k-1}\hat{g}, D^{k-1}\hat{g}] = 0$$

[Each $D^i\hat{g}$ is an ideal of $\hat{g}$ by Jacobi identity]

$D^{k-1}\hat{g} \neq 0$ is an abelian ideal of $\hat{g}$

****

Solvable Complete Flag:

In a solvable complete flag we have:

$$D^0\hat{g} = \hat{g} = \hat{a}_0 \supset \hat{a}_1 \supset \ldots \supset a_s = D^l\hat{g} \supset \ldots \supset \hat{a}_t = D^{l+1}\hat{g} \supset \ldots \supset a_n = D^k\hat{g} = 0$$

where $\dim \hat{g} = n = \dim a_0 = \dim a_1 + 1; \dim a_2 = \dim a_1 + 1; \ldots; \dim a_n = \dim a_{n-1} + 1$.

Also we have $\hat{g} \supset D^1\hat{g} \supset D^2\hat{g} \supset D^3\hat{g} \supset ... \supset D^{k-1}\hat{g} \supset D^k\hat{g} = 0$ where

$$\dim\ \hat{g} = \text{n} > \dim\ D^1\hat{g} > \dim\ D^2\hat{g} > \dim\ D^3\hat{g} > \ldots > \dim\ D^{k-1}\hat{g} > \dim\ D^k\hat{g} = 0$$

Also we a have

$[\hat{a}_i, \hat{a}_i] \subset \hat{a}_i$: thus $\hat{a}_i$ is a subalgebra

$[\hat{a}_i, \hat{a}_{i+1}] \subset \hat{a}_{i+1}$: thus $\hat{a}_{i+1}$ is an ideal in $\hat{a}_i$

****

Solvable Lie Algebra implies Solvable Complete Flag:

Here we have

$$\begin{gathered}
\dim \hat{g} = \mathrm{n} > \\
D^1\hat{g} := [D^0\hat{g}, D^0\hat{g}] = [\hat{g}, \hat{g}]; D^1\hat{g} = C^1\hat{g} \\
D^2\hat{g} := [D^1\hat{g}, D^1\hat{g}] \\
D^3\hat{g} := [D^2\hat{g}, D^2\hat{g}] \\
. \\
. \\
. \\
D^{k-1}\hat{g} := [D^{k-2}\hat{g}, D^{k-2}\hat{g}] \\
D^k\hat{g} := [D^{k-1}\hat{g}, D^{k-1}\hat{g}] = 0 \\
[\text{Each } D^i\hat{g} \text{ is an ideal of } \hat{g} \text{ by the Jacobi identity}] \\
D^{k-1}\hat{g} \neq 0 \text{ is an abelian ideal of } \hat{g}
\end{gathered}$$

$$\dim \hat{g} = \mathrm{n} > \dim D^1\hat{g} > \dim D^2\hat{g} > \dim D^3\hat{g} > \ldots > \dim D^{k-1}\hat{g} > \dim D^k\hat{g} = 0$$

Now having chosen $\hat{a}_0 = \hat{g} = D^0\hat{g}$, we choose $\hat{a}_0 = \hat{a}_1 \oplus \hat{h}_0$ where $\hat{a}_1 \supset D^1\hat{g}$ and is of dimension $n-1$ and $\hat{h}_0$ is an arbitrary complementary subspace of dimension $= 1$.

If $\hat{a}_1 \neq D^1\hat{g}$, we continue with $\hat{a}_1 = \hat{a}_2 \oplus \hat{h}_1$, where $\hat{a}_2 \supset D^1\hat{g}$ and is of dimension $n-2$ and $\hat{h}_1$ is an arbitrary complementary subspace of $\hat{a}_1$ of dimension $= 1$.

Continuing in this manner we obtain a complete flag

$$\hat{g} = \hat{a}_0 \supset \hat{a}_1 \supset \hat{a}_2 \supset \ldots \supset \hat{a}_{n-1} \supset \hat{a}_n = 0.$$

where each $D^i\hat{g}$ appears for some $\hat{a}_j$.

Thus we have

$$D^0\hat{g} = \hat{a}_0 \supset \hat{a}_1 \supset \ldots \supset a_s = D^l\hat{g} \supset \ldots \supset \hat{a}_t = D^{l+1}\hat{g} \supset \ldots \supset a_n = 0$$

Now choose an $\hat{a}_i$ such that

$$\ldots \supset \hat{a}_s = D^l\hat{g} \supseteq \ldots \supseteq \hat{a}_i \supset \hat{a}_{i+1} \supseteq \ \ldots \supseteq \hat{a}_t = D^{l+1}\hat{g} \supset \ldots$$

Then we have

$$[\hat{g}, \hat{a}_i] = [\hat{a}_0, \hat{a}_i] \subseteq [D^0\hat{g}, D^l\hat{g}] = D^{l+1}\hat{g} = \hat{a}_t \subseteq \hat{a}_{i+1} \subset \hat{a}_i$$

and thus $\hat{a}_i$ is an ideal in $\hat{g}$; and indeed this means $[\hat{a}_i, \hat{a}_i] \subset \hat{a}_i$, and thus $\hat{a}_i$ is a subalgebra. [We remark that solvability of the Lie algebra determines the critical inclusion $\hat{a}_{i+1} \subset \hat{a}_i$; and also this inclusion means $[\hat{g}/\hat{a}_{i+1}, \hat{a}_i/\hat{a}_{i+1}] \subseteq \hat{a}_{i+1}/\hat{a}_{i+1} = 0$ and that this says that $\hat{a}_i/\hat{a}_{i+1}$ is in the center of $\hat{g}/\hat{a}_{i+1}$.]

****

Solvable Complete Flag implies Solvable Lie Algebra:

We are given the solvable complete flag

$$\hat{g} = \hat{a}_0 \supset \hat{a}_1 \supset \hat{a}_2 \supset ... \supset \hat{a}_{n-1} \supset \hat{a}_n = 0$$

We use induction to prove that $D^i\hat{g} \subset \hat{a}_j$ for some $j > i$; that is, $D^i\hat{g} = [D^{i-1}\hat{g}, D^{i-1}\hat{g}] \subset \hat{a}_j$.

For $i = 0$, we have $D^0\hat{g} = \hat{g} = \hat{a}_0 = \hat{g}$.

Assume that the claim is true for i. This gives $D^i\hat{g} \subset \hat{a}_j$ for some $j > i$.

We prove that the claim is true for $i+1, i.e., D^{i+1}\hat{g} \subset \hat{a}_k$ for some $k > i+1$. Now since $\hat{a}_j = \hat{a}_{j-1} \oplus \hat{h}_j$, we have

$$\begin{gathered} D^{i+1}\hat{g} = [D^i\hat{g}, D^i\hat{g}] \subset [\hat{a}_j, \hat{a}_j] = [\hat{a}_{j-1} \oplus \hat{h}_j, \hat{a}_{j-1} \oplus \hat{h}_j] \\ = [\hat{a}_{j-1}, \hat{a}_{j-1}] + [\hat{h}_j, \hat{a}_{j-1}] \subset \hat{a}_{j-1} + [\hat{a}_j, \hat{a}_{j-1}] \subset \\ \hat{a}_{j-1} + \hat{a}_j \subset \hat{a}_{j-1} \end{gathered}$$

since $\hat{a}_{j-1}$ is a subalgebra and since h_j is in $\hat{a}_j$ and is an ideal in $\hat{a}_{j-1}$. Thus for some $k-1$, $D^{k-1}\hat{g} \subset \hat{a}_{n-1}$, and for some l, $D^l\hat{g} \subset \hat{a}_n = 0$, and we conclude that $\hat{g}$ is a solvable Lie algebra.

****

Appendices for the Whole Treatise

A.1. Linear Algebra

We begin with the definition of a Linear Algebra or Linear Space. When we talk about a Linear Algebra, we need to talk about two sets: a set V which is an abelian group and a set $\mathbb{F}$ which is a field – and we say that we are treating a Linear Algebra V [the abelian group] over a scalar field $\mathbb{F}$.

A.1.1 The Concept of a Group. We assume that our readers have some acquaintance with the concept of a group. Here, it is the set V which has a binary operation $V \times V \longrightarrow V$ that associates, has a unique identity, and in which every element has a unique inverse. If this operation also commutes, then we have an *abelian* group in which the binary operation is written in additive notation $[(u, v) \longmapsto u + v \in V]$, the identity is given the symbol 0, and the inverse of any element v in V is given the symbol $-v$. Here the group is abelian.

A.1.2 The Concept of a Field. The concept of a *field* $\mathbb{F}$ is rather complicated. A field $\mathbb{F}$ is a set which has two binary operations, called addition, symbolized by $+$, and multiplication, symbolized by just writing the two elements of the binary operation in juxtaposition. The addition operation in $\mathbb{F}$, symbolized again by $+$, is an abelian group, with identity written as 0, and with the inverse of an element c in $\mathbb{F}$ written as $-c$. [There is usually no confusion between these two addition operations and the two identities, one from V and one from $\mathbb{F}$, since one always knows when one is adding elements in V and when one is adding elements in $\mathbb{F}$.] The multiplication operation in $\mathbb{F}$ associates and commutes and has a multiplicative identity symbolized by 1. However this multiplication operation does not make $\mathbb{F}$ into a group. But the operation restricted to the subset $(\mathbb{F} \setminus \{0\})$ of $\mathbb{F}$ is again an abelian group with the identity 1, and with the multiplicative inverse of an element c in $(\mathbb{F} \setminus \{0\})$ written as c^{-1}. [Thus 0 is the only element in $\mathbb{F}$ which does not have a multiplicative inverse.] Now it is necessary to relate these two operations in $\mathbb{F}$ to one another, and this is done by the law of distribution. We say the multiplication distributes over addition on the left if for any elements a, b, and c in $\mathbb{F}$, we have $a(b + c) = ab + ac$. Since multiplication commutes in $\mathbb{F}$, this implies that multiplication also distributes over addition on the right, i.e., $(a + b)c = ac + bc$.

A.1.3 The Concept of a Linear Space V over a scalar field $\mathbb{F}$. A linear space V over a scalar field $\mathbb{F}$, where V is an abelian group and $\mathbb{F}$ is a field, defines a binary function $\mathbb{F} \times V \longrightarrow V$, written as $(c, v) \longmapsto cv \in V$, and this function is given the name of *scalar multiplication*. In order to complete our definition of a linear space V over the scalar field $\mathbb{F}$ we need to relate the two operations in the field $\mathbb{F}$ with the addition operation in the linear space V. Relating an element c in the field $\mathbb{F}$ with the addition of two elements u and v in V, we have the first relation for scalar multiplication $c(u + v) = cu + cv$. Relating two elements a and b in the field $\mathbb{F}$ with an element v in V, we have for addition in $\mathbb{F}$ the second relation for scalar multiplication $(a + b)v = av + bv$. Relating two elements a and b in the field $\mathbb{F}$ with an element v in V, we have for multiplication in $\mathbb{F}$ the third relation for scalar multiplication $(ab)v = a(bv)$. We also have the relationship of the identity 1 in $\mathbb{F}$ with any element v in V: $1v = v$, giving the fourth and final property of scalar multiplication. As seen in the above treatment, the elements of the linear space are in the set V. Since one example of this structure is a beautiful model for the Euclidean Plane with an arbitrary but fixed point called the *origin*, the elements of V are frequently called *vectors*, even though we might not be referring to this model. In this context, the elements of the scalar field $\mathbb{F}$ are called *scalars*. Another mathematical entity which exhibits this structure is the set of all n-tuples $\mathbb{F}^n$ of the field $\mathbb{F}$, with addition defined coordinate-wise $(a_1, \cdots, a_n)+(b_1, \cdots, b_n) = (a_1+b_1, \cdots, a_n+b_n)$, and scalar multiplication defined by $c(a_1, \cdot\cdot\cdot, a_n) = (ca_1, \cdot\cdot\cdot, ca_n)$. [Also we observe that the inverse c^{-1} of an element c in $\mathbb{F}$ does not enter into the definition of a linear space over a field $\mathbb{F}$. Thus we could weaken the structure on the set $\mathbb{F}$ to one in which 0 is not the only element without a multiplicative inverse. In this case that structure is called an associative, commutative *ring* R with an identity, and the abelian group V is called a *module* over the ring R. But we shall not pursue these concepts in this work.]

A.1.4 Bases for a Linear Space. From this point on we will consider only linear spaces which are <u>finite</u> dimensional. This means that there exists a linearly independent <u>finite</u> ordered subset $\mathcal{B}$ of V, called a <u>basis</u> of V, which spans V. This means that each vector can be expressed as a linear combination of the vectors in the basis and that these combinations are unique because the vectors in the basis are linearly independent. Thus suppose $\mathcal{B}$ $= (v_1, \cdot\cdot\cdot, v_n)$ [where the order is given by the integer subscripts]. Then for <u>any</u> element v in V, $v = c_1v_1 + \cdot\cdot\cdot + c_nv_n$, where the n-scalars $c_1, \cdot\cdot\cdot, c_n$ are uniquely determined. Now to prove that such a subset exists we would need the full force of the scalars being a field $\mathbb{F}$ We also know that every other basis for V has the same finite number of elements. [The proof of this fact is also omitted here.] Thus we can associate with V a non-negative integer n which describes the dimension of the linear space V. Note that we say that the dimension of V is 0 if the basis for V is the empty set and then V consists of only the vector 0.

We remark that in the linear space $\mathbb{F}^n$, where $n > 0$, there is a natural or canonical basis $(e_1, \cdots, e_n)$, where e_i is the n-tuple with 0 in each coordinate except the ith-coordinate, where it is 1: thus $e_i = (0, \cdots, 0, 1, 0, \cdots, 0)$, where 1 is the ith-coordinate. We can conclude that $\mathbb{F}^n$ has dimension n.

A.1.5 The Fields of Scalars: $\mathbb{R}$ and $\mathbb{C}$. The only fields of scalars that we will consider in this study are the field of real numbers $\mathbb{R}$ and the field of complex numbers $\mathbb{C}$. We pause here to make some remarks about these fields. The fundamental field is the field of real numbers. It is defined mathematically as the "complete ordered field". [The word "the" implies that there is essentially only one complete ordered field. Thus, no matter in what guises they may appear, all complete ordered fields are isomorphic to one another.] We have already discussed the meaning of a field. Now a field is ordered if for any two elements a and b in the field, with $a \neq b$, we can assert that either $a < b$ or $b < a$, where $a < b$ means that $b - a$ is a positive real number, i.e., is in the interval $(0, +\infty)$. [Obviously both relations cannot be true, for if $a < b$ is true, then $b < a$ means $a - b$ is a positive real number. But $a - b = -(b - a)$. Since $b - a$ is a positive real number, this means that $-(b - a)$ is a negative number, a contradiction.] Finally we say an ordered field is complete if any non-empty set which has an upper bound also has a least upper bound. [This concept is a concept from analysis, and not from algebra. We just quote it and will make no more comments about it. But it does give us the conclusion that $\mathbb{R} = \{set\ of\ rational\ numbers\} \cup \{set\ of\ irrational\ numbers\}$; and $\{set\ of\ rational\ numbers\} \cap \{set\ of\ irrational\ numbers\} = nullset$.]

What is important is that in an algebraic context, where we can write an algebraic equation such as $x^2 - 2 = 0$, where 2 is a rational number, x needs to belong to the set of real numbers for the equation to have a solution. Since $x^2 - 2 = 0$ means $x^2 = 2$, this says that $x = \pm\sqrt(2)$, which are irrational numbers. The fact that $\mathbb{R}$ is a **complete** ordered field allows us to conclude that it contains solutions to equations like $x^2 - 2 = 0$, but it does not furnish solutions to all polynomials over $\mathbb{R}$, for the real field is not algebraically closed. (The reader should review what "complete" means.)

Now it is a fact that the algebraic equation, the polynomial equation over $\mathbb{R}$ $x^2 + 1 = 0$, where 1 is a real number, has no solution in the field of real numbers. Since $x^2 + 1 = 0$ means $x^2 = -1$, and we know that the square of any real number is non-negative. But there is another field which is the **algebraic closure** of the field of real numbers, in which equations of this type have solutions. It is here where the field of **complex numbers** makes its appearance. The assertion that this field exists is the incredibly beautiful but difficult to prove *Fundamental Theorem of Algebra*. In so far as we will use this result we can state it as follows. Every polynomial in one indeterminate x with coefficients in the field of complex numbers can be factored into linear factors over the complex numbers, where the factoring is unique except for its order. Since the real field is contained in the complex field, this means that any such polynomial with coefficients in the field of real numbers can be factored over the the field of complex numbers into linear factors. Thus the above polynomial $x^2 + 1$ has the factorization $x^2 + 1 = (x + i)(x - i)$, where i is the so-called imaginary complex number such that $i^2 = -1$.

A.1.6 Linear Transformations. The Dimension Theorems and the Isomorphisms. Some of the most important results in Linear Algebra are the two dimension theorems and the corresponding isomorphism theorems. The context of the first dimension theorem concerns linear subspaces of a given linear space V. A subset W of a linear space V is a linear subspace if W is a linear space with respect to the same addition and scalar multiplication which defines V. This means that addition closes in W, i.e., the sum of any two elements in W remains in W and a scalar multiple of an element in W remains in W, that 0 is in W, and that the additive inverse of any element in W remains in W also. When W is a linear subspace of V, the concept of a quotient space or coset space also appears and is written as V/W. It is given a natural linear space structure: Two cosets $u + W$ and $v+W$ add together to give $(u+v)+W$; the 0 coset is $0+W$; and the additive inverse of the coset $u+W$ is $(-u)+W$; and for scalar multiplication we have: c(v+W) = cv + W.

The first dimension theorem and the corresponding isomorphism theorem considers two linear subspaces V_1 and V_2 of the same linear space V. Now the symbol $V_1 + V_2$ means all possible sums $v_1 + v_2$ with v_1 in V_1 and v_2 in V_2 It is an easy conclusion that $V_1 + V_2$ is a linear subspace of V. The first dimension theorem asserts:

$$\dim(V_1 + V_2) = \dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2)$$

or

$$\dim(V_1 + V_2) - \dim(V_1) = \dim(V_2) - \dim(V_1 \cap V_2)$$

which translates immediately into the isomorphism theorem between the quotient spaces, where the symbol $\cong$ indicates a relation of isomorphism:

$$(V_1 + V_2)/V_1 \cong V_2/(V_1 \cap V_2)$$

The second dimension theorem and the corresponding isomorphism theorem also put us in the context of a linear map ϕ between two linear spaces V and W over the same scalar field. Linear maps are the homomorphisms in the study of Linear Algebra. Thus a linear map between two linear spaces V and W over the same scalar field is a map that preserves addition and scalar multiplication:

$$\begin{aligned} V &\xrightarrow{\phi} W \\ v &\longrightarrow \phi(v) \end{aligned}$$

is such that

$$\phi(u + v) = \phi(u) + \phi(v) \qquad \textit{and} \qquad \phi(cu) = c\phi(u)$$

Now any linear transformation ϕ determines two subspaces: the kernel of ϕ, written as $\ker(\phi)$, is a subspace of the domain V of ϕ; and the image of ϕ, given by image(ϕ), is a subspace of the target space W of ϕ. Their definitions are:

$$\begin{aligned} \ker(\phi) &:= \{v \in V | \phi(v) = 0\} \\ \text{image}(\phi) &:= \{w \in W | \phi(v) = w \text{ for some v} \in \text{V}\} \end{aligned}$$

We remark that if ϕ is injective then $\ker(\phi) = 0$; while if ϕ is surjective then $\text{image}(\phi) = W$. Evidently if ϕ is bijective then $\ker(\phi) = 0$ and $\text{image}(\phi) = W$ and ϕ is an isomorphism.

The second dimension theorem and the corresponding isomorphism theorem are the following:

$$\dim(V) = \dim(\ker(\phi)) + \dim(\text{image}(\phi))$$

or

$$\dim(V) - \dim(\ker(\phi)) = \dim(\text{image}(\phi))$$

This equation, if ϕ surjects onto W, gives immediately the isomorphism between the quotient space and W or at least image(ϕ) if ϕ does not surject:

$$V/\ker(\phi) \cong \text{image}(\phi)$$

In our notation, given two linear spaces V and W over the same scalar field, the set of all linear maps from V to W is symbolized by $Hom(V,W)$, and the set of linear isomorphisms from V to W is symbolized by $Iso(V,W)$. [In this situation the dimension of V must be the same as the dimension of W.] However if $W = V$, i.e., if the domain and target spaces are the same, then we say that the set of all linear transformations from V to V is symbolized by $End(V)$, the linear endomorphisms on V. If these maps are all bijective, then we say that the set of these linear isomorphisms from V to V is symbolized by $Aut(V)$, the linear automorphisms of V. We also say that these maps are the nonsingular or invertible linear transformations. Within $End(V)$ we have some very special linear transformations, called *nilpotent* linear transformations. A linear transformation A is nilpotent if after k-iterations (k positive) it becomes the null transformation. i.e., $A^k = 0$, while $A^{k-1} \neq 0$.

We can also make $Hom(V,W)$ into a linear space over the same field $\mathbb{F}$ by a process called pointwise addition and pointwise scalar multiplication:

$$\begin{aligned} Hom(V,W)) \times Hom(V,W)) &\xrightarrow{\alpha+\beta} Hom(V,W)) \\ (\alpha,\beta) &\longmapsto \alpha+\beta : V \longrightarrow W \\ &\qquad v \longmapsto (\alpha+\beta)(v) := \alpha(v)+\beta(v) \end{aligned}$$

$$\begin{aligned} \mathbb{F} \times Hom(V,W) &\xrightarrow{c\alpha} Hom(V,W) \\ (c,\alpha) &\longmapsto c\alpha : V \longrightarrow W \\ &\qquad v \longmapsto (c\alpha)(v) := c(\alpha(v)) \end{aligned}$$

It is straightforward to show that $\alpha+\beta$ and $c\alpha$ are also linear transformations. It is interesting to note that $Iso(V,W)$ is not a linear subspace of $Hom(V,W)$. [If α is in $Iso(V,W)$, then $-\alpha$ is also in $Iso(V,W)$ but $\alpha + (-\alpha) = 0$, but the zero transformation is certainly not an isomorphism. However $c\alpha$, $c \neq 0$, is in $Iso(V,W)$.]

A.1.7 Matrices. Another concept from Linear Algebra is that of a matrix. An *mxn matrix* A in M_{mxn} is an array of field elements arranged in m rows and n columns:

$$A = [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ . & . & \cdots & . \\ . & . & \cdots & . \\ . & . & \cdots & . \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

We can give M_{mxn} the structure of a linear space over $\mathbb{F}$ by defining addition and scalar multiplication entrywise:

$$A + B = [a_{ij}] + [b_{ij}] := [a_{ij} + b_{ij}]$$
$$cA = c[a_{ij}] := [ca_{ij}]$$

A row matrix is an element in $M_{1\text{x}n}$, i.e., it is of the form

$$[a_{11}\ a_{12} \cdots a_{1n}]$$

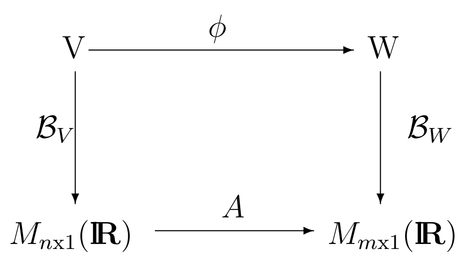
and a column matrix is an element in $M_{m\text{x}1}$, i.e., it is of the form

$$\begin{bmatrix} a_{11} \\ a_{21} \\ \cdot \\ \cdot \\ \cdot \\ a_{m1} \end{bmatrix}$$

There is another binary operation on matrices, that of row-by-column multiplication. This can be defined if the number of columns of the first matrix A equals the number of rows of the second matrix B, i.e., A is a mxn and B is a nxp. Then the product matrix AB is an mxp matrix, defined by

$$AB = [a_{ij}][b_{jk}] := [\sum_{j=1}^{n} a_{ij}b_{jk}]$$

The importance of matrices is that they can be representations of linear transformations. If we are given a finite dimensional linear space V over $\mathbb{F}$ of dimension n, and a finite dimensional linear space W over $\mathbb{F}$ of dimension m, and if we choose a basis $\mathcal{B}_V = (v_1, \cdots, v_n)$ for V and a basis $\mathcal{B}_W = (w_1, \cdots, w_m)$ for W, then any linear transformation ϕ from V to W can be given a matrix representation $A = [a_{ij}]$ in $M_{m\text{x}n}$. The following commutative diagram illustrates this situation.



The matrix $A = [a_{ij}]$ in $M_{m\text{x}n}$ is defined as follows. The j-th column is obtained by taking the j-th basis vector v_j for V, transforming it over to W by ϕ, and then writing $\phi(v_j)$ in terms of the basis for W, giving the m scalars which form the j-th column of the matrix A, i.e., $\phi(v_j) = \sum_{i=1}^{m} a_{ij}w_i$. In this correspondence between linear transformations and matrices, it is necessary to note that in the symbol M_{mxn} the target dimension m comes first and the domain dimension comes second. Thus there is a reversal in the manner of writing the dimensions in this correspondence.

We remark that in this interpretation an element v in V (or w in W) is represented by a column matrix whose entries are the scalar coefficients of v [or of w] written with respect to a basis $\mathcal{B}_V$ of V [$\mathcal{B}_W$ of W].

The above definitions of $(Hom(V, W), +, scalarmultiplication)$ and $(M_{mxn}, +, scalarmultiplication)$ give an isomorphism between these two linear spaces, and indeed this is the meaning of the word *representation* used above. We remark that in the case of $End(V)$, the corresponding matrices are square matrices.

The row-by-column multiplication of matrices was chosen so that the following relation is realized. If we take three linear spaces over the same scalar field, U of dimension p with basis $\mathcal{B}_U$, V of dimension n with basis $\mathcal{B}_V$, and W of dimension m with basis $\mathcal{B}_W$; and two linear maps $\phi_1 : U \longrightarrow V$ represented by the matrix A in M_{nxp}, and $\phi_2 : V \longrightarrow W$ represented by the matrix B in M_{mxn}, then the matrix C which represents the composition of the two maps $\phi_2 \circ \phi_1$ is exactly the product of the matrices B and A: $C = BA$. With these remarks one can also see that the non-singular matrices correspond to the invertible transformations.

A.1.8 The Trace and the Determinant of a Matrix. Now that we are considering square matrices, we can define two important functions of a linear transformation ϕ of a linear space V into itself: the trace and the determinant of a linear transformation. The target space of both of these functions is the field $\mathbb{F}$. Both of these will be defined by choosing a matrix representation of the linear transformation, which in this case will be a square matrix since both the domain and the target spaces are the same, and then we will prove that the function's value is independent of such a matrix representation.

The *trace* of a linear transformation $\phi : V \longrightarrow V$ in $End(V)$, with the dimension of V equal to n, is a linear function $tr(\phi) : V \longrightarrow \mathbb{F}$. We define it off of a matrix representation $A = [a_{ij}]$ of ϕ. After choosing a basis for V, giving us a matrix representation $A = [a_{ij}]$, $tr(A)$ is the sum of the diagonal elements of A: $tr(A) := \sum_{i=1}^{n} a_{ii}$. We know that we can give the set $End(V)$ a linear space structure, and we can then assert that the trace function on $End(V)$ is linear, i.e., $tr(A + B) = tr(A) + tr(B)$ and $tr(cA) = c(tr(A))$, where c is any scalar. It also has the property of being cyclic with respect to multiplication [or composition], i.e.: $tr(ABC) = tr(CAB) = tr(BCA)$. As a corollary we have $tr(AB) = tr(BA)$.

The *determinant* of a linear transformation $\phi : V \longrightarrow V$ in $End(V)$, with the dimension of V equal to n, is an alternating n-multilinear function $det(\phi) : V \times \cdots \times V \longrightarrow \mathbb{F}$. We define it off of a matrix representation $A = [a_{ij}]$ of ϕ. After choosing a basis for V, giving us a matrix representation $A = [a_{ij}]$, we read the n-columns of this matrix as the representation of n vectors $(v_1, \cdots, v_n)$ in V, or as we say, a representation of an n-multivector in V^n. Here multilinear means that in each term of the domain we have linearity with respect to addition and scalar multiplication, i.e.:

$$det(\phi)(v_1, \cdots, u_i+v_i, \cdots, v_n) = det(\phi)(v_1, \cdots, u_i, \cdots, v_n)+det(\phi)(v_1, \cdots, v_i, \cdots, v_n)$$

and

$$det(\phi)(v_1, \cdots, c(u_i), \cdots, v_n) = c(det(\phi)(v_1, \cdots, v_i, \cdots, v_n))$$

for any scalar c. An alternating multilinear function is one in which if two adjacent vectors are interchanged in the domain, then the value of the function changes sign.

$$det(\phi)(v_1, \cdots, v_i, v_j, \cdots, v_n) = -(det(\phi)(v_1, \cdots, v_j, v_i, \cdots, v_n))$$

The definition of the determinant function is complicated and at this level very unintuitive. We will just give it by an induction procedure on the dimension of V. Having chosen a basis for V, we now define the determinant of ϕ in terms of the matrix representation A of ϕ, that is, $det(A)$, which is an element of $\mathbb{F}$. If the dimension of V is one, then the $det([a_{11}]) := a_{11}$. If the dimension of V is 2, then the $det([a_{ij}]) := a_{11}det([a_{22}]) - a_{12}det([a_{21}]) = a_{11}a_{22} - a_{12}a_{21}$. If the dimension of V is 3, then

$$det([a_{ij}]) = det\left(\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\right) :=$$
$$+a_{11}det\left(\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}\right) - a_{12}det\left(\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}\right) + a_{13}det\left(\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}\right) =$$
$$+a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{12}(a_{21}a_{33} - a_{31}a_{23}) + a_{13}(a_{21}a_{32} - a_{31}a_{22}) =$$
$$+a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} - a_{12}a_{21}a_{33} + a_{12}a_{31}a_{23} + a_{13}a_{21}a_{32} - a_{13}a_{31}a_{22}$$

Guided by the definition for the first few dimensions, we can now give full inductive definition of the determinant function:

$$det([a_{ij}]) = det\left(\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}\right) :=$$

$$+a_{11}det\left(\begin{bmatrix} a_{22} & \cdots & a_{2n} \\ \cdot & \cdots & \cdot \\ a_{m2} & \cdots & a_{mn} \end{bmatrix}\right) - a_{12}det\left(\begin{bmatrix} a_{21} & a_{23} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{m1} & a_{m3} & \cdots & a_{mn} \end{bmatrix}\right) + \cdots$$

$$\pm a_{1n}det\left(\begin{bmatrix} a_{21} & \cdots & a_{2(n-1)} \\ \cdot & \cdots & \cdot \\ a_{m1} & \cdots & a_{m(n-1)} \end{bmatrix}\right)$$

where the determinants in the sum are taken on the submatrices which are missing the i-th row and the j-th column if a_{ij} is the coefficient of that determinant. In our definition the elements in the first row are used but it could be any row. We see that if the dimension of V is n, then the number of terms in the full expansion of the determinant is $n!$, which is the number of permutations on n objects and/or the order of the n-th order symmetric group. Now each element in the expanded defining equation carries with it a $\pm$ sign, and thus we see that the determinant is an alternating multilinear map, where half of the terms in the expansion carry a positive sign and half of the terms carry a negative sign. Each set of subscripts found in each term defines a permutation, and the sign of the permutation of the numbers 1, 2 ..., n is given to the term in the expansion of the determinant.

Another interpretation – a geometric one – can be given to the determinant. Suppose our linear space is $\mathbb{R}^n$ and our basis is the canonical basis. Each column of the nxn matrix representing a linear transformation is the image of one of these canonical vectors. These n vectors now determine an n-dimensional parallelepiped in $\mathbb{R}^n$. If they are linearly independent then the determinant of this matrix determines a real number which is the oriented Euclidean n-volume of this parallelepiped. If the vectors are not linearly independent, then the n-volume is zero. From these observations we can conclude that the determinant of a singular transformation is zero, while that of a non-singular transformation is not zero.

With respect to multiplication of matrices [or composition of functions] we have the following nice property of the determinant function:

$$det(AB) = det(A)\, det(B)$$

We do not prove this property but just here declare it.

A.1.9 Eigenspaces, Eigenvalues, Characteristic Polynomial and the Jordan Decomposition Theorem. The eigenspaces of linear transformations along with their eigenvectors and eigenvalues play an important part in our development. In this context we fix our attention on one linear transformation T in $End(V)$, where V is a linear space of dimension n.over the field $\mathbb{F}$. [Recall, again, that the only fields that we are considering are $\mathbb{R}$ and $\mathbb{C}$]. We seek vectors $v \neq 0$ in V with the following property:

$$Tv = \lambda v$$

where λ is a scalar in $\mathbb{F}$. This essentially says the T stabilizes v in a one-dimensional subspace of V. Such a vector is called an *eigenvector* with the *eigenvalue* $= \lambda$. Now we can rewrite this information as follows:

$$Tv = \lambda v \quad \text{or} \quad \lambda v - Tv = 0 \quad \text{or} \quad \lambda I_n v - Tv = 0 \quad \text{or} \quad (\lambda I_n - T)v = 0$$

where I_n is the identity function in $End(V)$. Since $v \neq 0$, this means that $(\lambda I_n - T)$ has a non-zero kernel. Thus this property identifies these special scalars, the eigenvalues, which in turn identify the eigenvectors. In order to obtain these scalars we use the fact that $(\lambda I_n - T)$ has a non-zero kernel implies that $(\lambda I_n - T)$ is a singular linear transformation, and thus its determinant is 0. Now we know that we need to choose a basis for V in order to calculate this determinant and its eigenvalues. What is remarkable is that these eigenvalues do not depend on the basis chosen to calculate this determinant. And this determinant will be an nth-degree polynomial in the indeterminate λ over the field $\mathbb{F}$. Thus we have this polynomial which is uniquely determined by a linear transformation T. It is called the *characteristic polynomial of the linear transformation* T. The roots of this polynomial in the field $\mathbb{F}$ are exactly the eigenvalues of the transformation T.

At this point it is necessary to bring in the algebraic closure of $\mathbb{R}$, which is $\mathbb{C}$. With this field of scalars we know that the characteristic polynomial of a linear transformation can be factored into linear polynomials. Also any linear space V over $\mathbb{R}$ can be considered a linear space over $\mathbb{C}$, since $\mathbb{R}$ is a subfield of $\mathbb{C}$. Thus among the linear polynomials of this factorization will again appear those with real factors, if there be any.

For a linear space V over $\mathbb{C}$ there is a beautiful theorem that states that for any linear transformation T of V, there exists a basis such that T is represented by a matrix A of the form

$$A = \begin{bmatrix} c_1 & a_{12} & a_{13} & a_{14} & \cdots & a_{1n} \\ 0 & c_2 & a_{23} & a_{24} & \cdots & a_{2n} \\ 0 & 0 & c_3 & a_{34} & \cdots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & 0 & \cdots & c_n \end{bmatrix}$$

In fact more can be said. This more is the content of the *Jordan Decomposition Theorem*. The linear transformation T of V over $\mathbb{C}$ can be written uniquely as a sum of two linear transformations, $T = S + N$, where S is diagonalizable and N is nilpotent, with $SN = NS$, and in which both S and N can be written as polynomials in T with zero constant term .

A.1.10 The Jordan Canonical Form. In this context we would also like to quote the theorem on the Jordan Canonical Form. First let us describe a kxk Jordan block matrix $J(c;k)$, where c is in $\mathbb{C}$ and k is a positive integer:

$$J(c;k) = \begin{bmatrix} c & 1 & & & & & \\ & c & 1 & & & & \\ & & c & 1 & & & \\ & & & \cdot & \cdot & & \\ & & & & \cdot & \cdot & \\ & & & & & \cdot & 1 \\ & & & & & & c \end{bmatrix} \in M_{k\text{xk}}$$

and where all other entries are zero. We remark that $J(c;k)$ is of the form $J(c;k) = S + N$, where S is the diagonal matrix with c on the diagonal, and N is nilpotent [an upper triangular matrix with a zero diagonal]. Also we see that $SN = (cI_k)(N) = cN = cNI_k = (N)(cI_k) = NS$, and we affirm that $S = cI_k$ and N can be written as polynomials in $J(c;k)$ <u>without constant terms</u>. [Since this result is not that well-known, we will give a detailed calculation of it in an example below. In the meantime let us continue in our context.] We now begin to put together these Jordan blocks. For $k_1 \geq k_2 \geq \cdots \geq k_p$, we define

$$J(c;k_1,\cdots,k_p) = \begin{bmatrix} J(c;k_1) & & & & & \\ & J(c;k_2) & & & & \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & & J(c;k_p) \end{bmatrix}$$

We remark that the size of this matrix is $(\sum_{i=1}^{p} k_i)$x$(\sum_{i=1}^{p} k_i)$. Finally we give the *Jordan Matrix* of the Jordan Canonical Form:

$$J = \begin{bmatrix} J(c_1;k_1^{(1)},\cdots,k_{p_1}^{(1)}) & & & & \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & J(c_m;k_1^{(m)},\cdots,k_{p_m}^{(m)}) \end{bmatrix} \in M_{n\text{x}n}$$

The theorem states that any linear transformation $T \in End(V)$, where V is an n-dimensional linear space over $\mathbb{C}$ and V has a basis which represents T as Jordan Matrix J, and where the $\{c_1,\cdots,c_m\}$ are the distinct eigenvalues of T with multiplicities $\{n_1,\cdots,n_m\}$ respectively and where $n_i = \sum_{j=1}^{p_i} k_j^{(i)}$ for $1 \leq i \leq m$. The dimension of V is $n = \sum_{i=1}^{m} n_i$. The characteristic polynomial of T is $(x - c_1)^{n_1} \cdots (x - c_m)^{n_m}$. Lastly the Jordan Canonical Form is unique up to the order of the blocks $J(c_i;k_i^{(i)},\cdots,k_{p_i}^{(i)})$.

A.1.11 An Example.

We now give in detail the calculations mentioned above, which show how we may write the matrices S and N as polynomials in $J(c;k)$. First we remark that if $c = 0$, $J(0;k) = N$ [and S is 0], and we have our conclusion trivially. However for $c \neq 0$, since we are writing these matrices in general form, it would be rather complicated to give the exact entries in these matrices. But we can give the flavor of these calculations by giving this decomposition for $k = 2$ and $k = 4$ and $c \neq 0$.

$$J(c;2) = \begin{bmatrix} c & 1 \\ 0 & c \end{bmatrix}$$

$$J(cI_2) = \begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix} = r \begin{bmatrix} c & 1 \\ 0 & c \end{bmatrix} + s \begin{bmatrix} c & 1 \\ 0 & c \end{bmatrix}^2$$

$$J(cI_2) = \begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix} = r \begin{bmatrix} c & 1 \\ 0 & c \end{bmatrix} + s \begin{bmatrix} c^2 & 2c \\ 0 & c^2 \end{bmatrix}$$

One observes that there are only two independent equations:

$$\begin{aligned} c &= r(c) + s(c^2) \\ 0 &= r + s(2c) \end{aligned}$$

Their solution is

$$r = 2 \qquad s = -\tfrac{1}{c}$$

giving

$$cI_2 = \begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix} = 2 \begin{bmatrix} c & 1 \\ 0 & c \end{bmatrix} - \tfrac{1}{c} \begin{bmatrix} c & 1 \\ 0 & c \end{bmatrix}^2$$

and

$$\begin{aligned} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} &= \begin{bmatrix} c & 1 \\ 0 & c \end{bmatrix} - cI_2 \\ &= - \begin{bmatrix} c & 1 \\ 0 & c \end{bmatrix} + \tfrac{1}{c} \begin{bmatrix} c & 1 \\ 0 & c \end{bmatrix}^2 \end{aligned}$$

The details for the $k = 4$ case are as follows:

$$J(c;4) = \begin{bmatrix} c & 1 & 0 & 0 \\ 0 & c & 1 & 0 \\ 0 & 0 & c & 1 \\ 0 & 0 & 0 & c \end{bmatrix}$$

$$cI_4 = \begin{bmatrix} c & 0 & 0 & 0 \\ 0 & c & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & c \end{bmatrix} =$$

$$r \begin{bmatrix} c & 1 & 0 & 0 \\ 0 & c & 1 & 0 \\ 0 & 0 & c & 1 \\ 0 & 0 & 0 & c \end{bmatrix} + s \begin{bmatrix} c & 1 & 0 & 0 \\ 0 & c & 1 & 0 \\ 0 & 0 & c & 1 \\ 0 & 0 & 0 & c \end{bmatrix}^2 + t \begin{bmatrix} c & 1 & 0 & 0 \\ 0 & c & 1 & 0 \\ 0 & 0 & c & 1 \\ 0 & 0 & 0 & c \end{bmatrix}^3 + u \begin{bmatrix} c & 1 & 0 & 0 \\ 0 & c & 1 & 0 \\ 0 & 0 & c & 1 \\ 0 & 0 & 0 & c \end{bmatrix}^4$$

$$cI_4 = \begin{bmatrix} c & 0 & 0 & 0 \\ 0 & c & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & c \end{bmatrix} = r \begin{bmatrix} c & 1 & 0 & 0 \\ 0 & c & 1 & 0 \\ 0 & 0 & c & 1 \\ 0 & 0 & 0 & c \end{bmatrix} + s \begin{bmatrix} c^2 & 2c & 1 & 0 \\ 0 & c^2 & 2c & 1 \\ 0 & 0 & c^2 & 2c \\ 0 & 0 & 0 & c^2 \end{bmatrix} +$$
$$t \begin{bmatrix} c^3 & 3c^2 & 3c & 1 \\ 0 & c^3 & 3c^2 & 3c \\ 0 & 0 & c^3 & 3c^2 \\ 0 & 0 & 0 & c^3 \end{bmatrix} + u \begin{bmatrix} c^4 & 4c^3 & 6c^2 & 4c \\ 0 & c^4 & 4c^3 & 6c^2 \\ 0 & 0 & c^4 & 4c^3 \\ 0 & 0 & 0 & c^4 \end{bmatrix}$$

One observes that there are only four independent equations:

$$\begin{aligned} c &= r(c) + s(c^2) + t(c^3) + u(c^4) \\ 0 &= r + s(2c) + t(3c^2) + u(4c^3) \\ 0 &= s + t(3c) + u(6c^2) \\ 0 &= t + u(4c) \end{aligned}$$

Their solution is

$r = 4$ $\quad s = -6\frac{1}{c}$ $\quad t = 4\frac{1}{c^2}$ $\quad u = -\frac{1}{c^3}$

giving:

$$cI_4 = \begin{bmatrix} c & 0 & 0 & 0 \\ 0 & c & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & c \end{bmatrix} = 4 \begin{bmatrix} c & 1 & 0 & 0 \\ 0 & c & 1 & 0 \\ 0 & 0 & c & 1 \\ 0 & 0 & 0 & c \end{bmatrix} - 6\frac{1}{c} \begin{bmatrix} c & 1 & 0 & 0 \\ 0 & c & 1 & 0 \\ 0 & 0 & c & 1 \\ 0 & 0 & 0 & c \end{bmatrix}^2 +$$
$$4\frac{1}{c^2} \begin{bmatrix} c & 1 & 0 & 0 \\ 0 & c & 1 & 0 \\ 0 & 0 & c & 1 \\ 0 & 0 & 0 & c \end{bmatrix}^3 - \frac{1}{c^3} \begin{bmatrix} c & 1 & 0 & 0 \\ 0 & c & 1 & 0 \\ 0 & 0 & c & 1 \\ 0 & 0 & 0 & c \end{bmatrix}^4$$

and

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} c & 1 & 0 & 0 \\ 0 & c & 1 & 0 \\ 0 & 0 & c & 1 \\ 0 & 0 & 0 & c \end{bmatrix} - cI_4$$

$$= -3 \begin{bmatrix} c & 1 & 0 & 0 \\ 0 & c & 1 & 0 \\ 0 & 0 & c & 1 \\ 0 & 0 & 0 & c \end{bmatrix} + 6\tfrac{1}{c} \begin{bmatrix} c & 1 & 0 & 0 \\ 0 & c & 1 & 0 \\ 0 & 0 & c & 1 \\ 0 & 0 & 0 & c \end{bmatrix}^2$$

$$-4\tfrac{1}{c^2} \begin{bmatrix} c & 1 & 0 & 0 \\ 0 & c & 1 & 0 \\ 0 & 0 & c & 1 \\ 0 & 0 & 0 & c \end{bmatrix}^3 + \tfrac{1}{c^3} \begin{bmatrix} c & 1 & 0 & 0 \\ 0 & c & 1 & 0 \\ 0 & 0 & c & 1 \\ 0 & 0 & 0 & c \end{bmatrix}^4$$

A.1.12 The Dual Space.

Another key notion in Linear Algebra is that of the dual V^* of a linear space V. This is a very subtle concept since sets of duals are special kinds of linear transformations and not just a set of elements as is V. We define the dual V^* of V to be the set of linear transformations from V to the scalar field $\mathbb{F}$. This is a clearly defined idea and thus there is no ambiguity about what the elements of V^* are. If we give V a basis $(e_1, \cdots, e_n)$, we can define a dual basis $((e^1)^*, \cdots, (e^n)^*)$ by $(e^i)^*(e_j) = \delta^i{}_j$

If we have a linear map ϕ between two linear spaces V and W over the same scalar field

$$V \xrightarrow{\phi} W$$
$$v \longrightarrow \phi(v)$$

then we can define immediately a dual linear map ϕ^* between the two dual spaces V^* and W^*, but we have to reverse the domain and the target spaces:

$$W^* \xrightarrow{\phi^*} V^*$$
$$w^* \longrightarrow \phi^*(w^*)$$

where $(\phi^*(w^*))(v) := (w^*)(\phi(v))$

Choosing dual bases, we can put this information in terms of matrices. Let V be n-dimensional and W be m-dimensional. Then we have the following two diagrams.

$$
\begin{array}{ccccccc}
V & \xrightarrow{\phi} & W & & W^* & \xleftarrow{\phi^*} & V^* \\
\mathcal{B}_V \downarrow & & \downarrow \mathcal{B}_W & & \mathcal{B}_{W^*} \downarrow & & \downarrow \mathcal{B}_{V^*} \\
M_{n\text{x}1}(\mathbb{R}) & \xrightarrow{A} & M_{m\text{x}1}(\mathbb{R}) & & M_{m\text{x}1}(\mathbb{R}) & \xleftarrow{A^T} & M_{n\text{x}1}(\mathbb{R})
\end{array}
$$

The mxn matrix A corresponding to ϕ becomes the nxm transpose matrix A^T corresponding to ϕ^*.

In the case where v^* is a dual in V^* acting on V, choosing a basis $(e_1, \cdot\cdot\cdot, e_n)$ in V means that v^* becomes a 1xn row matrix $[v^*] = [(v^*)_1, \cdot\cdot\cdot, (v^*)_n]$, where $(v^*)_i = v^*(e_i)$. In matrix notation this becomes

$$
\begin{bmatrix} (v^*)_1 & \cdot & \cdot & \cdot & (v^*)_n \end{bmatrix} \cdot \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 1_i \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} = [(v^*)_i]
$$

The incredible fact is that there is no natural or canonical transformation to map V to V^*. If we give V a basis and V^* the dual basis, we can map V to V^* by the map $e_i \longmapsto (e^i)^*$. But this map is not canonical and depends on the basis chosen. However if we give V a bilinear form B, then we have a natural map between V and V^*. Now a bilinear form B on V is a mapping

$$
\begin{aligned}
B : V \times V &\longrightarrow \mathbb{F} \\
(u, v) &\longmapsto B(u, v)
\end{aligned}
$$

such that B is a linear map in each variable. Then the linear map $\mathcal{B}$ from V to V^* is defined as follows.

$$
\begin{aligned}
V &\xrightarrow{\mathcal{B}} V^* \\
u &\longmapsto \mathcal{B}(u) : V \longrightarrow \mathbb{F} \\
&\qquad\qquad v \longmapsto \mathcal{B}(u)(v) := B(u, v)
\end{aligned}
$$

This map is called nondegenerate if the only zero map comes from 0 in V. This means that if the map $\mathcal{B}(u)$ acting on any v in V is zero, then u in V must be zero.

A.1.13 Schur's Lemma.

An important tool in representation theory is Schur's Lemma. [Below we follow the exposition of [N]]. Assume that we have two linear spaces V and W over a field $\mathbb{F}$. Let $\mathcal{A}$ be a set of linear transformations in $End(V)$ and $\mathcal{B}$ be a set of linear transformations in $End(W)$. Let U_V be a subspace of V, and U_W be a subspace of W. For each X in $\mathcal{A}$, assume that $X(U_V)$ is contained in U_V. [In this situation we say that $\mathcal{A}$ acts invariantly on U_V. Also we say that $\mathcal{A}$ acts irreducibly on the subspace U_V if $\mathcal{A}$ leaves no proper subspace of U_V invariant]. Likewise we assume that $\mathcal{B}$ acts invariantly on U_W. Then *Schur's Lemma* affirms the following.

> Let V and W be linear spaces over a field $\mathbb{F}$. Let $\mathcal{A}$ be a set of linear transformations in $End(V)$ acting invariantly on a subspace U_V of V; and let $\mathcal{B}$ be a set of linear transformations in $End(W)$ acting invariantly on a subspace U_W of W. Let C be a linear map from V to W which satisfies the following conditions:
>
> 1) For any A in $\mathcal{A}$ there exists a B in $\mathcal{B}$ such that $CA = BC$.
>
> 1) For any B in $\mathcal{B}$ there exists a A in $\mathcal{A}$ such that $CA = BC$.
>
> Assume further that both $\mathcal{A}$ and $\mathcal{B}$ act irreducibly on V and W respectively. Then C is either the zero map of V into W, or is an isomorphism of V onto W [and hence $\mathcal{B} = \{CAC^{-1}, A \in \mathcal{A}\}$].

We remark that in the situation expressed in Schur's Lemma, we can prove that both the $ker(C)$ in V is an invariant subspace of V by $\mathcal{A}$, and the $Im(C)$ in W is an invariant subspace of W by $\mathcal{B}$.

As a corollary of Schur's Lemma, we have the following important specification of the linear transformation C of Schur's Lemma.

> Let V be a linear space over $\mathbb{C}$ [an algebraically closed field]. Let $\mathcal{A}$ be a set of linear transformations in $End(V)$ acting irreducibly on V. Then every linear transformation C of V which commutes with every transformation A in $\mathcal{A}$ has the form cI, where c is an element of $\mathbb{C}$ and I is the identity transformation of V.

A. 2 Cartan's Classification of Simple Complex Lie Algebras

Note that in this appendix, many of the results are merely declared and are not proved. The reader is strongly advised to consult [K] or [FH] for fuller treatments.

We start with a semisimple complex Lie algebra $\hat{g}$. We seek what is called a *Cartan subalgebra of* $\hat{g}$. We know that every simple complex Lie algebra contains a Cartan subalgebra but the existence of such an algebra is not easy to prove. The problem is that its definition contains a complicated maximality property, and this property is not a trivial matter to verify. We give one description of this Cartan subalgebra, which here we shall call $\hat{h}$.

Thus, a subalgbra $\hat{h}$ is called a *Cartan subalgebra* if it is a maximal abelian subalgebra such that for each H in $\hat{h}$ the adjoint representation of H acts diagonally on $\hat{g}$. This gives us a characteristic equation for the operation $ad(H)$ acting on $\hat{g}$:

$$\det(ad(H) - \lambda I) = 0$$

and we have

$$\hat{g} = \oplus_{\alpha_i} \hat{g}_{\alpha_i}$$

where $\hat{g}_{\alpha_i}$ is an eigenspace parametrized by a linear form α_i in $\hat{h}^*$:

$$\hat{g}_{\alpha_i} = \{X \in \hat{g} | ad(H)X = [H, X] = \alpha_i(H)X \text{ for all } H \text{ in } \hat{h}\}$$

and where X is the simultaneous eigenvector in $\hat{g}$ with the corresponding eigenvalue $\alpha_i(H)$ for each H in $\hat{h}$. We say that α_i in $\hat{h}^*$ is a *root* of the Lie algebra $\hat{g}$. Also, we have $ad(H_i)H = [H_i, H] = 0(H)$ for all H in $\hat{h}$ since $\hat{h}$ is abelian, and thus we see that we have zero roots, and that the zero root space $\hat{g}_0$ is equal to $\hat{h}$. We usually want to think of only the non-zero roots and thus we reserve the α_i symbols for the <u>non-zero</u> roots. Also, except for the zero eigenspace, which has the dimension of $\hat{h}$ [since, of course, $\hat{h}$ is an abelian subalgebra], we know that each non-zero eigenspace is one-dimensional. It is interesting to observe that this means the trace of the linear transformation $ad(H_i)$ is zero.

We call these one-dimensional non-zero eigenspaces $\hat{g}_\alpha$ the *root spaces* of $\hat{g}$. [The zero root space is $\hat{h}$.] Thus we can conclude that we have a decomposition of $\hat{g}$:

$$\hat{g} = \hat{h} \oplus (\oplus_{\alpha \neq 0} \hat{g}_\alpha)$$

and that the roots span a real subspace of $\hat{h}^*$. The dimension of the Cartan subalgebra $\hat{h}$, which is the zero root space, is called the *rank* of the Lie algebra $\hat{g}$.

We have that if α_i and α_j are roots [now, by exception, including the 0 root] and $\alpha_i + \alpha_j$ is a root then

$$[\hat{g}_{\alpha_i}, \hat{g}_{\alpha_j}] \subset \hat{g}_{\alpha_i+\alpha_j}]$$

but if $\alpha_i + \alpha_j$ is not a root, then

$$[\hat{g}_{\alpha_i}, \hat{g}_{\alpha_j}] = 0$$

We see that if $\alpha_i = 0$, then

$$[\hat{g}_0, \hat{g}_{\alpha_j}] \subset \hat{g}_{0+\alpha_j} = \hat{g}_{\alpha_j}$$

and we are just repeating that $\hat{g}_{\alpha_j}$ is a simultaneous eigenspace for all of $ad(\hat{g}_0) = ad(\hat{h})$

Since our Lie algebras are finite dimensional, we know that there are only a finite number of roots. Thus the expression above says that the bracket of two elements from two root spaces $\hat{g}_{\alpha_i}$ and $\hat{g}_{\alpha_j}$ produces an element in another root space if and only if $\alpha_i + \alpha_j$ is another root [in this case again, by exception, including the zero root space].

We also have that for each root α, $[\hat{g}_\alpha, \hat{g}_{-\alpha}]$ is non-zero and one-dimensional and from the statements given above is in $\hat{h}$, and such that

$$\hat{s}_\alpha = \hat{g}_\alpha \oplus \hat{g}_{-\alpha} \oplus [\hat{g}_\alpha, \hat{g}_{-\alpha}]$$

is a 3-dimensional subalgebra of $\hat{g}$ isomorphic to $\hat{sl}(2,\mathbb{C})$. We also have that $[[\hat{g}_\alpha, \hat{g}_{-\alpha}], \hat{g}_\alpha] \neq 0$. In fact, we pick a basis X_α in $\hat{g}_\alpha$ and a basis $X_{-\alpha}$ in $\hat{g}_{-\alpha}$ and H_α in $\hat{h}$ such that $[X_\alpha, X_{-\alpha}] = H_\alpha$ and $[H_\alpha, X_\alpha] = ad(H_\alpha)X_\alpha = \alpha(H_\alpha)X_\alpha = 2X_\alpha$ and $[H_\alpha, X_{-\alpha}] = ad(H_\alpha)X_{-\alpha} = -\alpha(H_\alpha)X_{-\alpha} = -2X_{-\alpha}$. We see that H_α in $\hat{h}$ is uniquely determined by these choices, while X_α and $X_{-\alpha}$ are not. [However, there is no motivation here for choosing the eigenvalue of $ad(H_\alpha)$ acting on $X_{-\alpha}$ to be 2, or -2 for $ad(H_\alpha)$ acting on $X_{-\alpha}$. But a natural basis for $\hat{sl}(2,\mathbb{C})$ has this structure:

$$H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} E = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \text{ and } F = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

where $[H, E] = 2E$ and $[H, F] = -2F$ and $[E, F] = H$. Thus each $\hat{s}_\alpha$ is isomorphic to $\hat{sl}(2,\mathbb{C})$] The set of all non-zero roots will be called Δ. We also remark that we can show that all the eigenvalues associated with the roots are integers. Thus the roots determine an integer-valued lattice on $\hat{h}^*$ and thus we can restrict ourselves to considering $\hat{h}^*$ as a linear space over the *rational* field.

These roots are symmetric about the origin. We express this by choosing for any root α an involution W_α of $\hat{h}^*$, defined by

$$W_\alpha : \hat{h}^* \longrightarrow \hat{h}^*$$
$$\beta \longrightarrow W_\alpha(\beta) := \beta - \tfrac{2\beta_1(H_\alpha)}{\alpha(H_\alpha)}\alpha = \beta - \beta(H_\alpha)\alpha$$

[Recall that $\alpha(H_\alpha) = 2$.]

W_α is certainly linear since

$$W_\alpha(\beta_1 + \beta_2) = (\beta_1 + \beta_2) - \tfrac{2(\beta_1+\beta_2)(H_\alpha)}{\alpha(H_\alpha)} = \beta_1 - \tfrac{2(\beta_1)(H_\alpha)}{\alpha(H_\alpha)}\alpha + \beta_2 - \tfrac{2(\beta_2)(H_\alpha)}{\alpha(H_\alpha)}\alpha = W_\alpha(\beta_1) + W_\alpha(\beta_2)$$

and

$$W_\alpha(c\beta) = c\beta - \tfrac{2c\beta(H_\alpha)}{\alpha(H_\alpha)} = c(\beta(H_\alpha) - \tfrac{2\beta(H_\alpha)}{\alpha(H_\alpha)}) = c(W_\alpha(\beta))$$

It is also an involution since

$$W_\alpha(W_\alpha(\beta)) = W_\alpha(\beta - \beta(H_\alpha)\alpha) = W_\alpha(\beta) - W_\alpha(\beta(H_\alpha)\alpha) =$$
$$\beta - \beta(H_\alpha)\alpha - \beta(H_\alpha)\alpha + \beta(H_\alpha)\alpha(H_\alpha)\alpha =$$
$$\beta - \beta(H_\alpha)\alpha - \beta(H_\alpha)\alpha + 2\beta(H_\alpha)\alpha = \beta$$

[This says that W_α, acting twice on β, is equal to β and thus is the identity map, which means that W_α is an involution.]

The $+1$ eigenspace [i.e., the invariant hyperplane of the involution] is defined as:

$$\Omega_\alpha := \{\beta \in \hat{h}^* | \beta(H_\alpha) = 0\}$$

The negative eigenspace is the one-dimensional line determined by α. Let us verify both of these statements. Thus, for any β in Ω_α we have

$$W_\alpha(\beta) = \beta - \beta(H_\alpha)\alpha = \beta - 0\alpha = \beta$$

which shows everything in Ω_α remains fixed by W_α; and

$$W_\alpha(\alpha) = \alpha - \alpha(H_\alpha)\alpha = \alpha - 2\alpha = -\alpha$$

which shows that α is flipped to its negative by W_α.

The set of all these involutions, the W_α's, form a group called the *Weyl group* of the Lie algebra. And the set of roots is invariant by the Weyl group. To prove this we need to know $\alpha_j(H_{\alpha_i})$.

We work now in the integer-valued lattice of $\hat{h}^*$. We choose a root α in Δ and we let β be any other root in Δ. We define the α-series containing β to be all the roots of the form $\beta + n\alpha$, where n is an integer. This series has a lower bound $n = p$ and an upper bound $n = q$ and therefore includes all integers n such that $p \leq n \leq q$. Recall that for each α in Δ, there corresponds a well-determined H_α in $\hat{h}$. Now for the α-series containing β we know that

$$-2\tfrac{\beta(H_\alpha)}{\alpha(H_\alpha)} = p + q$$

Since $\alpha(H_\alpha) = 2$ we know that $q = -(p + \beta(H_\alpha))$. In particular the α-series containing $-\alpha$ is

$$\{-\alpha, 0, \alpha\} = \{-\alpha + 0\alpha, -\alpha + 1\alpha, -\alpha + 2\alpha\}$$

where we see that $p = 0$ and $q = 2$. Thus $-2\frac{\beta(H_\alpha)}{\alpha(H_\alpha)} = 2$ and we do have $2 = 0 + 2$.

Once again, we choose X_α in $\hat{g}_\alpha$ and $X_{-\alpha}$ in $\hat{g}_{-\alpha}$ such that $[X_\alpha, X_{-\alpha}] = H_\alpha$ and $[H_\alpha, X_\alpha] = 2X_\alpha$ and $[H_\alpha, X_{-\alpha}] = -2X_{-\alpha}$. Now for the α-series containing β we have that, with X_β in $\hat{g}_\beta$,

$$[[X_\beta, X_{-\alpha}], X_{-\alpha}] = \tfrac{(p+1)q}{2}\alpha(H_\alpha)X_\beta$$

Recalling that for the α-series containing $-\alpha, p = 0$ and $q = 2$, we see that we have

$$[[X_{-\alpha}, X_\alpha], X_{-\alpha}] = \tfrac{(p+1)q}{2}\alpha(H_\alpha)X_{-\alpha} = \tfrac{(0+1)2}{2}[H_\alpha, X_{-\alpha}] = -2X_{-\alpha}$$

and

$$[[X_{-\alpha}, X_\alpha], X_{-\alpha}] = [H_\alpha, X_{-\alpha}] = -\alpha(H_\alpha)X_{-\alpha} = -2X_{-\alpha}$$

and we observe that these two expressions are equal.

One of the most amazing facts about the roots is that if α and β are not zero roots, then the α-string containing β contains at most four roots. This gives us the possibilities:

$$\text{For } p = 0 : \beta, \alpha + \beta, 2\alpha + \beta, 3\alpha + \beta \ [q = 3]$$
$$\text{For } p = -1 : -\alpha + \beta, \alpha + \beta, 2\alpha + \beta \ [q = 2]$$
$$\text{For } p = -2 : -2\alpha + \beta, -\alpha + \beta, \beta, \alpha + \beta \ [q = 1]$$
$$\text{For } p = -3 : -3\alpha + \beta . -2\alpha + \beta, -\alpha + \beta, \beta \ [q = 0]$$

since for $p = 1$ $[q = 4]$ and for $p = -4$ $[q = -1]$, β would disappear from the string. This means in this case that

$$-2\tfrac{\beta(H_\alpha)}{\alpha(H_\alpha} = -\beta(H_\alpha) = p + q = 3, 1, -1, 3$$

If the α-string containing β contains <u>three</u> roots, this gives us the possibilities:

$$\text{For } p = 0 : \beta, \alpha + \beta, 2\alpha + \beta \ [q = 2]$$
$$\text{For } p = -1 : -\alpha + \beta, \beta, \alpha + \beta \ [q = 1]$$
$$\text{For } p = -2 : -2\alpha + \beta, -\alpha + \beta, \beta \ [q = 0]$$

since for $p = 1$ $[q = 2]$ and for $p = -3$ $[q = -1]$ we would have β disappear from the string. This means in this case that

$$-2\tfrac{\beta(H_\alpha)}{\alpha(H_\alpha)} = -\beta(H_\alpha) = p + q = 2, 0, -2$$

If the α-string containing β contains <u>two</u> roots, this gives us the possibilities:

$$\text{For } p = 0 : \beta, \alpha + \beta \ [q = 1]$$
$$\text{For } p = -1 : -\alpha + \beta, \beta \ [q = 0]$$

since for $p = 1$ $[q = 2]$ and for $p = -2$ $[q = -1]$, β would disappear from the string. This means that in this case

$$-2\tfrac{\beta(H_\alpha)}{\alpha(H_\alpha)} = -\beta(H_\alpha) = p + q = 1, -1$$

If the α-string containing β contains <u>one</u> root, this gives us the possibilities

$$\text{For } p = 0 : \beta[q = 0]$$

since for $p = 1$ and $q = -1$, β would disappear from the string. This means in that case that

$$-2\tfrac{\beta(H_\alpha)}{\alpha(H_\alpha)} = -\beta(H_\alpha) \text{ and } p + q = 0$$

There are two other ideas that help control the roots – that of a *simple* system of roots and that of an *indecomposable* system of roots. First we choose a basis of roots in $\hat{h}^* : (\beta_1, ..., \beta_l)$, where l is the dimension of the Cartan subalgebra [or rank of the Lie algebra]. Thus for any root α in $\hat{h}^*$ we have $\alpha = \sum_{i=1}^{n} a_i \beta_i$, where we know that the scalars a_i will be integers. We can now introduce an ordering on the set of roots by ordering $\hat{h}^*$. We say an element γ of $\hat{h}^*$ is *positive* if the first non-zero scalar c_i in this sum $\gamma = \sum_{i=1}^{n} c_i \beta_i$ is positive. The set of positive duals in $\hat{h}^*$ is closed under addition and by scalar multiplication by positive rationals. We can now order $\hat{h}^*$ by saying that $\sigma > \gamma$ if $\sigma - \gamma > 0$; and if $\sigma > \gamma$, then for any τ in $\hat{h}^*$ we have $\sigma + \tau > \gamma + \tau$; and if $c > 0$, then $c\sigma > c\gamma$; and if $c < 0$, then $c\sigma < c\gamma$.

Now we can call a root *simple* if α is positive and α cannot be written as the sum of two other positive roots. We then let Λ be the set of all *simple* roots relative to some fixed ordering of $\hat{h}^*$. Then we can make the following assertions:

(i) If α and β are in Λ and $\alpha \neq \beta$, then $\alpha - \beta$ in not a root.
(ii) If α and β are in Λ, and $\alpha \neq \beta$, then $\beta(H_\alpha) \leq 0$.
(iii) The set Λ is a basis for $\hat{h}^*$ over the rational field.
(iv) If α is any positive root, then $\alpha = \sum_{i=1}^{n} a_i\beta_i$ where the a_i are *non-negative integers*.
(v) If α is a positive root and α is not in Λ, then there exists a β in Λ such that $\alpha - \beta$ is a positive root.

We now write the simple system of roots $\Lambda = \{\beta_1, , , , \beta_l\}$ and we will call this the *simple system of roots* for the semisimple algebra $\hat{g}$ of rank l relative to the given ordering in $\hat{h}^*$. We know that this is a basis over the rational field for the linear space $\hat{h}^*$; but more than this, we know that every root α of $\hat{g}$ can be written as $\alpha = \sum_{i=1}^{n} a_i\beta_i$, where either the scalars a_i are positive integers or 0 [with at least one non-zero root] or all the scalars a_i are negative integers or 0 [with at least one non-zero root].

Now having fixed a Cartan subalgebra $\hat{h}$ of $\hat{g}$, a semisimple Lie algebra, and a simple system of roots, we can define the *Cartan matrix* of $\hat{g}$. The entries in this matrix are:

$$[A_{ij}] = [\tfrac{2(\alpha_i(H_{\alpha_j}))}{\alpha_i(H_{\alpha_j})}] = [\alpha_i(H_{\alpha_j})]$$

Each diagonal entry on the Cartan matrix $\alpha_i(H_{\alpha_i})$ is equal to 2. For the off-diagonal entries $\alpha_i(H_{\alpha_j})$, we know that they must be non-positive [see (ii) above], and also from above they can only be 0, -1, -2 or -3. Also, the determinant of the Cartan matrix is not zero. Now since α_i and α_j are basis vectors and thus independent, the angle θ_{ij} between them must satisfy $0 \leq cos^2(\theta_{ij}) < 1$. This is equivalent to saying that

$$0 \leq (\tfrac{2\alpha_i(H_{\alpha_j})}{\alpha_i(H_{\alpha_i})})(\tfrac{2\alpha_j(H_{\alpha_i})}{\alpha_j(H_{\alpha_j})}) < 4 \text{ or } \ 0 \leq \alpha_i(H_{\alpha_j})\alpha_j(H_{\alpha_i}) < 4 \text{ or } \ 0 \leq A_{ij}A_{ji} < 4$$

This, of course means that both A_{ij} and A_{ji} are 0 or that one is -1 and the other is -1, -2 or -3.

Everything that has been said above applies to semisimple Lie algebras. But we know that a semisimple Lie algebra is a direct sum of *simple* Lie algebras. To identify such a system of roots we need the concept of an *indecomposable* system of roots. We define it negatively as follows. An *indecomposable* system of roots $\Delta = \{\alpha_1, ...\alpha_l\}$ is a system of roots such that it is impossible to partition Δ into non-empty, non-overlapping sets Δ_1, Δ_2 such that $A_{ij} = 0$ for every α_i in Δ_1 and every α_j in Δ_2. And we can assert that $\hat{g}$ is *simple* if and only if the associated *simple* system of roots Δ is indecomposable.

We are now at the point where we can identify the simple Lie algebras $\hat{g}$. We choose a simple system of roots $\Lambda = \{\alpha_1, ..., \alpha_l\}$ where l is the rank of $\hat{g}$ [which now again may be semisimple and not necessarily simple] or the dimension of the Cartan subalgebra $\hat{h}$. Then each root space $\hat{g}_{\alpha_i}$ is one dimensional and it has a corresponding one-dimensional root space $\hat{g}_{-\alpha_i}$. Changing notation a little to conform with the standard notation, we choose a basis $(H_{\alpha_i}, ..., H_{\alpha_l})$ for $\hat{h}$ such that $\alpha_i(H_{\alpha_i}) = 2$, a basis of E_{α_i}'s for $\hat{g}_\alpha$, and a basis of $F_{-\alpha_i}$'s for $\hat{g}_{-\alpha}$, such that $[E_{\alpha_i}, F_{-\alpha_i}] = H_{\alpha_i}$. As we have said above, this does determine H_{α_i} but it does not uniquely determine E_{α_i} or $F_{-\alpha_i}$. With these choices we have

$$[E_{\alpha_i}, F_{-\alpha_j}] = 0$$

and

$$\begin{gathered}
[E_{\alpha_i}, F_{-\alpha_i}] = ad(H_{\alpha_i})(E_{\alpha_i}) = \alpha_i(H_{\alpha_i})E_{\alpha_i} = 2E_{\alpha_i} \\
[H_{\alpha_i}, F_{-\alpha_i}] = ad(H_{\alpha_i})(F_{-\alpha_i}) = -\alpha_i(H_{\alpha_i})F_{-\alpha_i} = 2F_{-\alpha_i} \\
[E_{\alpha_i}, F_{-\alpha_i}] = H_{\alpha_i}
\end{gathered}$$

Since $[E_{\alpha_i}, F_{-\alpha_j}]$, $i \neq j$ is in $\hat{g}_{\alpha_i - \alpha_j}$ and $\alpha_i - \alpha_j$ is not a root, we have

$$[E_{\alpha_i}, F_{-\alpha_j}] = 0$$

and

$$\begin{gathered}
[H_{\alpha_i}, H_{\alpha_j}] = 0 \\
[H_{\alpha_i}, E_{\alpha_j}] = ad(H_{\alpha_i})(E_{\alpha_j}) = \alpha_j(H_{\alpha_i})E_{\alpha_j} = A_{ij}E_{\alpha_j} \\
[H_{\alpha_i}, F_{-\alpha_j}] = ad(H_{\alpha_i})(F_{-\alpha_j}) = -\alpha_j(H_{\alpha_i})F_{-\alpha_j} = -A_{ji}F_{-\alpha_j} \\
[H_{\alpha_j}, E_{\alpha_i}] = ad(H_{\alpha_j})(E_{\alpha_i}) = \alpha_i(H_{\alpha_j})E_{\alpha_j} = A_{ij}E_{\alpha_i}
\end{gathered}$$

With this notation in place we can now assert the following. We have chosen a simple system of roots $\Lambda = \{\alpha_1, ..., \alpha_l\}$. This system determines the triples $H_{\alpha_i}, E_{\alpha_i}$ and $F_{-\alpha_i}$.Then the $3l$ elements of $\hat{g}$, namely, $H_{\alpha i}$,E_{α_i} and $F_{-\alpha_i}$, generate $\hat{g}$. Now for each positive root β [not necessarily simple] we can select a representation of $\beta = \alpha_{i1} + ... + \alpha_{im}$ so that $\alpha_{i1} + ... + \alpha_{ik}$ is also a root for all $k \leq m$. Now we know that these sequences can be determined from the Cartan matrix $[A_{ij}]$. Then the elements

H_{α_i} and $[...[[E_{\alpha_{i1}}, E_{\alpha_{i2}}], E_{\alpha_{i3}}, ..., E_{\alpha_{im}}]$ and $[...[[F_{-\alpha_{i1}}, F_{-\alpha_{i2}}], F_{-\alpha_{i3}}, .., .F_{-\alpha_{im}}]$

determined by β form a basis for $\hat{g}$ and the multiplication table for this basis has rational coefficients determined by the Cartan matrix $[A_{ij}]$.

And thus if we want to identify the *simple* Lie algebras, all we have to do is choose an *indecomposable* simple system of roots.

We pause here to give three low dimensional examples. The basic building blocks of the simple complex Lie algebra is the algebra a_1, the $2x2$ matrices of trace zero in $\hat{sl}(2,\mathbb{C})$. The subscript 1 on a_1 says that the Cartan subalgebra of a_1 is one-dimensional.] We have already exposed this algebra above, giving the basis:

$$H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} E = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \text{ and } F = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

with the multiplication table:

$$[H, E] = \alpha(H)E = 2E \text{ and } [H, F] = -\alpha(H)F = -2F \text{ and } [E, F] = H$$

We now show how what was discussed above ideas apply to $\hat{sl}(2,\mathbb{C})$. We see that the Cartan subalgebra $\hat{h}$ of $\hat{sl}(2,\mathbb{C})$ is one-dimensional with basis H, and thus $\hat{sl}(2,\mathbb{C})$ has rank one. Dualizing, we let β be a basis for $\hat{h}^*$ such that $\beta(H) = 1$. Thus for the two roots of $\hat{sl}(2,\mathbb{C})$ we have $\alpha = 2\beta$ and $-\alpha = -2\beta$. And we see that α is a positive root. We also see that α is simple, since it is the only root, that is, the set of positive roots Λ consists of one element α. Now the Cartan matrix for $\hat{sl}(2,\mathbb{C})$ is a 1x1 matrix $[A_{11}] = [\alpha(H)] = [2]$. We also see that the simple system of roots Λ is also indecomposable, for there is no way that we can partition a set of only one member. Thus Λ is indeed a Δ. This says that $\hat{sl}(2,\mathbb{C})$ is not only semisimple but indeed is simple.

The second low dimensional complex Lie algebra that we examine is the algebra d_2. These are the 4x4 skew symmetric matrices $\hat{so}(4,\mathbb{C})$. [Again the subscript 2 of d_2 says that the Cartan subalgebra of d_2 has dimension 2.] Its elements have the form:

$$\begin{bmatrix} 0 & -a & -b & -c \\ a & 0 & -d & -e \\ b & d & 0 & -f \\ c & e & f & 0 \end{bmatrix}$$

We see that it is six-dimensional. We choose the following matrices as it basis:

$$H_1 = \begin{bmatrix} 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{bmatrix} H_2 = \begin{bmatrix} 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \end{bmatrix}$$

$$E_1 = \begin{bmatrix} 0 & 0 & -1/2 & -i/2 \\ 0 & 0 & i/2 & 1/2 \\ 1/2 & -i/2 & 0 & 0 \\ i/2 & -1/2 & 0 & 0 \end{bmatrix} E_2 = \begin{bmatrix} 0 & 0 & 1/2 & -i/2 \\ 0 & 0 & -i/2 & -1/2 \\ -1/2 & i/2 & 0 & 0 \\ i/2 & 1/2 & 0 & 0 \end{bmatrix}$$

$$F_1 = \begin{bmatrix} 0 & 0 & -1 & -i \\ 0 & 0 & -i & -1 \\ 1 & i & 0 & 0 \\ -i & 1 & 0 & 0 \end{bmatrix} F_2 = \begin{bmatrix} 0 & 0 & -1 & -i \\ 0 & 0 & -i & 1 \\ 1 & i & 0 & 0 \\ i & -1 & 0 & 0 \end{bmatrix}$$

We calculate the brackets:

$$[H_1, H_2] = ad(H_1)(H_2) = 0$$

$$[H_1, E_1] = ad(H_1)(E_1) = \alpha_1(H_1)(E_1) = 2E_1$$
$$[H_1, F_1] = ad(H_1)(F_1) = -\alpha_1(H_1)(F_1) = -2F_1$$
$$[H_1, E_2] = ad(H_1)(E_2) = \alpha_2(H_1)(E_2) = 0$$
$$[H_1, F_2] = ad(H_1)(F_2) = -\alpha_2(H_1)(F_2) = 0$$

$$[H_2, E_1] = ad(H_2)(E_1) = \alpha_1(H_2)(E_1) = 0$$
$$[H_2, F_1] = ad(H_2)(F_1) = -\alpha_1(H_2)(F_1) = 0$$
$$[H_2, E_2] = ad(H_2)(E_2) = \alpha_2(H_2)(E_2) = 2E_2$$
$$[H_2, F_2] = ad(H_2)(F_2) = -\alpha_2(H_2)(F_2) = -2F_2$$

$$[E_1, E_2] = ad(E_1)(E_2) = 0$$
$$[E_1, F_1] = ad(E_1)(F_1) = H_1$$
$$[E_1, F_2] = ad(E_1)(F_1) = 0$$
$$[E_2, F_1] = ad(E_2)(F_1) = 0$$
$$[E_2, F_2] = ad(E_2)(F_2) = H_2$$
$$[F_1, F_2] = ad(F_1)(F_2) = 0$$

This says that the rank of $\hat{so}(4,\mathbb{C})$ is two, the dimension of the Cartan subalgebra with basis (H_1, H_2). Correspondingly we choose a dual basis for $\hat{h}^* : (\beta,\beta_2)$ dual to (H_1, H_2). We write the roots with respect to this dual basis. We know that

$$\alpha_1(H_1) = 2; \alpha_1(H_2) = 0 \text{ and } \alpha_2(H_1) = 0;\ \alpha_2(H_2) = 2$$

Thus

$$\alpha_1 = 2\beta_1 + 0\beta_2 \text{ and } \alpha_2 = 0\beta_1 + 2\beta_2$$

since

$$\alpha_1(H_1) = 2 \text{ and } (2\beta_1 + 0\beta_2)H_1 = (2\beta_1)H_1 + 0(\beta_2)H_1 = 2(1) + 0 = 2$$
$$\alpha_1(H_2) = 0 \text{ and } (2\beta_1 + 0\beta_2)H_2 = (2\beta_1)H_2 + 0(\beta_2)H_2 = 0 + 0(1) = 0$$
$$\alpha_2(H_1) = 0 \text{ and } (0\beta_1 + 2\beta_2)H_1 = (0\beta_1)H_1 + 2(\beta_2)H_2 = 0(1) + 0 = 0$$
$$\alpha_2(H_2) = 2 \text{ and } (0\beta_1 + 2\beta_2)H_2 = (0\beta_1)H_2 + 2(\beta_2)H_2 = 0(1) + 2(1) = 2$$

We choose $\beta_1.\beta_2$ as a basis corresponding to the roots. We see that with respect to this basis both α_1 and α_2 are both positive and simple [since they cannot be written as a sum of two other positive roots]. Finally, we know that this set of positive and simple roots is a basis for $\hat{h}^*$.

We now compute the Cartan matrix:

$$A_{11} = \alpha_1(H1) = 2$$
$$A_{21} = \alpha_1(H2) = 0$$
$$A_{12} = \alpha_2(H1) = 0$$
$$A_{22} = \alpha_2(H2) = 2$$

and thus the 2 x 2 Cartan matrix is

$$[A_{ij}] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

We now show how the ideas given above apply to $\hat{so}(4,\mathbb{C})$. Thus, for the 4 roots of $\hat{so}(4,\mathbb{C})$, written on the dual basis (β_1, β_2), we have

$$\alpha_1 = 2\beta_1 + 0\beta_2 \text{ or } -\alpha_1 = -2\beta_1 + 0\beta_2$$
$$\alpha_2 = 0\beta_1 + 2\beta_2 \text{ or } -\alpha_2 = 0\beta_1 - 2\beta_2$$

Thus, we see that α_1 and α_2 are positive roots. We also see that α_1 and α_2 are simple roots since neither can be written as a sum of two other positive roots. Thus the set of positive roots Λ is the set $\{\alpha_1, \alpha_2\}$. We observe that $\alpha_1 \neq \alpha_2$ and thus $\alpha_1 - \alpha_2$ is a root; and that $-\alpha_2(H_{\alpha 1}) \leq 0$. In fact it is equal to 0. And, of course, the set Λ is decomposable. And thus we can conclude that $\hat{so}(4,\mathbb{C})$ is semisimple [we can show that it has no radical] but not simple. In fact, from the Cartan matrix, we see that is it just two copies of $\hat{sl}(2,\mathbb{C})$.

This concludes our discussion of Cartan subalgebras. The reader is strongly encouraged to pursue this topic in greater depth in [K] or [FH].