

Lecture Notes on Lie Algebras

Allen Bell

These notes are incomplete and still changing.

A version of the notes should be available at sites.uwm.edu/adbell

L^AT_EXed May 4, 2026

Contents

Chapter 0. Introduction	7
1. Why Lie Algebras?	7
1.1. Classifications	8
1.2. Representations	10
2. Outline of Material	11
3. Conventions, Notation, and Quirks	12
3.1. Conventions	12
3.2. Old German letters	13
3.3. List of notation	13
3.4. Quirks	13
4. List of notable propositions	14
Chapter 1. Desirable Forms for Matrices and Operators	17
1. Single operators	19
1.1. Diagonalizability and triangularizability	19
1.2. Eigenspace decompositions	21
2. The Jordan-Chevalley Decomposition	22
2.1. The additive Jordan-Chevalley decomposition	22
2.2. The multiplicative Jordan-Chevalley decomposition	25
3. Sets of Operators	25
3.1. General remarks	26
3.2. Diagonalizability	28
3.3. Triangularizability	28
4. Other Constructions	34
4.1. Bilinear forms and the trace form	34
4.2. Extension of scalars	36
4.3. The matrix exponential	37
4.4. Lagrange interpolation and the Chinese remainder theorem	38
Chapter 2. Basics of Lie Algebras	39
1. Definition and Examples	41

1.1.	The definition of a Lie algebra	41
1.2.	Examples of linear Lie algebras	43
1.3.	Lie algebras of low dimension	45
2.	Morphisms and Subobjects	47
2.1.	Homomorphisms and Isomorphisms	47
2.2.	Subalgebras and Ideals	48
2.3.	Derivations	49
3.	Types of Lie Algebras	52
3.1.	Solvable and Nilpotent Lie algebras	52
3.2.	Examples of derived and central series	55
3.3.	The radical	56
3.4.	Semisimple and reductive Lie algebras	59
4.	Adjoint	60
4.1.	Basics of adjoints	60
4.2.	Engel's theorem	63
5.	Characteristic Subspaces	64
6.	Cartan's Criterion for solvability	66
7.	Invariant Forms	70
7.1.	The Killing form	71
7.2.	Cartan's criterion for semisimplicity	74
7.3.	The structure of semisimple and reductive Lie algebras	77
8.	Extension of Scalars	80
Chapter 3. Representations		83
1.	Definitions and Basics	84
1.1.	Basics of modules and representations	84
1.2.	Two examples	87
1.3.	The Jordan-Hölder and Krull-Schmidt theorems	89
1.4.	Completely reducible representations	90
1.5.	New modules from old	93
1.6.	Semidirect products	95
2.	Lie's Theorem	97
2.1.	Irreducible modules over solvable Lie algebras	97
2.2.	Solvability and triangularizability	100
2.3.	Solvability and nilpotence	102
3.	Representations of $\mathfrak{sl}(2)$	103
3.1.	Concrete construction of irreducible representations	103

3.2.	Abstract construction of irreducible representations	104
3.3.	Vista: infinite dimensional irreducible representations	106
3.4.	Generating irreducible representations from the natural representation	107
4.	Representations of $\mathfrak{sl}(3)$	107
4.1.	Roots, weights, and representations for $\mathfrak{sl}(3)$	108
4.2.	Highest weight vectors for $\mathfrak{sl}(3)$	111
4.3.	Dimensions of irreducible representations	114
5.	Weyl's Theorem on Complete Reducibility	114
5.1.	The Casimir element and operator	115
5.2.	Weyl's theorem	117
5.3.	Abstract Jordan-Chevalley decomposition	119
5.4.	Reductive algebras	121
Chapter 4.	Semisimple Lie algebras and Root Space Decomposition	125
1.	Toral, Cartan, and Borel subalgebras	125
2.	Roots and Root Spaces	128
2.1.	Root space decomposition	129
2.2.	Root space decomposition for $\mathfrak{sl}(n)$ and $\mathfrak{gl}(n)$	132
2.3.	Root space decomposition for semisimple Lie algebras	134
2.4.	The dual Killing form	138
Chapter 5.	Root Systems	141
1.	Coxeter Groups and Root Systems	142
2.	Two-dimensional Root Systems	145
3.	Positive Roots and Bases	147
4.	Coxeter Graphs and Matrices	151
5.	The Crystallographic Condition	157
Chapter 6.	Classification of Semisimple Lie Algebras	159
1.	Objects associated to Semisimple Lie Algebras	159
2.	Classification of Split Semisimple Lie Algebras	164
Chapter 7.	More Representation Theory	165
1.	Basics of Weights	165
2.	The Theorem of the Highest Weight	168
3.	Representations of $\mathfrak{sl}(n)$	169
4.	The Weyl Character Formula	170
Chapter 8.	Miscellaneous Calculations and Observations	173
1.	Squarefree Polynomials and the Derivative	173

2. Multiplication and Commutator of Triangular Matrices	174
3. Matrices for the Orthogonal and Symplectic Groups	175
4. The a, b, c Lie Algebras	176
4.1. The four cases and the derived subalgebra	176
4.2. Change of variables and isomorphisms	177
4.3. The adjoint representation and linear realizations	179
4.4. The Killing form and trace form	179
4.5. The Jordan-Chevalley decomposition	180
4.6. The Casimir element	181
5. Killing and Trace Form Computations	181
5.1. Forms for $\mathfrak{sl}(2)$	181
5.2. Examples where the Killing form is zero	181
6. Bases of L and $\text{ad } L$	183
7. More on Bilinear Forms	183
7.1. Degenerate forms	183
7.2. Invariant forms on some algebras	184
8. Derivations of the 2-dimensional and the Heisenberg Lie Algebras	184
9. Derivation Formulas	184
9.1. Powers of a derivation on a commutator	185
9.2. Powers of a derivation minus scalars on a commutator	185
9.3. Powers of ad in $\mathfrak{gl}(V)$	185
10. Jordan-Chevalley Computations	186
11. Lie algebra Plus its Dual	188
12. The Transporter	189
13. $\mathfrak{sl}(2)$ Representations	191
13.1. Small dimension $\mathfrak{sl}(2)$ representations	191
13.2. E, F, H relations	192
13.3. Irreducibility of V_n	193
14. Dynkin Diagrams and Coxeter Graphs	193
14.1. Dynkin and Coxeter diagrams	194
14.2. The $\text{\LaTeX}$ code	202
15. Root System Pictures	206
Bibliography	211
Index of Definitions	213
Index of Notation and Symbols	217

CHAPTER 0

Introduction

1. Why Lie Algebras?

We take it for granted that the reader believes Representation Theory (whatever that means!) is both important and interesting. Most of us first see representation theory as the theory of finite groups acting on sets. For us, though, representations will always be linear, and so the corresponding notion is the theory of finite groups acting on vector spaces. This is a beautiful theory in characteristic 0, and we will restrict ourselves in these notes to characteristic 0. Over $\mathbb{C}$, we can use characters as a basis for the theory, and we can immediately put restrictions on the number and dimensions of irreducible representations. Below we compare some results from the theory of representations of finite groups with the theory of representations of finite-dimensional Lie algebras.

When we go beyond finite groups, representation theory becomes vastly more complicated. We usually restrict ourselves to some nice classes of groups: matrix groups (subgroups of $\mathbf{GL}(n, \mathbb{R})$ or $\mathbf{GL}(n, \mathbb{C})$), Lie groups, or algebraic groups. Representations of such groups are much more complicated than those of finite groups, but they arise in many areas of mathematics and applications such as physics, so it is worthwhile to try to understand them. We will not cover this part of the representation theory story, at least not directly.

One of Sophus Lie's ideas was to consider the invariant vector fields (infinitesimal generators) of Lie groups and so turn the study of these representations into (still quite complicated) linear algebra. In the translation from group to algebra, the multiplicative commutator $ABA^{-1}B^{-1}$ in the group corresponds to the additive commutator $AB - BA$ in the Lie algebra. The map from the Lie algebra to the Lie group is the exponential map, which can be defined concretely or abstractly. Kaplansky ([**Saaty**, p. 117]) lists these results for matrices A, B that show the connection. (He calls them glorified calculus exercises.)

$$\begin{aligned}\lim_{n \rightarrow \infty} \left(e^{tA/n} e^{tB/n} \right)^n &= e^{t(A+B)} \\ \lim_{n \rightarrow \infty} \left(e^{tA/n} e^{tB/n} e^{-tA/n} e^{-tB/n} \right)^{n^2} &= e^{t(AB-BA)}\end{aligned}$$

The reader can find many sources describing the correspondence. Among our references, see [**Bourbaki 1-3**, Chapter III; **Fulton-Harris**, Lecture 8; **Hall**, Chapter 2; **Knapp**, §I.10].

The purpose of this introduction is to pique the reader's interest in the study of Lie algebras. For most, the connection between Lie groups and Lie algebras briefly noted above is the strongest motivation to study the theory of Lie algebras. However, we will give two other motivations to hopefully stimulate the reader's interest.

First, we give two quotes from Irving Kaplansky, from [**Saaty**, pp. 115 & 122].

A distillation of the research of the last two decades appears to indicate that there are three classes of rings worthy of close study: alternative (a slight generalization of associative), Jordan, and Lie.

The tools employed are not deep; they are just, so to speak, linear algebra raised to the n th power. The apprentice algebraist who wishes to see the finer points of linear algebra put to work is strongly urged to work through the theory of Lie algebras. He will find everything he knows used in a dazzling array of spectacular arguments.

We will concentrate in the rest of this introduction on a comparison between the theory of finite groups and the theory of finite-dimensional Lie algebras. We hope the reader will find that the similarities and differences between the two theories make the study of Lie algebras appealing.

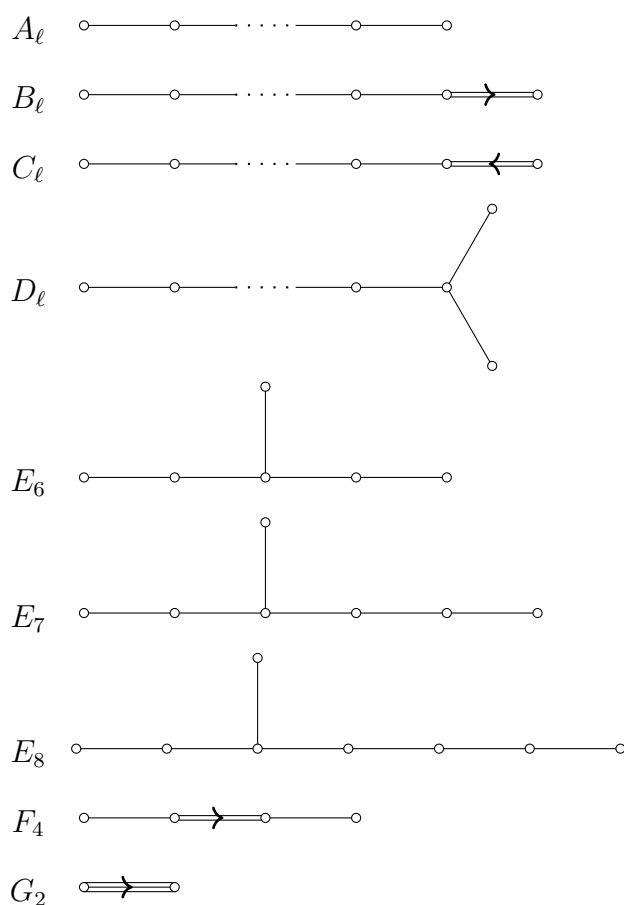
1.1. Classifications. Let's compare the theory of finite groups to the theory of finite-dimensional Lie algebras. The structure of finite Abelian groups is the subject of a well-known theorem. The theory of finite-dimensional Abelian Lie algebras is identical to the theory of finite-dimensional vector spaces, and all such Lie algebras are direct sums of 1-dimensional subalgebras. However, an Abelian subalgebra of a Lie algebra can be extremely important.

We define nilpotent and solvable Lie algebras exactly as we define nilpotent and solvable groups. A finite-dimensional Lie algebra contains a largest nilpotent ideal and a largest solvable ideal (but not a largest Abelian ideal), and these are called the nil radical and (solvable) radical respectively. Modulo its solvable radical, a finite-dimensional Lie algebra is semisimple, or a direct sum of simple Lie algebras.

Thus we are led to the question of classifying simple Lie algebras, just as we study simple groups. The following table contrasts the cases of finite groups and finite-dimensional Lie algebras. We will say more after the table.

Classification of	Finite groups	Matrix Lie Algebraic groups	finite-dimensional Lie algebras
Abelian simple	direct sums of cyclic of prime power 17 families plus 26 sporadic		direct sums of 1-dimensional 4 (or 3) families plus 5 sporadic

Let's dig into the classification of simple Lie algebras and simple groups (everything finite-dimensional/finite). The classification of simple Lie algebras is as follows. Every simple (finite-dimensional) Lie algebra is described uniquely by one of the diagrams below. These Dynkin diagrams picture the simple roots of the Lie algebra. In the table, ℓ is a positive integer; to ensure simplicity and uniqueness, we restrict ℓ as follows. For B_ℓ , we require $\ell \geq 2$; for C_ℓ , we require $\ell \geq 3$; for D_ℓ , we require $\ell \geq 4$. The algebras A_ℓ are the $(\ell + 1) \times (\ell + 1)$ matrices of trace 0, the algebras B_ℓ and D_ℓ are the skew-symmetric matrices of a size given by ℓ , and the algebras C_ℓ correspond to a symplectic form. The other 5 algebras are called exceptional (like “sporadic” in the finite group case).

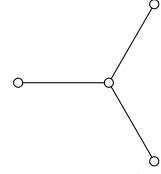


In 1955, Claude Chevalley (studying algebraic groups) invented a new way to create simple Lie algebras. His method could be applied to creating finite simple groups based over finite fields, groups we now refer to as “of Lie type”. In 1958, Robert Steinberg extended Chevalley’s method using automorphisms of the Dynkin diagram. Other efforts followed, some only possible over fields of order 2^n and 3^n . Thus we get to the current (complete) list of finite simple groups: the alternating groups, the 26 sporadic groups, and the following (q refers to the finite field $\mathbb{F}_q$ and n is a positive integer):

$$A_n(q), B_n(q)(n \geq 2), C_n(q)(n \geq 3), D_n(q)(n \geq 4), E_6(q), E_7(q), E_8(q), F_4(q), G_2(q), \\ {}^2A_n(q^2)(n \geq 2), {}^2D_n(q^2)(n \geq 4), {}^2E_6(q^2), {}^3D_4(q^3), {}^2B_2(2^{2n+1}), {}^2F_4(2^{2n+1}), {}^2F_4(2)', {}^2G_2(3^{2n+1})$$

If we examine these groups and the Dynkin diagrams, we see that the pre-superscripts match orders of automorphisms of the Dynkin diagrams. For example, A_ℓ has an automorphism of order 2 flipping the diagram about its center, as does E_6 , while D_ℓ has an

automorphism of order 2 exchanging the two endpoints. The diagram D_4



has a rotational automorphism of order 3, leading to a phenomenon known as triality.

1.2. Representations. One of the main reasons we study Lie algebras is to study their representations and thereby to study the representations of Lie groups, algebraic groups, and matrix groups. Let’s compare the theory of representations of finite groups and finite-dimensional Lie algebras (over $\mathbb{C}$, generally).

(Abbreviations in the tables: char for characteristic; rep for representation; f.d. for finite-dimensional; dim for dimension; diml for dimensional.)

Field	Finite group G	Matrix Lie Algebraic group	finite-dimensional Lie algebra L
char 0	all reps completely reducible		all f.d. reps completely reducible $\iff L$ semisimple
char 0	all irreducibles appear in regular rep		no f.d. rep contains all irreducibles
$\mathbb{C}$	# irreducibles = # conjugacy classes		# irreducibles = ∞
$\mathbb{C}$	dim of irreducible divides $ G/Z(G) $		dim of f.d. irreducibles unbounded if(f) L semisimple
$\mathbb{C}$	all irreducibles 1 diml $\iff G$ Abelian		all irreducibles 1 diml $\iff L$ solvable
char >0	not so nice		not so nice

2. Outline of Material

Much of the theory of Lie algebras is a vast generalization of linear algebra, so we start with some linear algebra. The set of $n \times n$ matrices over $\mathbb{F}$ is denoted either $M_n(\mathbb{F})$ or $\mathfrak{gl}(n, \mathbb{F})$ or usually just $\mathfrak{gl}(n)$; we use the latter notation when we are thinking of the set as a Lie algebra. We denote the set of all linear operators on the vector space V by $\text{End}(V)$ or $\mathfrak{gl}(V)$. Our base field $\mathbb{F}$ is always of characteristic 0 and sometimes algebraically closed.

There are two types of Lie algebras (though in fact they are the same). First are the linear ones, the subalgebras of $\mathfrak{gl}(V)$ with $[A, B] = AB - BA$ as their Lie bracket. Then come the abstract ones, defined as having a skew-symmetric product satisfying the Jacobi identity. A surprisingly difficult theorem of Ado says that all finite-dimensional Lie algebras are isomorphic to linear ones. Nevertheless, we treat them a bit differently.

Here is an example: Engel's theorem. One (concrete) version states that a linear Lie algebra consisting of nilpotent matrices is nilpotent as a Lie algebra. A second (abstract) version states that any Lie algebra for which $\text{ad } x$ is nilpotent for all x is nilpotent as a Lie algebra.

The map $\text{ad} : L \rightarrow \mathfrak{gl}(L)$ takes the abstract Lie algebra L to the concrete Lie algebra $\text{ad } L \subseteq \mathfrak{gl}(L)$, although it loses information about the center of the original L . This enables us to use linear algebra techniques as in Engel's theorem and is absolutely fundamental to the study of Lie algebras.

In Chapter 1 (Desirable Forms for Matrices and Operators, p. 17), we develop the necessary results on putting matrices/operators in simultaneous diagonal or upper triangular form. These results are then used in Engel's theorem, Cartan's criteria, and Lie's theorem. We also recall other material from linear algebra, such as the Jordan-Chevalley decomposition (a variation on the Jordan form) and bilinear forms, that will be used heavily. In subsequent chapters, we will apply these results.

In Chapter 2 (Basics of Lie Algebras, p. 39), we formally introduce Lie algebras and give a number of examples. As in the theory of finite groups, we have basic types of Lie algebras, such as Abelian, nilpotent, solvable, simple, and we define these types. Results such as Engel's theorem and Lie's theorem show that (over an algebraically closed field) nilpotent and solvable Lie algebras are isomorphic to subalgebras of the Lie algebra of upper triangular matrices.

We also define semisimple and reductive Lie algebras, which are direct sums of simple Lie algebras plus Abelian Lie algebras. One new concept in the theory of Lie algebras is that of a derivation, which frequently replaces the use of automorphisms in group theory. For example, a subgroup is called "characteristic" if it is invariant under every

automorphism of the group. The same concept is available for Lie algebras, but it is much more common to call a subalgebra characteristic if it is invariant under every derivation of the Lie algebra. We also introduce the adjoint mapping, an inner derivation, as mentioned above, and prove the second version of Engel's theorem. We introduce the concept of an invariant (also called associative) bilinear form, as well as the main example, the Killing form. This enables us to prove more "triangularization theorems" and to give Cartan's criterion for semisimplicity.

In Chapter 3 (Representations, p. 83), we begin our study of representation theory. We use representations, particular $\mathfrak{ad}$, to prove Lie's theorem and Cartan's criterion for solvability. We then completely describe the representation theory of $\mathfrak{sl}(2)$ and (still waiting) $\mathfrak{sl}(3)$. We finish by proving Weyl's theorem that all (finite-dimensional) representations of a semisimple Lie algebra are completely reducible and we draw some corollaries. To do this, we introduce the Casimir element of a semisimple Lie algebra.

In Chapter 4 (Semisimple Lie algebras and Root Space Decomposition, p. 125), we describe the structure of semisimple Lie algebras via root spaces and prove some basic properties of the root space decomposition. This enables us to turn the study of semisimple Lie algebras into a geometric problem, the study of root systems. In Chapter 5 (Root Systems, p. 141), we define and classify root systems. This topic is central to the classification of simple Lie algebras, but it also is crucial to the study of Coxeter groups. In this chapter, we introduce Dynkin diagrams.

Chapter 6 (Classification of Semisimple Lie Algebras, p. 159) and Chapter 7 (More Representation Theory, p. 165) are only outlines at this point. This reflects the seminar, where we will run out of time and give just only a high-level view of the material. In Chapter 6 (Classification of Semisimple Lie Algebras, p. 159), we discuss the classification of simple Lie algebras, based on our work on root systems. In Chapter 7 (More Representation Theory, p. 165), we discuss the theory of weights and highest weight vectors and classify the irreducible representations of semisimple Lie algebras via dominant integral weights and Verma modules. This chapter was presented directly after Chapter 4 (Semisimple Lie algebras and Root Space Decomposition, p. 125) in the seminar. Because of this, it does not assume the reader has read Chapter 5 (Root Systems, p. 141) and so it repeats some definitions from earlier chapters.

3. Conventions, Notation, and Quirks

This section is very much a work in progress and hence quite incomplete.

3.1. Conventions. We work throughout over a field $\mathbb{F}$ of characteristic 0. We use $\overline{\mathbb{F}}$ to denote the algebraic closure of $\mathbb{F}$ and $\mathbb{E}$ to denote an extension field of $\mathbb{F}$. We warn

that many results may fail for Lie algebras over fields of positive characteristic. All vector spaces are assumed to be finite-dimensional unless otherwise stated.

We will use the word “operator” to refer to either a matrix or a linear operator. If A is an operator, we will generally use Av to indicate the action of A on the vector v (matrix multiplication if A is a matrix), though we may occasionally use $A(v)$ for clarity. When dealing with matrices, it is understood that the vector space V is $\mathbb{F}^n$.

When referring to references, we use the section symbol § to refer to any subdivision smaller than a chapter. Thus §III.4.5 is in chapter 3, probably section 4, and subsection, paragraph, etc., 5.

3.2. Old German letters. If you read older texts, such as [Jacobson], you will run into old German letters, known to L^AT_EX as `fraktur` or `mathfrak`. Here is a list at large size, followed by a table.

Ⓐ Ⓑ Ⓒ Ⓓ Ⓔ Ⓕ Ⓖ Ⓗ Ⓙ Ⓢ Ⓚ Ⓛ Ⓜ Ⓝ Ⓟ Ⓠ Ⓡ Ⓢ Ⓣ Ⓤ Ⓡ Ⓦ Ⓧ Ⓨ Ⓩ

ⓐ ⓑ ⓒ ⓓ ⓔ ⓕ ⓖ ⓗ ⓘ ⓙ ⓚ ⓛ ⓜ ⓝ ⓞ ⓟ ⓠ ⓡ ⓢ ⓣ ⓤ ⓥ ⓦ ⓧ ⓨ ⓩ

⓪ ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ Bonus: the Weierstrass ρ (rho) is ⓪

a	ⓐ	A	Ⓐ	b	ⓑ	B	Ⓑ	c	ⓒ	C	Ⓒ
d	ⓓ	D	Ⓓ	e	ⓔ	E	Ⓔ	f	ⓕ	F	Ⓕ
g	ⓖ	G	Ⓖ	h	ⓗ	H	Ⓗ	i	ⓘ	I	Ⓙ
j	ⓙ	J	Ⓢ	k	ⓚ	K	Ⓚ	l	Ⓛ	L	Ⓛ
m	Ⓜ	M	Ⓜ	n	Ⓝ	N	Ⓝ	o	ⓞ	O	Ⓟ
p	Ⓟ	P	Ⓠ	q	ⓠ	Q	Ⓡ	r	ⓣ	R	Ⓢ
s	Ⓢ	S	Ⓢ	t	ⓣ	T	Ⓣ	u	Ⓤ	U	Ⓡ
v	Ⓡ	V	Ⓡ	w	Ⓦ	W	Ⓦ	x	Ⓧ	X	Ⓧ
y	Ⓨ	Y	Ⓨ	z	Ⓩ	Z	Ⓩ				

3.3. List of notation.

There is an index of definitions and an index of notation and symbols at the end of these notes.

3.4. Quirks. Following long-ago advice of Knuth (who knows if he still thinks this way), quotation marks precede punctuation. Thus: We refer to such algebras as “associative”, and we call any other algebras “cray-cray”. Allegedly, this is called “logical punctuation” or “British punctuation”.

Though it is understood that algebras and vector spaces are finite-dimensional (unless the opposite is stated) and the base field is of characteristic 0, in the formal statement of propositions we explicitly list all hypotheses. This is because your author hates to have to hunt down the exact hypotheses for a result in a book or paper where those hypotheses are part of a convention stated “somewhere”. But perhaps you should not trust that the perfect hypotheses are stated in each of our propositions.

4. List of notable propositions

The statements are made compact and sometimes mildly altered.

1.2.1. Jordan-Chevalley decomposition, p. 23

Every operator is a sum of commuting semisimple and nilpotent operators.

1.3.4. Diagonalizability of a set of operators, p. 28

A set of operators is simultaneously diagonalizable if and only if the operators commute pairwise and are individually diagonalizable.

1.3.5. Frobenius’s theorem on triangularizability, p. 29

- (1) A set of commuting operators is simultaneously triangularizable iff each operator is individually triangularizable.
- (2) A set of commuting operators is simultaneously strictly triangularizable iff each operator is nilpotent.

1.3.7. Generalized Jacobson’s theorem, p. 31

A set X of nilpotent operators is simultaneously triangularizable if for every $A, B \in X$ there is a scalar c with $AB - cBA \in X$.

1.3.8. Strong Engel’s theorem, p. 33

A set X of operators with the property that $A, B \in X$ implies $AB - BA \in X$ is simultaneously triangularizable iff each operator is individually triangularizable and $AB - BA$ is nilpotent for all $A, B \in X$.

2.1.1. Ado’s theorem, p. 43

Every finite-dimensional Lie algebra is isomorphic to a linear Lie algebra.

2.2.3. Jordan-Chevalley decomposition of derivations, p. 51

The semisimple and nilpotent parts in the Jordan-Chevalley decomposition of a derivation are themselves derivations.

2.4.3. Concrete Engel’s theorem, p. 63

If each element of a linear Lie algebra L is a nilpotent operator, then L is a nilpotent Lie algebra and L is simultaneously strictly triangularizable.

2.4.4. Engel's theorem for $\mathfrak{ad}$, p. 64

The Lie algebra L is nilpotent iff $\mathfrak{ad} x$ is a nilpotent operator for every $x \in L$.

2.7.6. Cartan's criterion for semisimplicity, p. 74

A Lie algebra L is semisimple iff the Killing form on L is nondegenerate.

2.7.11. Semisimple is equivalent to sum of simples, p. 77

A Lie algebra is semisimple iff it is a direct sum of simple Lie algebras.

2.7.13. All derivations are inner, p. 78

All derivations on a semisimple Lie algebra are inner.

3.5.5. Reductive algebras as direct sums, p. 121

A reductive Lie algebra is the direct sum of its derived subalgebra and its center.

3.1.13. Levi decomposition, p. 96

If R is the solvable radical of L , then there is a semisimple subalgebra K of L with $L = R \oplus K$. Moreover, any two such K are conjugate via a special automorphism of L .

3.2.2. Irreducibles are 1-dimensional for solvable Lie algebras, p. 98

If L is solvable, any split irreducible L -module has dimension 1.

3.2.6. Concrete Lie's theorem, p. 100

If L is a split linear Lie algebra, then L is solvable iff L is simultaneously triangularizable.

3.2.7. Lie's theorem on complete solvability, p. 101

If L is solvable and $\mathbb{F}$ is algebraically closed, then there is a full flag in L of ideals.

3.2.8. Solvable is close to nilpotent, p. 102

If L is solvable, then $[L, L]$ is nilpotent.

3.2.9. Derivations take solvable to nilpotent, p. 103

If N is the nil radical and R the solvable radical and δ is a derivation of L , then $\delta(R) \subseteq N$.

2.6.2. Concrete Cartan's criterion for solvability, p. 67

If L is a linear Lie algebra, then L is solvable if either of the following hold.

- (1) $\operatorname{tr} x^2 = 0$ for all $x \in [L, L]$.
- (2) $[L, L]$ is contained in the radical of the concrete trace form.

2.6.3. Cartan's criterion for solvability, p. 69

If L is a Lie algebra, then L is solvable if either of the following hold.

- (1) $\operatorname{tr}(\mathfrak{ad} x)^2 = 0$ for all $x \in [L, L]$.
- (2) $[L, L]$ is contained in the radical of the Killing form.

3.3.1. Irreducible representations of $\mathfrak{sl}(2)$, p. 106

- (1) The Lie algebra $\mathfrak{sl}(2)$ has exactly one irreducible module of each positive dimension.

- (2) Any irreducible module has a basis of h -eigenvectors and each h -eigenspace has dimension 1.
- (3) The h -eigenvalues of an irreducible module form an arithmetic progression with difference 2, and the dimension of the module is 1 plus the largest h -eigenvalue.

3.4.1. Irreducible representations of $\mathfrak{sl}(3)$, p. 113

- (1) Irreducible $\mathfrak{sl}(3)$ -modules are generated by (unique) highest weight vectors and are the direct sum of weight spaces.
- (2) There is an irreducible module $\Gamma_{a,b}$ of dimension $(a+1)(b+1)(a+b+2)$ for every pair of nonnegative integers a, b . Moreover, $\Gamma_{a,b} \cong \Gamma_{a',b'}$ if and only if $a = a'$ and $b = b'$.
- (3) Every finite-dimensional irreducible L -module is isomorphic to $\Gamma_{a,b}$ for some a, b .

3.5.3. Weyl's theorem on complete reducibility, p. 118

If L is semisimple, then every (finite-dimensional) representation of L is completely reducible.

3.5.4. Abstract Jordan-Chevalley decomposition, p. 120

If L is semisimple and $x \in L$, then there exist $x_s, x_n \in L$ such that $\text{ad } x_s, \text{ad } x_n$ are the semisimple and nilpotent parts in the Jordan-Chevalley decomposition of $\text{ad } x$ in $\mathfrak{gl}(L)$.

4.2.1. Root space decomposition, p. 129

Let H be a split toral subalgebra of L and let L_α be the weight space for $\alpha \in H^*$.

- (1) $L = \bigoplus_{\alpha \in H^*} L_\alpha$ and $[L_\alpha, L_\beta] \subseteq L_{\alpha+\beta}$ for all $\alpha, \beta \in H^*$.
- (2) If $\alpha, \beta \in H^*$ and $\beta \neq -\alpha^*$, then $\kappa(L_\alpha, L_\beta) = 0$.
- (3) Suppose L is semisimple. If $x \in L_\alpha$ and $\kappa(x, L_{-\alpha}) = 0$, then $x = 0$.

4.2.3. Maximal toral subalgebras self-normalize, p. 131

If L is semisimple and H is a maximal toral subalgebra of L , then $H = N(H)$ is its own normalizer.

4.2.5. Properties of the root space decomposition, p. 135

Let H be a split maximal toral subalgebra of L and let Φ be the set of roots of L relative to H . Let L_α be the weight space for $\alpha \in \Phi$. Many properties of L are listed, including:

- (2) Φ spans H^* .
- (5) For each $\alpha \in \Phi$, the subalgebra $L_\alpha \oplus [L_\alpha, L_{-\alpha}] \oplus L_{-\alpha}$ is an isomorphic copy of $\mathfrak{sl}(2)$.
- (9) If $\lambda \in \mathbb{F}$, $\alpha \in \Phi$, then $\lambda\alpha \in \Phi$ iff $\lambda = \pm 1$.
- (10) If $\alpha, \beta \in \Phi$, then the roots of the form $\beta + i\alpha$, $i \in \mathbb{Z}$, form a string with no gaps containing at most 4 elements.

CHAPTER 1

Desirable Forms for Matrices and Operators

The title of this chapter refers to “desirable form”. For us, that means one of three forms: diagonal, (upper) triangular, and strictly (upper) triangular. When we have a set X of operators, we are interested in simultaneously putting them in the desired form. When X consists of matrices, that means finding a single P with PAP^{-1} in the desired form for all $A \in X$. When X consists of linear transformations from V to V , that means finding a basis of V , such that all operators in X are in the desired form with respect to this basis.

We begin with a small example of what is to come. Set $A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Then A is already in diagonal form, but B is not. We can immediately read off the eigenvalues of A ($0, 1$) and its eigenvectors are the standard basis vectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. It’s not obvious what the eigenvalues of B are, although they are not hard to compute. Also if B were another diagonal matrix, we would immediately know $AB = BA$, but in our case, we don’t know if this equality holds. (If we’ve had some experience with matrices, we might realize that since A has distinct eigenvalues and B is not diagonal, we in fact have $AB \neq BA$.)

However, B is diagonalizable (with diagonal form $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$). We can see this either by noting that B has distinct eigenvectors $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ or by noting that the minimal (=characteristic in this case) polynomial of B is $x^2 - 1 = (x - 1)(x + 1)$.

After some discussion of single operators, we will turn to our main focus: can we simultaneously diagonalize A and B ? If not, can we least simultaneously put them in upper triangular form? (Simultaneously means is there a single P with both PAP^{-1}, PBP^{-1} in the desired form.) The answer to both questions is “no”. For if we could put A and B into upper triangular form with the same change of basis, the new basis would have to start with a common eigenvector for A and B . But the eigenvectors of A are (scalar multiples of) $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and the eigenvectors of B are (scalar multiples of) $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}$, and so there are no eigenvectors in common. We will be interested in conditions on A and B that guarantee that we can simultaneously diagonalize or upper triangularize them. If we keep our current A and allow ourselves another B , one can show that A, B can simultaneously be diagonalized if and only if B is already diagonal, and this in turn is equivalent to

the condition that $AB = BA$. This A and another B can simultaneously put into upper triangular form if and only if B is either upper triangular or lower triangular. This in turn is equivalent to the condition that $(AB - BA)^2 = 0$.

We call $AB - BA$ the (additive) **commutator** of A and B , and we denote it $[A, B]$. We will frequently be interested in products and commutators, so let's compute $[A, B]$ for this example. We have $AB = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ and $BA = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, neither of which is diagonalizable. Thus $[A, B] = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. This is an interesting matrix: it is diagonalizable if $\mathbb{F} = \mathbb{C}$, but it has no eigenvalues and no eigenvectors if $\mathbb{F} = \mathbb{R}$, so it is not even upper triangularizable in the real case. This hints that some of our results will depend on the base field (and in particular on whether it's algebraically closed).

In Section 1 (Single operators, p. 19), we review of what is known about diagonalizing and (upper) triangularizing single matrices. Before turning to sets of matrices, in Section 2 (The Jordan-Chevalley Decomposition, p. 22), we consider a decomposition $A = S + N$ for matrices that is related to the Jordan form, called the Jordan-Chevalley decomposition. Here S corresponds to the diagonal part of the Jordan form and N to the above diagonal part.

In Section 3 (Sets of Operators, p. 25), we turn to sets of operators. We also introduce some terminology to discuss the case where a matrix could be diagonalized if only its eigenvalues lay in our field of interest. When all the matrices commute (pairwise) and individually can be put in a given form, we will see that they can simultaneously put in the given form. If our goal is to put matrices in diagonal form, this result is an “iff”: we can only put them in simultaneous diagonal form if they commute. If our goal is to put matrices in a triangular form, we can relax the condition that they commute, but finding a necessary and sufficient condition is not easy. Instead, our goal is to find a sufficient conditions, and often this involves putting a condition on the commutators $AB - BA$ for A, B in the set. We will also see that nilpotent operators are easier to work with in some cases and that they are related to strict triangularizability. Some of our results in the triangularizable (but not strictly) case will have to wait for later chapters, when we have studied Lie algebras a little bit.

In Section 4 (Other Constructions, p. 34), we recall a number of topics from linear algebra that we will need. In particular, we discuss bilinear forms and we introduce the trace form, which leads to the Killing form and will be enormously useful in the study of Lie algebras.

1. Single operators

1.1. Diagonalizability and triangularizability. An operator A is diagonal if the current basis ($\mathbf{e}_1, \dots, \mathbf{e}_n$ for matrices) consists of A -eigenvectors. We say A is **diagonalizable** if there is *some* basis consisting of eigenvectors (equivalently, PAP^{-1} is diagonal for some invertible P). When working over fields that are not algebraically closed, we are also interested in operators that are **potentially diagonalizable**, that is, that are diagonalizable if we extend the base field $\mathbb{F}$ (e.g., to the algebraic closure $\overline{\mathbb{F}}$). We will call such operators **semisimple**, though not all sources use this term. (Some even call such operators diagonalizable.) **Warning:** The word “semisimple” has other meanings, even in Lie theory.

Consider the following examples: $A = \begin{pmatrix} 1 & 2 \\ 0 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $C = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Are these matrices diagonalizable? Are they semisimple? Can they be put in Jordan form?

Are there tests for diagonalizability? Associated to any operator A are the **minimal polynomial** $\text{mp}(A)$ and the **characteristic polynomial** $\text{cp}(A)$. One can show A is diagonalizable if and only if $\text{mp}(A)$ is a product of distinct linear factors (in $\mathbb{F}[x]$). (See, e.g., [Erdmann-Wildon, Theorem 16.4, p. 195].) It follows that A is semisimple if and only if its minimal polynomial is a product of distinct linear factors in $\overline{\mathbb{F}}[x]$. Can we give a condition for semisimplicity that doesn't require the algebraic closure? Recall that a polynomial is called **squarefree** if it is not divisible by any nonconstant square, or equivalently if it is not divisible by the square of an irreducible polynomial. A polynomial that factors completely into monic linear factors is squarefree if and only if the linear factors are distinct. Recall that a polynomial (over a field of characteristic 0) is squarefree if and only if $\gcd(f, f') = 1$, where f' is the formal derivative of f . (See Proposition 8.1.1 (A polynomial is squarefree iff it is relatively prime to its derivative, p. 173).) This last condition doesn't depend on the field we are working over. Thus we conclude that A is semisimple if and only if $\text{mp}(A)$ is squarefree. We can also give another characterization of diagonalizability: A is diagonalizable if and only if $\text{mp}(A)$ is squarefree and every one of its irreducible factors is linear (or: $\text{mp}(A)$ is squarefree and all the eigenvalues of A lie in $\mathbb{F}$).

Not every matrix is diagonalizable or semisimple, so we must also consider other forms. Over an algebraically closed field, every operator can be put in Jordan form. More generally, if all the eigenvalues of A lie in $\mathbb{F}$, i.e., if $\text{cp}(A)$ factors into powers of linear factors, then A can be put in Jordan form over the base field $\mathbb{F}$. Jordan form is too special a form to expect when we wish to put multiple operators into a nice form, so we will weaken it to upper triangular form. In addition, it's easier to prove a matrix with all

eigenvalues in $\mathbb{F}$ can be put into upper triangular form than into Jordan form. To see this, let v_1 be any eigenvector for the operator $A : V \rightarrow V$ with eigenvalue λ_1 and let W be the subspace of V spanned by v_1 (i.e., $W = Fv_1$). Let $\bar{V} = V/W$, and let $\bar{A}$ be the induced operator on $\bar{V}$ (i.e., $\bar{A}(v + W) = Av + W$). The eigenvalues of $\bar{A}$ are precisely the eigenvalues of A , possibly minus λ_1 , so they all lie in $\mathbb{F}$. By induction on $\dim V$, there is a basis of $\bar{V}$ with respect to which $\bar{A}$ has upper triangular form. We can take that basis to be $v_2 + W, \dots, v_n + W$ where $v_1, v_2, \dots, v_n$ is a basis of V . The matrix of A with respect to this basis is upper triangular.

We can give alternative conditions for an operator to be upper triangularizable, as we did for diagonalizability. Thus A is upper triangularizable if and only if $\text{mp}(A)$ factors into (not necessarily distinct) linear factors, and this in turn is equivalent to $\text{cp}(A)$ factoring into linear factors.

We now introduce some terminology to make it shorter to express ourselves. We say a polynomial is **split** if all of its roots lie in the base field $\mathbb{F}$, or equivalently, if it is a product of linear factors (in $\mathbb{F}[x]$). This is fairly standard terminology, but the following terminology — which we use throughout these notes — is **not standard**. We say a matrix or operator A is **split** if A is upper triangularizable. As we've noted above, this is equivalent to $\text{mp}(A)$ and $\text{cp}(A)$ being split, and to the condition that all the eigenvalues of A lie in $\mathbb{F}$. All polynomials and operators are split if $\mathbb{F}$ is algebraically closed.

We are also interested in operators that can be put into strictly upper triangular form. It's an easy exercise to see that for a strictly upper triangular A , we have $A^n = 0$ and sometimes $A^k = 0$ for $k < n$. An operator A with the property that $A^k = 0$ for some k is called **nilpotent**. The nice result here is that an operator is strictly upper triangularizable if and only if it is nilpotent. This is true if and only if its minimal polynomial is x^k for some k , and this is equivalent to the condition that the characteristic polynomial is x^n .

Before continuing, let us summarize some of the above discussion.

Proposition 1.1.1 (Operator forms and polynomials).

Let A be an operator on a finite-dimensional vector space over a field of characteristic 0. Then the following statements hold.

- (1) A is diagonalizable $\iff \text{mp}(A)$ is a product of distinct linear factors.
- (2) A is semisimple $\iff \text{mp}(A)$ is squarefree.
- (3) A is triangularizable $\iff \text{mp}(A)$ is split $\iff \text{cp}(A)$ is split.
- (4) A is strictly triangularizable $\iff A$ is nilpotent $\iff \text{mp}(A)$ is a power of $x \iff \text{cp}(A) = x^n$.

PROOF. See the above discussion. ■

Here is one more fact that will be used several times. We use $\text{tr } A$ to denote the trace of the matrix or operator A .

Proposition 1.1.2 (Nilpotency and trace).

Let A be an operator on an n -dimensional vector space over a field of characteristic 0. Then A is nilpotent if and only if $\text{tr } A^k = 0$ for $1 \leq k \leq n$.

PROOF. This is a well-known fact. We leave it to the reader to provide a proof or look it up. ■

The following observation is very useful.

Proposition 1.1.3 (Desirable forms on subspaces and quotients: single operators).

Let V be a vector space over a field of characteristic 0. Let $A : V \rightarrow V$ be a linear operator and let W be a A -invariant subspace. Let $A|_W$ denote the restriction of A to an operator on W and let $\bar{A}$ denote the induced operator on V/W (i.e., $\bar{A}(v + W) = A(v) + W$).

- (1) *If A is any one of diagonalizable, triangularizable, or strictly triangularizable, then so are $A|_W$ and $\bar{A}$.*
- (2) *A is triangularizable or strictly triangularizable if and only if both $A|_W$ and $\bar{A}$ are the same.*
- (3) *If $V = \bigoplus_{i=1}^k V_i$ and each V_i is A -invariant, then A is any one of diagonalizable, triangularizable, or strictly triangularizable if and only if $A|_{V_i}$ is the same for all i .*

PROOF. (1) If $f = \text{mp}(A)$, then $f(A|_W) = f(A)|_W = 0$, so $\text{mp}(A|_W)$ divides $\text{mp}(A)$. Likewise, $\text{mp}(\bar{A})$ divides $\text{mp}(A)$. Now apply Proposition 1.1.1 (Operator forms and polynomials, p. 20).

(2) There is a basis $v_1, \dots, v_k$ of W such that $A|_W$ is [strictly] triangularizable with respect to this basis, and there is a basis $u_1, \dots, u_\ell$ of $\bar{A}$ such that $\bar{A}$ is [strictly] triangularizable with respect to this basis. If $u_i = v_{k+i}$ for some $v_{k+i} \in V$, then $v_1, \dots, v_k, v_{k+1}, \dots, v_{k+\ell}$ is a basis of V and A is [strictly] triangular with respect to this basis.

(3) We leave this to the reader. ■

1.2. Eigenspace decompositions. Let A be an operator on V and let $\lambda \in \mathbb{F}$. We define $V_\lambda = \{v \in V \mid Av = \lambda v\}$ and $V_{[\lambda]} = \{v \in V \mid (A - \lambda I)^k v = 0 \text{ for some } k\}$. Both of these are subspaces of V that are invariant under A . We call them, respectively, the **λ -eigenspace** and **generalized λ -eigenspace** for A . Some sources require that λ be an eigenvalue for A for these to qualify as eigenspaces; if λ is not an eigenvalue, both spaces are 0.

The notation may vary with sources; for example, some use V_λ for the generalized eigenspace. The terminology may also vary. The generalized eigenspaces are also called **primary components** and **weight spaces**, although the latter term is usually reserved for a more general setting. [Bourbaki 7-9, §VII.1.1, p. 1] uses V^λ for the generalized eigenspace and calls it the primary subspace relative to λ .

Suppose the minimal polynomial of A is $\prod_{j=1}^k (x - \lambda_j)^{n_j}$ for distinct $\lambda_1, \dots, \lambda_k$. Then one can show (exercise!) that if $\dim V = n$, then $V_{[\lambda_j]} = \{v \in V \mid (A - \lambda_j I)^{n_j} v = 0\} = \{v \in V \mid (A - \lambda_j I)^n v = 0\}$.

Proposition 1.1.4 (Eigenspace and generalized eigenspace decomposition).

Let V be an n -dimensional vector space and let A be an operator on V . The sums below may be taken over all $\lambda \in F$ or just over the eigenvalues λ of A .

- (1) *A is diagonalizable if and only if $V = \bigoplus_\lambda V_\lambda$.*
- (2) *A is split if and only if $V = \bigoplus_\lambda V_{[\lambda]}$.*

PROOF. This is standard. **Provide a proof?** ■

Proposition 1.1.5 (Consequences of distinct eigenvalues).

Let V be a vector space and let A be an operator on V . Let $v_1, \dots, v_n \in V$ be eigenvectors for A with corresponding eigenvalues $\lambda_1, \dots, \lambda_n$. Suppose $\lambda_1, \dots, \lambda_n$ are distinct.

- (1) *$v_1, \dots, v_n$ are linearly independent.*
- (2) *If $v_1, \dots, v_n$ span V , then every nonzero A -invariant subspace of V contains one of the v_i .*

REFERENCES AND REMARKS. The results in (1) and (2) are valid even for infinite collections of v_i .

PROOF. (1) This is standard. You can also imitate the proof of (2).

(2) Let W be a nonzero invariant subspace of V and suppose v is a nonzero element of W of the form $\sum_{i=1}^n \alpha_i v_i$ with the fewest number of nonzero coefficients α_i . If only one α_i is nonzero, then multiplying by a scalar gives us $v_i \in W$. Thus we can suppose at least two α_i are nonzero. We can renumber and assume $\alpha_1 \neq 0 \neq \alpha_n$.

Consider $w = \lambda_n v - h(v) \in W$. We see that $w = \sum_{i=1}^{n-1} \alpha_i (\lambda_n - \lambda_i) v_i$. This is not 0, since $\alpha_1 (\lambda_n - \lambda_1) \neq 0$, but it has fewer nonzero coefficients than v . This contradicts our choice of v , and so we conclude that the minimal number of nonzero coefficients is one. Thus W contains some v_i . ■

2. The Jordan-Chevalley Decomposition

2.1. The additive Jordan-Chevalley decomposition. Consider a matrix A in Jordan form. If we let S be the diagonal part of A and N the upper triangular part, we

see that $A = S + N$ and $SN = NS$. A bit more work (see below), shows that $S = f(A)$ and $N = g(A)$ for some polynomials $f, g \in \mathbb{F}[x]$ with 0 constant terms. This means that any matrix commuting with A also commutes with S and N . Since conjugation preserves diagonalizability, nilpotence, and commutativity, we see that any operator A that has a Jordan form (e.g., any A if $\mathbb{F}$ is algebraically closed) can be written $A = S + N$ where S is diagonalizable, N is nilpotent, and S, N commute. This is called the additive **Jordan-Chevalley decomposition** of A and it is unique, as we will see shortly. In these notes, we will drop the adjective “additive”. Some sources call this the **Jordan decomposition**.

To see that $S = f(A), N = g(A)$ for some polynomials f, g with 0 constant terms, suppose $\text{mp}(A) = \prod_{i=1}^k (x - \lambda_i)^{e_i}$ for the distinct eigenvalues $\lambda_1, \dots, \lambda_k$ of A . By Proposition 1.4.5 (The Chinese remainder theorem and Lagrange interpolation, p. 38), there is a polynomial $f \in \mathbb{F}[x]$ such that $f \equiv \lambda_i \pmod{(x - \lambda_i)^{e_i}}$ for $i = 1, \dots, k$ and $f \equiv 0 \pmod{x}$. (If some $\lambda_i = 0$, the last congruence is superfluous.) We see that $f(A)$ acts like multiplication by λ_i on the generalized eigenspace $\ker(A - \lambda_i)^{e_i}$. Thus $f(A) = S$. If we set $g = x - f$, we see $g(A) = N$.

What if A does not have a Jordan form, i.e., some eigenvalues do not lie in $\mathbb{F}$? The general (additive) **Jordan-Chevalley decomposition** of an operator A is a decomposition $A = S + N$ where S is semisimple, N is nilpotent, and S, N commute. Such a decomposition always exists and is unique.

Proposition 1.2.1 (Jordan-Chevalley decomposition).

Let A be an operator on a finite-dimensional vector space over a field $\mathbb{F}$ of characteristic 0. Then there exist a semisimple operator S and a nilpotent operator N with $A = S + N$ and $SN = NS$. Moreover, there are polynomials $f, g \in \mathbb{F}[x]$ with 0 constant terms such that $S = f(A), N = g(A)$. If $A = S' + N'$ for semisimple S' and nilpotent N' with $S'N' = N'S'$, then $S = S'$ and $N = N'$.

REFERENCES AND REMARKS. See [Humphreys, §4.2, Proposition, p. 17; Jacobson, §III.11, Theorem 16, p. 98; Erdmann-Wildon, §16.6, p. 200; Wikipedia].

PROOF. We first prove uniqueness, assuming two results we will verify soon. Suppose $S + N = T = S' + N'$ as above. Since S', N' commute, they commute with T . As S, N are polynomials in T , we see that S', N' commute with S, N . We will show in the near future that the sum of commuting semisimple operators is semisimple (Proposition 1.3.4 (Diagonalizability of a set of operators, p. 28)) and the sum of commuting nilpotent operators is nilpotent (Proposition 1.3.5 (Frobenius’s theorem on triangularizability, p. 29) or prove it directly). Thus $S - S'$ is semisimple and $N - N'$ is nilpotent, and we have

$S - S' = N - N'$. We leave it to the reader to show that the only semisimple and nilpotent operator is 0. Hence we have $S = S'$ and $N = N'$.

For existence, we give a proof from **Wikipedia**, due to Geck, but dating back to Chevalley. To be definite, let us suppose A is a matrix and let R be the (associative) subalgebra of $M_n(\mathbb{F})$ generated by A , that is, $R = \mathbb{F}[A]$. Let q be the product of the distinct irreducible factors of $\text{mp}(A)$ (sometimes called the radical of $\text{mp}(A)$). If 0 is not an eigenvalue of A , replace q by xq : this does not affect the claims below. If m is the largest power of an irreducible factor of $\text{mp}(A)$, we have $\text{mp}(A) \mid q^m$. Because q is squarefree, we know there exist $u, v \in \mathbb{F}[x]$ such that $uq + vq' = 1$. We now define a sequence of elements of R .

$$A_0 = A \quad \text{and} \quad A_{n+1} = A_n - v(A_n)q(A_n)$$

We will prove three statements below:

- (a) Each $A_n \in AR = \{Ar \mid r \in R\}$.
- (b) $A_n - A \in q(A)R$ for each n .
- (c) $q(A_n) \in q(A)^{2^n}R$ for each n .

Assume these statements hold. Since the minimal polynomial of A divides q^m , we see that if $2^n \geq m$, then $q(A)^{2^n} = 0$. By (c), this means $q(A_n) = 0$, so $\text{mp}(A_n)$ divides q . This means that $\text{mp}(A_n)$ is squarefree, i.e., A_n is semisimple. Moreover, $A_n = f(A)$ for some polynomial f divisible by x . Set $g(x) = x - f(x)$ and we see $g(A) = A - A_n \in q(A)R$ and g is divisible by x . Since $q(A)^m = 0$, we see that $g(A)^m = 0$ and so $g(A)$ is nilpotent. Finally, $A = f(A) + g(A)$ by definition, and $f(A) = S, g(A) = N$ certainly commute.

Proof of (a), (b), (c). For (a), we have $A_0 \in AR$. We have $q(A_n) \in A_n R$ since $x \mid q$. Thus $A_n \in AR$ implies $A_{n+1} \in A_n R \subseteq AR$.

For statement (c), let t be an indeterminate. By an algebraic version of Taylor's theorem (expand $q(x)$ in terms of $x - t$), we see $q(x) = q(t) + q'(t)(x - t) + h(x, t)(x - t)^2$ for some $h \in \mathbb{F}[x, t]$. Set $x = A_{n+1}$ and $t = A_n$. In the calculation below, set $r = h(A_{n+1}, A_n) \in R$ and note $q'(A_n)v(A_n) = I - u(A_n)q(A_n)$ by the definition of u, v .

$$\begin{aligned} q(A_{n+1}) &= q(A_n) + q'(A_n)(A_{n+1} - A_n) + (A_{n+1} - A_n)^2 r = \\ &= q(A_n) - q'(A_n)v(A_n)q(A_n) + ((v(A_n)q(A_n))^2 r = \\ &= q(A_n) - (I - u(A_n)q(A_n))q(A_n) + ((v(A_n)q(A_n))^2 r = \\ &= u(A_n)q(A_n)^2 + v(A_n)^2 q(A_n)^2 r = q(A_n)^2 (u(A_n) + v(A_n)^2 r). \end{aligned}$$

Thus we see that $q(A_{n+1}) = q(A_n)^2 s$ for some $s \in R$. Since $q(A_0) = q(A) \in q(A)R$, it follows by induction that $q(A_n) \in q(A)^{2^n}R$ for every n .

For statement (b), apply statement (c) to get $A_{n+1} - A = (A_{n+1} - A_n) + (A_n - A) = -v(A_n)q(A_n) + (A_n - A) = -v(A_n)q(A)^{2^n}r + A_n - A$ for some $r \in R$. Applying induction, we see that both terms on the right hand side are R -multiples of $q(A)$, and so $A_{n+1} - A$ is such a multiple as well. ■

2.2. The multiplicative Jordan-Chevalley decomposition. The Jordan-Chevalley decomposition has a multiplicative version. Let A be an invertible operator and write $A = S + N$, the additive Jordan-Chevalley decomposition. We claim S must be invertible. This rests on the following fact: we know that A, N commute. If A, N commute, A is invertible, and N is nilpotent, then $A - N$ is invertible. Indeed, $(A - N)(A^{-1} + A^{-2}N + A^{-3}N^2 + \cdots + A^{-k}N^{k-1}) = I$ when $N^k = 0$. Thus $S = A - N$ is invertible and of course still semisimple. Now $A = S(I + S^{-1}N)$, and $S^{-1}N$ is nilpotent (since S, N commute). We say an operator is **unipotent** if it equals $I + B$ for a nilpotent B . It's not hard to see that A is unipotent if and only if its characteristic polynomial is $(x - 1)^n$, or equivalently, its minimal polynomial is $(x - 1)^k$ for some k .

Proposition 1.2.2 (Multiplicative Jordan-Chevalley decomposition).

Let A be an invertible operator on a finite-dimensional vector space over a field $\mathbb{F}$ of characteristic 0. Then there exist a semisimple S and a unipotent U with $A = SU$ and $SU = US$. Moreover, there are polynomials $f, g \in \mathbb{F}[x]$, with f having constant term 0 and g having constant term 1, such that $S = f(A), N = g(A)$. If $A = S'U'$ for semisimple S' and unipotent U' with $S'U' = U'S'$, then $S = S'$ and $U = U'$.

PROOF. This is outlined before the proposition, with one point remaining. As before, $S = f(A)$ for a polynomial f and $N = g(A)$ for a polynomial g , where f, g have 0 constant terms. Using the minimal or characteristic polynomial, one can show $S^{-1} = h(S) = h(f(A))$ for some polynomial h . Thus $U = I + S^{-1}N = I + h(f(A))g(A)$.

We leave the proof of uniqueness to the reader. ■

3. Sets of Operators

We now consider the problem of putting multiple operators simultaneously into a nice form. First let us note that the cases of upper and lower triangular form are symmetric and one is obtainable from the other via the transpose. Henceforth we will use triangular or triangularizable to mean upper triangular or upper triangularizable. Given a set X of operators, we say X is **diagonalizable**, **triangularizable**, or **strictly triangularizable** if the elements of X can be *simultaneously* put in the given form. That is, if there is a basis with respect to which all elements of X are in the desired form.

3.1. General remarks. Clearly a set X of operators can only be put in a desirable form if all of its elements individually can be put in that form. Thus we follow Section 1 (Single operators, p. 19) and make the following **non-standard** definition. We say X is **split** if each $A \in X$ is split, i.e., each $A \in X$ can individually be triangularized. Without other hypotheses, X being split does not imply X is triangularizable. Of course, all sets of operators are split if $\mathbb{F}$ is algebraically closed.

A nilpotent operator is strictly triangularizable if and only if it is triangularizable. The same is true of sets, of course. But for sets, we have some new notions. We say a set X of operators is **nil** if every element is nilpotent. Of course, any nil set is split. We say X is **nilpotent** if there is a k such that every (matrix/composition) product of k elements of X is 0. We write $X^k = 0$ when the last condition holds. Any nilpotent set is nil, but being nilpotent is a much stronger condition. For example, the set $\left\{ \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\}$ is nil since both matrices square to 0. However, no product of these matrices which alternates the two matrices is zero, so the set is not nilpotent.

If elements of X can be put in a desirable form, so can other operators.

Proposition 1.3.1 (Subalgebra generated by desirable operators is desirable).

Let X be a set of operators on a finite-dimensional vector space. If X is simultaneously diagonalizable, triangularizable, or strictly triangularizable, then so are the following sets.

- (1) *The subspace of $M_n(F)$ or $\mathfrak{gl}(V)$ spanned by X .*
- (2) *The associative subalgebra (with or without 1) of $M_n(F)$ or $\text{End}(V)$ generated by X .*
- (3) *The Lie subalgebra of $\mathfrak{gl}(n)$ or $\mathfrak{gl}(V)$ generated by X . (This is the smallest subspace of $\mathfrak{gl}(n)$ or $\mathfrak{gl}(V)$ containing X and closed under the operation $[A, B] = AB - BA$.)*

PROOF. All parts follow from (2), and (2) follows from the fact that each class of operators is closed under products and linear combinations. ■

Next we make an observation that enables inductive proofs.

Proposition 1.3.2 (Desirable forms on subspaces and quotient spaces).

Let X be a collection of operators on a finite-dimensional vector space V over a field of characteristic 0 and let W be an X -invariant subspace, that is, a subspace with $A(W) \subseteq W$ for every $A \in X$. For $A \in X$, we define $\bar{V} = V/W$ and $\bar{A} : \bar{V} \rightarrow \bar{V}$ by $\bar{A}(v+W) = Av+W$. Define $X|_W = \{A|_W \mid A \in X\}$ and $\bar{X} = \{\bar{A} \mid A \in X\}$.

- (1) *If X is any one of simultaneously diagonalizable, triangularizable, or strictly triangularizable, then so are $X|_W$ and $\bar{X}$.*
- (2) *X is simultaneously triangularizable or strictly triangularizable if and only if both $X|_W$ and $\bar{X}$ are.*

(3) If $V = \bigoplus_{i=1}^k V_i$ and each V_i is X -invariant, then X is any one of simultaneously diagonalizable, triangularizable, or strictly triangularizable if and only if $X|_{V_i}$ is for all i .

REFERENCES AND REMARKS. This result should be true over a field of any characteristic. The problem is that I do not know of an easy direct proof of the result for $X|_W$ in the triangularizable case, and thus we use theorems with a characteristic 0 hypothesis.

PROOF. (1) Let $v_1, \dots, v_n$ be a basis for V with respect to which every $A \in X$ has the desired form. Then $v_1 + W, \dots, v_n + W$ spans $\bar{V}$ and so it contains a basis for $\bar{V}$. It is easy to see the matrix of $\bar{A}$ with respect to this basis has the desired form for every $A \in X$.

When X consists of a single operator, we have proven the result for $X|_W$ in Proposition 1.1.3 (Desirable forms on subspaces and quotients: single operators, p. 21).

When X is diagonalizable, Proposition 1.3.4 (Diagonalizability of a set of operators, p. 28) tells us first that the elements of X and hence $X|_W$ commute and then tells us that $X|_W$ is diagonalizable.

The triangularizable cases follow from Proposition 1.3.8 (Strong Engel's theorem, p. 33). Suppose for example that X is triangularizable. By Proposition 1.3.1 (Subalgebra generated by desirable operators is desirable, p. 26), the Lie algebra L generated by X is triangularizable. Proposition 1.3.8 (Strong Engel's theorem, p. 33) thus implies that $[A, B]$ is nilpotent for every $A, B \in L$. Thus $[A|_W, B|_W]$ is always nilpotent, and so by the reverse direction of Proposition 1.3.8 (Strong Engel's theorem, p. 33), $L|_W$ is triangularizable. Is there a more direct proof?

(2) Let $v_1, \dots, v_n$ be a basis for V such that $v_1, \dots, v_k$ is a basis for W with respect to which the matrices of elements of $X|_W$ are [strictly] upper triangular and such that $v_{k+1} + W, \dots, v_n + W$ is a basis for $\bar{V}$ with respect to which the matrices of $\bar{X}$ are [strictly] upper triangular. This means that $Av_i \in \sum_{j=k+1}^i \mathbb{F}v_j + W = \sum_{j=1}^i \mathbb{F}v_j$ for $i \geq k+1$. This containment holds for $i \leq k$ since $A|_W$ is [strictly] upper triangular. Thus A is [strictly] triangular with respect to the basis $v_1, \dots, v_n$ for V .

(3) If $\mathcal{B}_i$ is an ordered basis for each V_i , then $\mathcal{B} = \bigcup_{i=1}^k \mathcal{B}_i$ is a basis for V . Moreover, the matrix of $A \in X$ with respect to $\mathcal{B}$ is a block matrix of the matrices with respect to the bases $\mathcal{B}_i$. Since blocks of diagonal or (strictly) triangular matrices are diagonal or (strictly) triangular, this proves (3). ■

The next result is the set version of Proposition 1.1.5 (Consequences of distinct eigenvalues, p. 22). To state it, we need some terminology. Let X be a set of operators on a vector space V . An **eigenvector** for X is a nonzero $v \in V$ such that there is a map $\lambda : X \rightarrow \mathbb{F}$ with the property that $Av = \lambda(A)v$ for all $A \in X$. We call λ the **eigenvalue**

of v . Thus to say that v is an eigenvector for X with eigenvalue λ is simply to say that v is an eigenvector for each $A \in X$ with eigenvalue $\lambda(A)$.

Proposition 1.3.3 (Consequences of distinct eigenvalues for sets).

Let X be a collection of operators on a vector space V over the field $\mathbb{F}$. Let $v_1, \dots, v_n \in V$ be X -eigenvectors for A with corresponding eigenvalues $\lambda_1, \dots, \lambda_n : X \rightarrow \mathbb{F}$. Suppose $\lambda_1, \dots, \lambda_n$ are distinct (as functions).

- (1) $v_1, \dots, v_n$ are linearly independent.
- (2) If $v_1, \dots, v_n$ span V , then every nonzero X -invariant subspace of V contains one of the v_i .

REFERENCES AND REMARKS. The results in (1) and (2) are valid even for infinite collections of v_i .

PROOF. Imitate the proof of Proposition 1.1.5 (Consequences of distinct eigenvalues, p. 22). ■

3.2. Diagonalizability. Here is our first result. Note that one corollary is that a sum of commuting semisimple operators is semisimple.

Proposition 1.3.4 (Diagonalizability of a set of operators).

Let X be a set of operators on a finite-dimensional vector space. Then X is simultaneously diagonalizable if and only if each element of X is diagonalizable and the elements of X commute pairwise.

PROOF. Since any two diagonal matrices commute, simultaneous diagonalization implies individual diagonalization plus commuting.

Suppose now that each element of X can be diagonalized and the elements of X commute. If every element of X is a scalar transformation, then X is simultaneously diagonal. Suppose $A \in X$ is not scalar. Since A is diagonalizable, this means that there are eigenvalues $\lambda_1, \dots, \lambda_k$ ($k \geq 2$) with corresponding eigenspaces $W_1, \dots, W_k$ such that $V = W_1 \oplus \dots \oplus W_k$. Fix an i and let $B \in X$, $w \in W_i$. Then $A(B(w)) = B(A(w)) = B(\lambda_i w) = \lambda_i B(w)$, so $B(w) \in W_i$. Thus each W_i is X -invariant and has dimension less than n . By induction, we can assume $X|_{W_i}$ is diagonalizable for each i . Since $V = W_1 \oplus \dots \oplus W_k$, this implies X is diagonalizable. ■

3.3. Triangularizability. We can imitate the proof above to obtain the analog of Proposition 1.3.4 (Diagonalizability of a set of operators, p. 28) for triangularizability. It is not the best possible result, however. We will explore improvements below. Note that one corollary is that a sum of commuting nilpotent operators is nilpotent; one can prove that result directly.

Proposition 1.3.5 (Frobenius's theorem on triangularizability).

Let X be a set of operators on a finite-dimensional vector space that commute pairwise.

- (1) If X is split, then X is simultaneously triangularizable.*
- (2) If X is nil, then X is simultaneously strictly triangularizable and X is nilpotent.*

REFERENCES AND REMARKS. See [Radjavi-Rosenthal, Theorem 1.1.5, p. 2].

PROOF. We give two proofs. The first follows the lines of the proof of Proposition 1.3.4 (Diagonalizability of a set of operators, p. 28). Recall that the generalized eigenspace for a scalar λ and operator A is $\{v \in V \mid (A - \lambda I)^k(v) = 0 \text{ for some } k\}$. In fact, we can bound the k s that occur by n , so the generalized eigenspace is $\ker(A - \lambda I)^n$.

If every element of X is a scalar transformation, then X is simultaneously diagonal. Suppose $A \in X$ is not scalar. Since A is [strictly] upper triangularizable, this means that there are eigenvalues $\lambda_1, \dots, \lambda_k$ ($k \geq 2$) with corresponding generalized eigenspaces $V_1, \dots, V_k$ such that $V = V_1 \oplus \dots \oplus V_k$. Fix an i and let $B \in X$, $w \in V_i$. Since A, B commute, B commutes with $(A - \lambda I)^k$ for any scalar λ . Then $(A - \lambda_i I)^k(B(w)) = B((A - \lambda_i I)^k(w)) = B(0) = 0$, so $B(w) \in V_i$. Thus each V_i is X -invariant and has dimension less than n . By induction, we can assume $X|_{V_i}$ is [strictly] upper triangularizable for each i . Since $V = V_1 \oplus \dots \oplus V_k$, this implies X is [strictly] upper triangularizable.

For the second proof, let A be a non-scalar element of X , let λ be an eigenvalue of A , and let W be the λ -eigenspace. As in the proof of Proposition 1.3.4 (Diagonalizability of a set of operators, p. 28), W is X -invariant. Let $\bar{V} = V/W$ and note that $X|_W$ and $\bar{X}$ satisfy the hypotheses of the proposition. Thus by induction, $X|_W$ and $\bar{X}$ are [strictly] triangularizable. By Proposition 1.3.2 (Desirable forms on subspaces and quotient spaces, p. 26), it follows that X is [strictly] triangularizable. Note that this proof does not work for diagonalizability because diagonalizability cannot be pulled up from a subspace and a quotient space.

For the fact that X is nilpotent in (2), note that any product of n strictly upper triangular matrices is 0. See the next proposition for more detail. ■

The reader can verify that any product of n strictly upper triangular matrices is 0. Thus if A is strictly upper triangular, $A^n = 0$. Further, if A, B are upper triangular, then $AB - BA$ is strictly upper triangular and hence nilpotent. If you want to see the details, see Proposition 8.2.1 (Multiplication of triangular matrices, p. 174) in Chapter 8 (Miscellaneous Calculations and Observations, p. 173).

Before stating the next result, which is partially an abstraction and generalization of the previous paragraph, we need a definition. Define a **full flag**, or **complete flag**, on an n -dimensional vector space V to be a chain $0 = V_0 \subseteq V_1 \subseteq \dots \subseteq V_n = V$ such

that $\dim V_i = i$ (equivalently, each V_i/V_{i-1} is one-dimensional). A **partial flag** is a strictly ascending chain of subspaces from 0 to V , where the quotients V_i/V_{i-1} may have dimension greater than 1. We will only deal with full flags, and some sources (like [Humphreys, p.13]) use **flag** to mean “full flag”. [Note: the indexing on a flag must be reversed when discussing lower triangular form, or else some other conventions about matrices must be adopted.]

Proposition 1.3.6 (Equivalent conditions for strict triangularizability).

Let X be a set of operators on an n -dimensional vector space V . The following conditions are equivalent.

- (1) X is simultaneously strictly triangularizable.
- (2) X is simultaneously triangularizable and nil.
- (3) X is nilpotent.
- (4) The product of any n elements of X is 0.
- (5) There is a chain $0 = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_k = V$ of subspaces of V such that $A(V_i) \subseteq V_{i-1}$ for $i = 1, \dots, k$ and all $A \in X$.
- (6) There is a full flag $0 = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_n = V$ on V such that $A(V_i) \subseteq V_{i-1}$ for $i = 1, \dots, n$ and all $A \in X$.

PROOF. (1) $\Rightarrow$ (2) and (2) $\Rightarrow$ (4) are matrix calculations.

(4) $\Rightarrow$ (3). This is immediate.

(3) $\Rightarrow$ (5). Let X^i be the set of all products of i elements of X and let $X^i(V)$ denote the sum of the images of the elements of X^i , with $X^0(V) = V$. Since $A_{i+1}A_i \cdots A_2A_1v = A_{i+1}A_i \cdots A_2(A_1v)$, we see that $X^{i+1}(V) \subseteq X^i(V)$. Thus if we set $V_i = X^{k-i}(V)$, we have $A(V_i) = A(X^{k-i}(V)) \subseteq X^{k-(i-1)}(V) = V_{i-1}$. Moreover, $V_k = X^0(V) = V$ and $V_0 = X^k(V) = 0$, since $X^k = 0$. It follows that the V_i give us the desired chain.

(5) $\Rightarrow$ (6). Given a chain as in (5) with all inclusions strict, we can simply fill in spaces between the V_i to make each step codimension 1. For example, suppose the chain $0 = W_0 \subseteq W_1 \subseteq \cdots \subseteq W_k = V$ with $A(W_i) \subseteq W_{i-1}$ for all $A \in X$ is given and suppose $\dim W_1 = d$. Take any chain $0 = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_d = W_1$ with $\dim V_i = i$. Then $A(V_1) \subseteq A(W_1) = W_0 = V_0$. A similar construction works at every step.

(6) $\Rightarrow$ (1). Take a basis $v_1, \dots, v_n$ of V so that $v_1, \dots, v_i$ is a basis for V_i . The matrix of every $A \in X$ with respect to this basis is strictly triangular. ■

The next result is a generalization of a result due to [Jacobson, §2.2, Theorems 1 and 1'; pp. 31-36]. He proved that if X is a collection of nilpotent operators such that $AB - cBA \in X$ for any $A, B \in X$ and some scalar c , then the operators in X can be

simultaneously triangularized. The version below was stated by Radjavi in 1987, and there is an unpublished proof of Bergman from 1978.

Proposition 1.3.7 (Generalized Jacobson's theorem).

Let X be a set of nilpotent operators on a finite-dimensional vector space V over the field $\mathbb{F}$ and suppose that for any $A, B \in X$, the following condition holds:

() There is an operator C in the associative subalgebra generated by A, B such that $AB - CA \in X$.*

Then X is simultaneously strictly triangularizable.

REFERENCES AND REMARKS. See [Jacobson, §2.2, Theorems 1 and 1'; pp. 31-36; Radjavi-Rosenthal, Theorem 1.7.3 and Corollary 1.7.4, pp. 21-22].

An element of the subalgebra generated by A, B is a noncommutative polynomial in A, B , such as $A - BAB + 3A^2BAB^3 + 4I$. In Jacobson's original theorem, $C = cB$.

Jacobson's original theorem applies to any collection of operators that is a semigroup (closed under multiplication; take $c = 0$), Lie algebra (closed under the commutator $[A, B] = AB - BA$; take $c = 1$) or Jordan algebra (closed under the Jordan product $A \circ B = \frac{1}{2}(AB + BA)$ or $A \circ B = AB + BA$; take $c = -1$). If the elements of X commute, then (perhaps adding 0) X is closed under the Lie bracket. Thus this result is a generalization of the case where the elements of X commute.

The result for Lie algebras is one version of Engel's theorem, which was known in the late 19th century and the result for semigroups is due to Levitzki in 1931.

PROOF. We proceed by induction on $n = \dim V$ and all such sets X .

If X contains no elements or only contains 0, then X is already triangular. Suppose X contains a nonzero element B . Since B is nilpotent, we have $0 \subsetneq \ker B \subsetneq V$. For any $Y \subseteq X$, define $K(Y) = \{v \in V \mid Av = 0 \text{ for all } A \in Y\} = \cap_{A \in Y} \ker A$. The above observation shows that there exist subsets Y (e.g., $Y = \{B\}$) with $K(Y) \neq 0$. Choose a Y_0 with $W = K(Y_0) \neq 0$ having minimal dimension. The example of B shows $W \subsetneq V$. Now set $Y = \{A \in X \mid Aw = 0 \text{ for all } w \in W\}$. Minimality of W implies that $K(Y) = W$.

Let $\bar{V} = V/W$ and let $\bar{Y} = \{\bar{A} \mid A \in Y\}$, where $\bar{A}$ is defined on $\bar{V}$ by $\bar{A}(v + W) = Av + W$. $\bar{Y}$ satisfies condition (*), although this is somewhat subtle. Certainly $\bar{Y}$ consists of nilpotent operators and if $a, b \in \bar{Y}$, we can write $a = \bar{A}, b = \bar{B}$ for $A, B \in Y$. By (*), there is a C in the subalgebra generated by A, B with $AB - CA \in X$. Clearly $AB - CA \in Y$. Here is the subtle point: we wish to create $\bar{C}$. Since Y is a noncommutative polynomial in A, B , it leaves invariant any subspace left invariant by A and B . In particular, C leaves W invariant. Thus we can define $c = \bar{C}$ and we have $ab - ca = \overline{AB - CA} \in \bar{Y}$. Since $W \subsetneq V$, induction tells us $\bar{Y}$ is triangularizable. Since $Y|_W$ is identically 0, it is already triangular, and hence we see that Y is triangularizable. Proposition 1.3.6 (Equivalent

conditions for strict triangularizability, p. 30) tells us Y is strictly triangularizable and any product of n elements of Y is 0.

Our goal now is to show $Y = X$, which will of course prove X is triangularizable. First suppose $B \in X$ and W is invariant under B . Since B is nilpotent, this implies either $B \in Y$ or there is a $w \in W$ with $w' \neq 0, Bw' = 0$. Suppose the latter holds and set $W' = K(Y \cup \{B\}) \subseteq W$. Since $w' \in W'$, we see $W' \neq 0$ and so we must have $W' = W$. This implies $B(W) = 0$, i.e., $B \in Y$. We've shown if W is invariant under B , then $B \in Y$.

Now suppose $B \in X$ and W is not invariant under B , say $Bw \notin W$ for some $w \in W$. Then there is a matrix $A \in Y$ with $ABw \neq 0$. Set $A_1 = A, B_1 = B, w_1 = w$. We will construct sequences $A_1, \dots, A_n \in Y, B_1, \dots, B_n \in X, w_1, \dots, w_n \in W$ such that

- (i) $A_i B_i w_i \neq 0$;
- (ii) $B_{i+1} = A_i B_i - C_i A_i$ for some C_i as in (*).
- (iii) $A_i A_{i-1} \cdots A_1 B w = A_i A_{i-1} \cdots A_j B_j w$ for all $w \in W$ and all $0 \leq j \leq i \leq n$.

We start with $i = 2$; the confident reader can skip this paragraph. We apply (*) to get $B_2 = A_1 B_1 - C_1 A_1$ for some C_1 , giving (ii). Since $B_2 w_1 = A_1 B_1 w_1 - C_1 A_1 w_1 \neq 0$, we see $B_2 \notin Y$ and hence W is not invariant under B_2 . As above, this implies there exist $w_2 \in W, A_2 \in Y$ with $A_2 B_2 w_2 \neq 0$, giving (i). For $w \in W$, we have $A_2 A_1 B w = A_2 A_1 B_1 w = A_2 (B_2 + C_1 A_1) w = A_2 B_2 w$, giving (iii).

The general construction from i to $i + 1$ is much like the $i + 1 = 2$ case. We apply (*) to get $B_{i+1} = A_i B_i - C_i A_i$ for some C_i , giving (ii). Since $B_{i+1} w_i = A_i B_i w_i - C_i A_i w_i \neq 0$, we see $B_{i+1} \notin Y$ and hence W is not invariant under B_{i+1} . As above, this implies there exist $w_{i+1} \in W, A_{i+1} \in Y$ with $A_{i+1} B_{i+1} w_{i+1} \neq 0$, giving (i). If $w \in W$ and $j \leq i$, we have $A_{i+1} A_i A_{i-1} \cdots A_1 B w = A_{i+1} A_i A_{i-1} \cdots A_j B_j w$ by induction. For $j = i + 1$, we have $A_{i+1} A_i A_{i-1} \cdots A_1 B w = A_{i+1} A_i B_i w = A_{i+1} (B_{i+1} + C_i A_i) w = A_{i+1} B_{i+1} w$, giving (iii).

In the case $i = n, j = 1$ and $w = w_n$, (iii) becomes $A_n A_{n-1} \cdots A_1 B w_n = A_n B_n w_n$. The right hand side is nonzero by (i), but the left hand side is 0 since any product $A_n A_{n-1} \cdots A_1$ from Y is 0. This contradiction shows no $B \in X$ can exist with W not invariant under B , and so every element of X is in Y . This completes the proof. ■

The case of semigroups has been much studied, and we state two results in that case. Kolchin proved in 1948 that if every member of a semigroup $X \subseteq M_n(\mathbb{F})$ is unipotent, then X is triangularizable (see [Radjavi-Rosenthal, Theorem 2.1.8, p. 29]). Here is another result of interest, probably due to Radjavi, 1986 (see [Radjavi-Rosenthal, Corollary 4.2.14, p. 88]). A semigroup X is triangularizable if and only if $\{A, B\}$ is triangularizable for all $A, B \in X$ (possibly with the assumption $\mathbb{F} = \overline{\mathbb{F}}$).

The following example (from Bergman and [Radjavi-Rosenthal, Example 1.7.11, p. 24]) shows that a vector space of nilpotent matrices need not be triangularizable. Let

$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ and let X be the (vector space) span of A, B , that is,

$X = \left\{ \begin{pmatrix} 0 & a & 0 \\ -b & 0 & a \\ 0 & b & 0 \end{pmatrix} \mid a, b \in \mathbb{F} \right\}$. A direct calculation shows the cube of any element of X

is 0. However $AB = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ and no power $(AB)^k$ is zero, so X is not nilpotent as a

set. The nullspace of A is spanned by $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ and the nullspace of B is spanned by $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, so there is no common null vector for X and hence X is not triangularizable.

We can use Engel's theorem to give an "if and only if" characterization of triangularizability in the Lie algebra case. In its proof we use the fact — whose proof we leave to the reader — that if A, B are upper triangular matrices, then $[A, B] = AB - BA$ is strictly upper triangular.

Proposition 1.3.8 (Strong Engel's theorem).

Suppose X is a set of split operators on a finite-dimensional vector space and suppose that $AB - BA \in X$ whenever $A, B \in X$. Then X is simultaneously triangularizable if and only if $AB - BA$ is nilpotent for each $A, B \in X$.

REFERENCES AND REMARKS. See [Radjavi-Rosenthal, Corollary 1.7.8, p. 23].

PROOF. First suppose X is a triangularizable set of matrices and let $A, B \in X$. Choose a matrix P such that both PAP^{-1}, PBP^{-1} are upper triangular. As noted above $[PAP^{-1}, PBP^{-1}]$ is strictly upper triangular and hence nilpotent. Since $[PAP^{-1}, PBP^{-1}] = P[A, B]P^{-1}$, it follows that $[A, B]$ is nilpotent.

Now let $Y \subseteq X$ be the set of $AB - BA$ such that $A, B \in X$ and suppose every element of Y is nilpotent. By hypothesis, Y is closed under commutators. Thus by Proposition 1.3.7 (Generalized Jacobson's theorem, p. 31) or Proposition 2.4.3 (Concrete Engel's theorem, p. 63), the set Y is triangularizable and nilpotent. Proposition 1.3.6 (Equivalent conditions for strict triangularizability, p. 30) tells us that there is a nonzero subspace of V that is annihilated by all $A \in Y$. Thus $W = \cap_{A \in Y} \ker A \neq 0$. Suppose $A \in Y, B \in X, w \in W$. Then $A(Bw) = [A, B]w + BAw = 0 + 0$, and so $Bw \in W$. This shows that W is X -invariant. Since the conditions that individual operators are triangularizable and that commutators are nilpotent is inherited by $X|_W$ and $\bar{X}$ (on V/W), we may assume by induction on $\dim V$ that $X|_W$ and $\bar{X}$ are triangularizable.

Proposition 1.3.2 (Desirable forms on subspaces and quotient spaces, p. 26)(2) thus implies that X is triangularizable. ■

Radjavi in 1986 proved the following result. He says X has **permutable trace** when the trace of any product of elements of X remains unchanged however the order is permuted. If X is closed under multiplication, X has permutable trace whenever $\text{tr } ABC = \text{tr } BAC$ for all $A, B, C \in X$.

Proposition 1.3.9 (Permutable trace and triangularizability).

Let X be a set of split operators on a finite-dimensional vector space over a field of characteristic 0. Then X is simultaneously triangularizable if and only if X has permutable trace. **Need to make sure the split condition holds when needed.**

REFERENCES AND REMARKS. See [Radjavi-Rosenthal, Theorem 2.2.1, p. 33].

We can drop the condition that $\mathbb{F} = \overline{\mathbb{F}}$ if we assume X is closed under commutators and the elements of X are split.

PROOF. Suppose first that X is triangularizable. The diagonal entries of the products of upper triangular matrices do not depend on the order of the product. Thus the traces of those products do not depend on the order.

Now assume X has permutable trace. This condition doesn't change under taking products and sums, so we may assume X is a subspace closed under multiplication (even an associative subalgebra, since we can add I). In particular, we can assume X is closed under commutators. If $A, B \in X$ and k is a positive integer, then every term in the expansion of $(AB - BA)^k$ has k copies of A and k copies of B . Since the trace is permutable, each such term has trace $\text{tr } A^k B^k$. Thus $\text{tr } (AB - BA)^k = \sum_{i=0}^k \binom{k}{i} (-1)^i \text{tr } A^k B^k = (\sum_{i=0}^k \binom{k}{i} (-1)^i) \text{tr } A^k B^k = 0$. Since an operator T is nilpotent if and only if the trace of any power of T is 0, we see that all commutators in X are nilpotent. We may apply Proposition 1.3.8 (Strong Engel's theorem, p. 33) to conclude that X is triangularizable.

We leave the remark to the reader. ■

4. Other Constructions

4.1. Bilinear forms and the trace form. Recall that a **bilinear form** on V is a bilinear map $\langle, \rangle : V \times V \rightarrow \mathbb{F}$. We say it is **symmetric** if $\langle v, w \rangle = \langle w, v \rangle$ for all $v, w \in V$ and we say it is **skew-symmetric** if $\langle v, w \rangle = -\langle w, v \rangle$ for all $v, w \in V$. We say a form $\langle, \rangle$ is **alternating** if $\langle v, v \rangle = 0$ for all v . Since we assume $\text{char } \mathbb{F} = 0$, the skew-symmetric and alternating conditions are equivalent. The **radical** of a symmetric or skew-symmetric bilinear form $\langle, \rangle$ is $\{v \in V \mid \langle v, w \rangle = 0 \text{ for all } w \in V\}$. A form is **nondegenerate**

or **nonsingular** when its radical is 0. From now on, we will assume forms are either symmetric or skew-symmetric.

More generally, if W is a subset of V , we define its **orthogonal complement** to be $W^\perp = \{v \in V \mid \langle v, w \rangle = 0 \text{ for all } w \in W\}$; this is always a subspace. Thus the radical of $\langle, \rangle$ is $V^\perp$.

Here are some useful properties of orthogonal complements.

Proposition 1.4.1 (Orthogonal complements).

Let $\langle, \rangle$ be a symmetric or skew-symmetric nondegenerate bilinear form on a finite-dimensional vector space V and let U, W be subspaces of V . Then the following statements hold.

- (1) $(U \cap W)^\perp = U^\perp + W^\perp$.
- (2) $(U + W)^\perp = U^\perp \cap W^\perp$.
- (3) $\dim W + \dim W^\perp = \dim V$.
- (4) $(W^\perp)^\perp = W$.
- (5) If $W \cap W^\perp = 0$, then $V = W \oplus W^\perp$.

PROOF. These are well-known facts. We leave it to the reader to supply a proof or look one up. ■

For results on degenerate forms, see Subsection 7.1 (Degenerate forms, p. 183) in Chapter 8 (Miscellaneous Calculations and Observations, p. 173).

If $\langle, \rangle$ is a bilinear form on V , we can define a linear map $\Phi : V \rightarrow V^*$ as follows. For $x \in V$, let $\Phi(x) \in V^*$ be given by $\Phi(x)(y) = \langle x, y \rangle$ for $y \in V$. The kernel of Φ is precisely the radical of the form $\langle, \rangle$. Conversely, if $\Phi : V \rightarrow V^*$ is a linear map, we can define a bilinear form on V by $\langle x, y \rangle = \Phi(x)(y)$ for $x, y \in V$. However, it is not clear in this direction if the form $\langle, \rangle$ defined is symmetric or skew-symmetric.

Proposition 1.4.2 (Nondegenerate forms and the dual space).

Let V be a finite-dimensional vector space and let $\langle, \rangle$ be a symmetric or skew-symmetric bilinear form on V . Then $\Phi : V \rightarrow V^*$ defined by $\Phi(x)(y) = \langle x, y \rangle$ is an isomorphism if and only if $\langle, \rangle$ is nondegenerate.

PROOF. This is virtually proved in the comments before the proposition. ■

We can define a bilinear form τ on $M_n(\mathbb{F}) = \mathfrak{gl}(n, \mathbb{F})$ or $\text{End}(V) = \mathfrak{gl}(V)$ by $\tau(A, B) = \text{tr } AB$. We call this the **concrete trace form**; we will define other trace forms later. Since $\text{tr } AB = \text{tr } BA$, this form is symmetric. It also has properties called **associativity** or **invariance**. These mean that $\langle AB, C \rangle = \langle A, BC \rangle$ and $\langle [A, B], C \rangle = \langle A, [B, C] \rangle$, where $[A, B] = AB - BA$ is the (additive) commutator. The first equality is trivial; for the

second, note $\tau([A, B], C) = \operatorname{tr} ABC - \operatorname{tr} BAC = \operatorname{tr} ABC - \operatorname{tr} ACB + \operatorname{tr} ACB - \operatorname{tr} B(AC) = \operatorname{tr} A[B, C] + \operatorname{tr} ACB - \operatorname{tr} ACB = \tau(A, [B, C])$.

The concrete trace form is nondegenerate since if $a_{ij} \neq 0$ and e_{rs} is the matrix unit with a 1 in the r, s entry and 0 elsewhere, we have $\tau(A, e_{ji}) = a_{ij} \neq 0$.

We will often consider the concrete trace form on subspaces of $\mathfrak{gl}(n)$, in which case it may no longer be nondegenerate. For example, if N is the subspace of strictly upper triangular matrices, then the restriction of the concrete trace form to N is identically zero. If W is the space of all $n \times n$ upper triangular matrices, the radical of the concrete trace form restricted to W is the space of strictly upper triangular matrices. However, the concrete trace form is nondegenerate on the space $\mathfrak{sl}(n)$ of all matrices with trace 0. To see this, note that the proof for $\mathfrak{gl}(n)$ applies to any matrix $A \in \mathfrak{sl}(n)$ that is not diagonal. Suppose $A \in \mathfrak{sl}(n)$ is diagonal: note $\tau(A, e_{ii} - e_{jj}) = a_{ii} - a_{jj}$ for all i, j . If this is always 0, then all diagonal entries of A are equal, let us say to a . In this case $0 = \operatorname{tr} A = na$. Since $\mathbb{F}$ has characteristic 0, this implies $a = 0$, whence $A = 0$.

We state a result from linear algebra we will need later.

Proposition 1.4.3 (Trace of a restriction).

Let V be a finite-dimensional vector space with subspace W and let $\phi : V \rightarrow V$ be a linear operator with $\operatorname{im} \phi \subseteq W$. Then $\operatorname{tr} \phi = \operatorname{tr}(\phi|_W)$.

PROOF. Take a basis $v_1, \dots, v_k$ of W and extend it to a basis $v_1, \dots, v_k, v_{k+1}, \dots, v_n$ of V . Since $\phi(v_i)$ is a linear combination of $v_1, \dots, v_k$ for any i , the matrix of ϕ with respect to the full basis contains only zeros on the diagonal for $i > k$. This proves $\operatorname{tr} \phi = \operatorname{tr} \phi|_W$. ■

4.2. Extension of scalars. Let $\mathbb{E}$ be an extension field of $\mathbb{F}$. We sometimes prefer to work over $\mathbb{E}$ instead of $\mathbb{F}$; for example, $\mathbb{E}$ might be the algebraic closure of $\mathbb{F}$. Given a vector space V over $\mathbb{F}$, we define the vector space $V_{\mathbb{E}}$ over $\mathbb{E}$ to be $\mathbb{E} \otimes_{\mathbb{F}} V$. The $\mathbb{E}$ -vector space structure is given by $\lambda(\mu \otimes v) = \lambda\mu \otimes v$ for $v \in V$, $\lambda, \mu \in \mathbb{E}$. Elements of $V_{\mathbb{E}}$ and sums of simple tensors $\lambda_i \otimes v_i$. If $v_1, \dots, v_n$ is a basis of V , then $1 \otimes v_1, \dots, 1 \otimes v_n$ is a basis of $V_{\mathbb{E}}$.

If $A : V \rightarrow V$ is a linear operator on V , then we define $A_{\mathbb{E}} = \operatorname{id}_{\mathbb{E}} \otimes A : V_{\mathbb{E}} \rightarrow V_{\mathbb{E}}$: this is a linear operator on $V_{\mathbb{E}}$. (A similar construction works for linear maps from V to W .) The matrix of $A_{\mathbb{E}}$ relative the extended basis $1 \otimes v_1, \dots, 1 \otimes v_n$ is identical to the matrix of A relative to $v_1, \dots, v_n$. For this reason, we often omit the $\mathbb{E}$ subscript for matrices. It follows that A and $A_{\mathbb{E}}$ (for operator or matrix) have the same minimal polynomial and characteristic polynomial, the same determinant and trace, and the same eigenvalues in $\mathbb{E}$ (i.e., some eigenvalues may lie in $\mathbb{E}$ but not in $\mathbb{F}$, but they are not “new”). The conditions of being semisimple or nilpotent are the same for A and $A_{\mathbb{E}}$.

The process of passing from V to $V_{\mathbb{E}}$ and A to $A_{\mathbb{E}}$ is called **extension of scalars**. We are interested in this because there are many properties that are preserved in both directions by extension of scalars. We noted several examples above, and we will see more when we study Lie algebras. This enables us, in some situations, to assume that $\mathbb{F}$ is algebraically closed. For such a property, we pass from $\mathbb{F}$ to $\mathbb{E} = \overline{\mathbb{F}}$, prove that the Lie algebra $L_{\mathbb{E}}$ has the property in the algebraically closed case, and then conclude that the original Lie algebra L over $\mathbb{F}$ has that property.

Proposition 1.4.4 (Injectivity/surjectivity and extension of scalars).

Injectivity/surjectivity and extension of scalars Let V, W be vector spaces over $\mathbb{F}$ and let $\mathbb{E}$ be an extension field of $\mathbb{F}$. Let $\phi : V \rightarrow W$ be an $\mathbb{F}$ -linear map, so $\phi_{\mathbb{E}} : V_{\mathbb{E}} \rightarrow W_{\mathbb{E}}$ is $\mathbb{E}$ -linear. Then $\ker \phi_{\mathbb{E}} = (\ker \phi)_{\mathbb{E}}$ and $\operatorname{im} \phi_{\mathbb{E}} = (\operatorname{im} \phi)_{\mathbb{E}}$. Hence ϕ is injective or surjective or bijective if and only if $\phi_{\mathbb{E}} : V_{\mathbb{E}} \rightarrow W_{\mathbb{E}}$ is.

PROOF. Omitted. ■

4.3. The matrix exponential. If A is a linear operator, a matrix in $M_n(\mathbb{F})$ or an operator in $\operatorname{End}(V)$, then we define the **exponential** of A to be $e^A = \exp A = \sum_{i=0}^{\infty} \frac{1}{i!} A^i$. For example, $e^0 = I$. If A is nilpotent, say $A^n = 0$, then this is a finite sum: $\exp(A) = \sum_{i=0}^{n-1} \frac{1}{i!} A^i$. If $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$, this infinite series converges absolutely for all A .

It's easy to see that e^A is again a matrix or linear operator. If $AB = BA$ for matrices A, B , then $e^A e^B = e^{A+B}$ follows by manipulating power series as follows. When we expand e^A, e^B and multiply the series, the terms of the product having total degree n in A, B sum to $\sum_{i=0}^n \frac{1}{i!(n-i)!} A^i B^{n-i}$. Since A, B commute, we have $\frac{1}{n!} (A+B)^n = \sum_{i=0}^n \frac{\binom{n}{i}}{n!} A^i B^{n-i} = \sum_{i=0}^n \frac{1}{i!(n-i)!} A^i B^{n-i}$. Thus the two series are equal. Thus when it exists, e^A is nonsingular, i.e., an isomorphism.

It follows that $e^A e^{-A} = e^{-A} e^A = e^0 = I$. If $AB \neq BA$, then this equality is not valid, but a more complicated formula $e^A e^B = e^C$, the Baker-Campbell-Hausdorff formula, may hold for some C . This C is usually given by an infinite series like $C = A + B + \frac{1}{2}[A, B] + \frac{1}{12}([[A, B], B] - [[A, B], A]) + \dots$. The name “Baker-Campbell-Hausdorff” is not universal, especially in older sources. Some call it “Campbell-Hausdorff” and some just refer to Hausdorff. (See [Jacobson, §V.5, pp. 170-174; Bourbaki 1-3, §II.6-8, p. 155; Historical note, §V, middle of p. 425; Hall, Chapter 3, p. 63; Wikipedia].)

If N is nilpotent, then e^N is equal to I plus a multiple of N that commutes with N . Thus e^N is unipotent and we see that $\operatorname{tr} N = 0$ and $\det e^N = 1 = e^{\operatorname{tr} N}$. In fact, this is true for all matrices, whenever the exponential makes sense: $\det e^A = e^{\operatorname{tr} A}$. To see this for $\mathbb{F}$ algebraically closed, assume A is in Jordan form and note the exponential of an upper triangular matrix is upper triangular with the exponentials of the original diagonal elements on the diagonal.

4.4. Lagrange interpolation and the Chinese remainder theorem. Several times, we will wish to construct a polynomial with certain properties. One of those properties will sometimes be that the constant term is 0. The following results are used to do this.

Proposition 1.4.5 (The Chinese remainder theorem and Lagrange interpolation).

Let $\mathbb{F}$ be a field.

- (1) Let $p_1, \dots, p_n$ be pairwise relatively prime polynomials in $\mathbb{F}[t]$ and let $f_1, \dots, f_n \in \mathbb{F}[t]$. Then there is a polynomial $f \in \mathbb{F}[t]$ with $f \equiv f_i \pmod{p_i}$ for each $i = 1, \dots, n$. If $p_i(0) \neq 0$ for all i or if some $p_i(0) = 0$ and the corresponding $f_i(0) = 0$, we can take f with $f(0) = 0$.
- (2) Let $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{F}$ and suppose $b_i = b_j$ whenever $a_i = a_j$. Then there is a polynomial $f \in \mathbb{F}[t]$ of degree less than n with $f(a_i) = b_i$ for each $i = 1, \dots, n$. If no $a_i = 0$ or if some $a_i = 0$ and the corresponding $b_i = 0$, we can take f with $f(0) = 0$ and $\deg f \leq n$.

REFERENCES AND REMARKS. The Chinese remainder theorem is often abbreviated as CRT, and the result in (2) is called Lagrange interpolation.

If $\text{char } \mathbb{F} = 0$, we can specify not just $f(a_i)$ but also some derivatives of f at a_i in (2) by choosing $p_i = (x - a_i)^k$ and $f_i = b_i + c_i(x - a_i) + \frac{1}{2}d_i(x - a_i)^2 + \dots$ for appropriate k and f_i .

PROOF. (1) This is well known. For the claim about $f(0) = 0$, add the polynomial x to the p_i . You can also imitate the proof of (2) below.

(2) We can assume the a_i are distinct by eliminating duplicates. This follows from (1) by taking $p_i = x - a_i$ and $f_i = b_i$ (except perhaps for the degree requirement). Here is a direct proof. For $1 \leq j \leq n$, set $g_j = \prod_{i \neq j} (x - a_i)$, $c_j = b_j/g_j(a_j)$. The definition of c_j makes sense since $g_j(a_j) \neq 0$. Then $f = \sum_{j=1}^n c_j g_j$ satisfies the requirements for f .

For the additional condition, if some $a_i = b_i = 0$, we automatically have $f(0) = 0$. If no $a_i = 0$, add the condition $a_0 = b_0 = 0$. ■

CHAPTER 2

Basics of Lie Algebras

In Chapter 1 (Desirable Forms for Matrices and Operators, p. 17) we saw several situations where hypotheses on the additive commutator $[A, B] = AB - BA$ led to nice results. This is one of the motivations for focusing on vector spaces closed under $[\cdot, \cdot]$. As usual in algebra, we go from this concrete notion to a more abstract version of $[\cdot, \cdot]$ that satisfies enough identities to allow us to create a useful theory and also applies to a number of examples.

In Section 1 (Definition and Examples, p. 41), we give the formal definitions of concrete (linear) and abstract (general) Lie algebras. We then give numerous examples. Quintessential examples include the full general linear Lie algebra $\mathfrak{gl}(n)$ of $n \times n$ matrices, the special linear Lie algebra of matrices of trace 0, the Abelian Lie algebra of diagonal matrices $\mathfrak{d}(n)$, the solvable Lie algebra of upper triangular matrices $\mathfrak{t}(n)$, and the nilpotent Lie algebra of strictly upper triangular matrices $\mathfrak{n}(n)$. In addition to these Lie algebras, we also have (mostly simple) Lie algebras that almost preserve a bilinear form. Among these are the orthogonal Lie algebra $\mathfrak{o}(n)$ of skew-symmetric matrices and the symplectic Lie algebra $\mathfrak{sp}(n)$. The special linear, orthogonal, and symplectic Lie algebras give us almost all of the simple Lie algebras over an algebraically closed field of characteristic 0. (Symmetric matrices are not closed under $[\cdot, \cdot]$.) We then give examples of some abstract Lie algebras of dimensions 2 and 3. In dimension 2, there are only two examples, but in dimension 3, there are many examples. It turns out that every (finite-dimensional) Lie algebra is isomorphic to a linear one. We state this theorem but do not prove it here.

In Section 2 (Morphisms and Subobjects, p. 47) we describe the usual paraphernalia of an algebraic theory such as homomorphisms, subalgebras, quotients, and ideals. The theory of Lie algebras features a special type of map that is not a homomorphism but often takes the place of automorphisms in the theory. This map is a derivation and we define it, define an inner derivation, and describe how get an automorphism from a (special) derivation.

In Section 3 (Types of Lie Algebras, p. 52), we define various classes of Lie algebras: nilpotent, solvable, and semisimple. Nilpotent and solvable Lie algebras are defined in the same way nilpotent and solvable groups are, via derived and central series, and the theory is developed in a very similar way. Thus if the reader is aware of a theorem for

nilpotent or solvable groups, it is every likely an analogous result holds for Lie algebras. We show that any Lie algebra contains a largest ideal that is nilpotent as a Lie algebra (the nil radical) and a largest ideal that is solvable as a Lie algebra (the solvable radical, or just the radical). The solvable radical plays a crucial role in the theory of Lie algebras, in some ways pointing the theory in the largely disjoint directions of solvable Lie algebras and semisimple Lie algebras. For example, a Lie algebra is semisimple precisely when its solvable radical is 0. To many people, the theory of Lie algebras is almost synonymous with the theory of semisimple Lie algebras, which we will explore in later chapters.

In Section 4 (Adjoint, p. 60), we introduce the adjoint, which plays the role that conjugation plays in the theory of groups. The adjoint of x , denoted $\text{ad } x$, is the map from L to itself that takes y to $[x, y]$. This is (the definition of) an inner derivation and it maps any L to a linear Lie algebra. When the center $Z(L)$ of L is zero, ad provides an isomorphism from L to a linear Lie algebra. One example of the use of the adjoint is Engel's Theorem. The concrete version (from Chapter 1 (Desirable Forms for Matrices and Operators, p. 17) says that a linear Lie algebra of nilpotent operators can be simultaneously strictly upper triangularized and so is nilpotent in the Lie algebra sense. The general version says that a general Lie algebra L is nilpotent (as a Lie algebra) if and only if $\text{ad } x$ is a nilpotent operator for each $x \in L$. (Sophus Lie was a Norwegian mathematician who introduced Lie algebras in the 1870s. Friedrich Engel was a collaborator of Lie. According to **Wikipedia**, he sketched a proof of Engel's theorem in 1890 and his student Umlauf gave a complete proof in his dissertation in 1891. We will encounter Lie's theorem, which characterizes solvability in terms of triangularizability much as Engel's theorem characterizes nilpotence, in Chapter 3 (Representations, p. 83). Lie seems to have proved his theorem on solvability around 1876.)

One of the interesting, and often confusing, features of Lie algebras is the interplay between the concrete and the abstract. The adjoint frequently provides the connection between the two. We will see this connection strongly in this chapter and the next.

There is an elegant way of proving the derived and central series and various other constructions in group theory are normal subgroups: the notion of a characteristic subgroup. Recall that a subgroup of G is characteristic if it is invariant under every automorphism of G . For Lie algebras, there are two notions of characteristic: invariance under automorphisms and invariance under derivations. In Section 5 (Characteristic Subspaces, p. 64), we explore these two types of characteristic subspaces. The derivation-invariant version of characteristic is much more prominent in the theory of Lie algebras.

In Section 6 (Cartan's Criterion for solvability, p. 66), we state criteria for a Lie algebra to be solvable based on conditions on the trace of products. If L is a linear Lie algebra, Cartan's criterion says that L is solvable if $\text{tr } xy = 0$ for all $x \in L, y \in [L, L]$ or even if

$\text{tr } x^2 = 0$ for all $x \in [L, L]$. For a general L , Cartan's criterion says that L is solvable if and only if $\text{tr } \text{ad } x \circ \text{ad } y = 0$ for all $x \in L, y \in [L, L]$ or even if $\text{tr}(\text{ad } x)^2 = 0$ for all $x \in [L, L]$. These results can also be stated in terms of the trace form or Killing form.

In Section 7 (Invariant Forms, p. 70), we introduce a enormously important tool in Lie theory, the Killing form (named after Wilhelm Killing, who invented Lie algebras in the 1880s, independently of Sophus Lie who had introduced them in the 1870s). The theory of Lie algebras relies heavily on the properties of the trace and the Killing form is one example of this. It is trivially true that $\text{tr}(AB)C = \text{tr } A(BC)$, but a less trivial fact is true $\text{tr}[A, B]C = \text{tr } A[B, C]$ (prove it!). This property is called invariance and invariant forms have many nice properties. The Killing form is defined by $\kappa(x, y) = \text{tr}(\text{ad } x \text{ ad } y)$ (where the multiplication $\text{ad } x \text{ ad } y$ is composition, i.e., matrix multiplication.) Orthogonal complements of ideals with respect to the Killing form are ideals of special interest. One version of Cartan's criterion for solvability can be stated as: L is solvable if and only if $[L, L]$ is contained in the radical of the Killing form. Cartan's criterion for semisimplicity states that a Lie algebra L (finite-dimensional over a field of characteristic 0) is semisimple if and only if the Killing form on L is nondegenerate. (This is due to Élie Cartan, who proved a rigorous classification of semisimple Lie algebras in his dissertation in 1894, which included this criterion. Some sources refer to the Cartan-Killing form.) This can be used to show that semisimple Lie algebras are (uniquely) direct sums of simple (non-Abelian) Lie algebras. The Killing form can also be used to show all derivations on a semisimple Lie algebra are inner.

In Section 8 (Extension of Scalars, p. 80), we note some properties that are preserved when the field $\mathbb{F}$ is replaced by a larger field $\mathbb{E}$. Examples include nilpotence, solvability, and semisimplicity. When we have preservation both up and down, we can frequently do proofs under the assumption that the base field is algebraically closed and then transport the result to the general case.

1. Definition and Examples

1.1. The definition of a Lie algebra. A subspace of some $\mathfrak{gl}(n)$ or $\mathfrak{gl}(V)$ that is closed under the commutator $[A, B] = AB - BA$ is called a **linear Lie algebra**. (It is not uncommon to loosen this definition to allow any Lie algebra isomorphic to one of these to be called linear, but we will not do so.)

The following definition is **not standard**. We say a linear Lie algebra is **split** if each operator $x \in L$ is split, i.e., if each element of x is triangularizable as an operator on V . Of course, if $\mathbb{F}$ is algebraically closed, every linear Lie algebra is split. "Split" often has different meanings. One example is: a linear Lie algebra that contains the semisimple and nilpotent parts of its elements. "Splittable" is also used for this concept;

[**Jacobson**, §III.11, Definition 3, p. 98] uses the term “almost algebraic” for this concept and [**Bourbaki 7-9**, §VII.5.1, Definition 1, p. 34] uses the term “decomposable”. However, “almost algebraic” also has other meanings. A much more important example is that a “split Lie algebra” may refer to a pair (L, H) such that H is a splitting Cartan subalgebra of L .

We now define a general **Lie algebra** to be a vector space L equipped with a product $[\cdot, \cdot] : V \times V \rightarrow V$ that satisfies the following properties:

- (Bilinearity) $[x + y, z] = [x, z] + [y, z]$ and $[x, y + z] = [x, y] + [x, z]$ for all $x, y, z \in L$, and $[\lambda x, y] = \lambda[x, y]$ and $[x, \lambda y] = \lambda[x, y]$ for all $x, y \in L, \lambda \in \mathbb{F}$.
- (Alternation or Skew-Symmetry) $[x, x] = 0$ for all $x \in L$.
- (Jacobi identity) $[[x, y], z] + [[y, z], x] + [[z, x], y] = 0$ for all $x, y, z \in L$.

We call the product $[\cdot, \cdot]$ the **Lie bracket** or just **bracket** or **Lie product**.

Note: Some sources, such as [**Humphreys**] and [**Jacobson**], omit the comma and write $[x y]$ in place of $[x, y]$.

A few remarks on this definition are in order. The alternating condition implies skew-symmetry. To see this, note $0 = [x + y, x + y] = [x, x] + [x, y] + [y, x] + [y, y]$, so $[y, x] = -[x, y]$. Since $\text{char } \mathbb{F} \neq 2$, skew-symmetry implies alternation since skew-symmetry implies $[x, x] = -[x, x]$. The definition we have given is the standard one in all characteristics. Note that if we know $[\cdot, \cdot]$ is skew-symmetric, then we only need to check linearity in the first variable.

The Jacobi identity is the most complicated condition and it is stated in numerous ways by different sources. The way we have stated it emphasizes the cyclic permutation of x, y, z . That can also be emphasized by the equivalent condition $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$ for all $x, y, z \in L$. Some folks prefer the “derivation/product rule” version: $[x, [y, z]] = [[x, y], z] + [y, [x, z]]$. It is a great exercise to use skew-symmetry to show all these versions are the same.

If L is an associative algebra, such as $M_n(\mathbb{F})$ or $\text{End}(V)$, then we can define a bracket on L by the commutator $[x, y] = xy - yx$. It’s easy to see this bracket is bilinear, alternating, and skew-symmetric. The Jacobi identity takes longer to verify, but careful calculation and persistence will show that it holds. In particular, a linear Lie algebra is a Lie algebra with the commutator as Lie bracket.

Every group of finite order n is isomorphic to a subgroup of $\mathbf{GL}(n, \mathbb{F})$ for any field $\mathbb{F}$. Likewise, such group is isomorphic to a subgroup of the symmetric group Sym_n . Either of these results is easy to prove. The situation for Lie algebras is similar, but the proof is much more complicated. For completeness we state this result; we may or may not get to

it. When $\mathbb{F}$ has characteristic 0, this was proved by Ado in 1935. Iwasawa extended it to positive characteristic, so it is sometimes known as the Ado-Iwasawa Theorem.

Proposition 2.1.1 (Ado’s theorem).

Every finite-dimensional Lie algebra is isomorphic to a linear Lie algebra.

REFERENCES AND REMARKS. See [Jacobson, Chapter VI, pp.199-206; Bourbaki 1-3, §I.7.3, pp.71-72].

PROOF. **Maybe later – or maybe not.** ■

1.2. Examples of linear Lie algebras. We begin with some examples of linear Lie algebras. We’ve already seen the first example, $\mathfrak{gl}(n)$, the **general linear Lie algebra**. The subspace of all matrices of trace 0 is called the **special linear Lie algebra** and denoted $\mathfrak{sl}(n)$. This is a Lie algebra, since $\text{tr}[A, B] = 0$ for all matrices A, B . The algebra $\mathfrak{sl}(\ell + 1)$ may also be denoted or referred to as the Lie algebra of type A_ℓ , for reasons we will explain later. Its dimension is $n^2 - 1$.

In general, the adjective “special” will refer to matrices of trace 0. Thus if we have defined the fantabulous Lie algebra, the special fantabulous Lie algebra will consist of the elements of the fantabulous Lie algebra of trace 0. For groups, “special” usually means of determinant 1.

The spaces of upper and strictly upper triangular matrices are closed under commutators, so we get the **upper triangular Lie algebra** $\mathfrak{t}(n)$ and **strictly upper triangular Lie algebra** $\mathfrak{n}(n)$ (the $\mathfrak{n}$ suggests nilpotent). The dimensions of these algebras are $1 + \dots + n$ and $1 + \dots + n - 1$ respectively. The space of skew-symmetric matrices is closed under commutators; the space of symmetric matrices is not. We call this the **orthogonal Lie algebra** and denote it by $\mathfrak{o}(n)$. Suppose A is skew-symmetric: then $\text{tr } A = -\text{tr } A^t = -\text{tr } A$, so $\text{tr } A = 0$. Thus we also call $\mathfrak{o}(n)$ the **special orthogonal Lie algebra** and denote it by $\mathfrak{so}(n)$. The dimension of $\mathfrak{so}(n) = \mathfrak{o}(n)$ is $1 + \dots + n - 1$. **Warning:** [Humphreys, §1.2, p.3] and other sources define the orthogonal Lie algebra differently. Their version is isomorphic to ours when $\mathbb{F}$ is algebraically closed, but the matrices that occur are different.

We will now introduce a very important batch of linear Lie algebras, each of them defined two ways. For the “abstract” version of these Lie algebras, suppose $\langle \cdot, \cdot \rangle$ is a bilinear form on V . We define the **skew-symmetric Lie algebra relative to $\langle \cdot, \cdot \rangle$** to be $\mathfrak{gl}(V, \langle \cdot, \cdot \rangle) = \{ A \in \mathfrak{gl}(V) \mid \langle Ax, y \rangle = -\langle x, Ay \rangle \text{ for all } x, y \in V \}$. This is a linear Lie algebra, since if $A, B \in \mathfrak{gl}(V, \langle \cdot, \cdot \rangle)$ and $x, y \in V$, then

$$\begin{aligned} \langle [A, B](x), y \rangle &= \langle ABx, y \rangle - \langle BAx, y \rangle = -\langle Bx, Ay \rangle + \langle Ax, By \rangle = \langle x, BAy \rangle - \langle x, AB y \rangle \\ &= -\langle x, (AB - BA)y \rangle = -\langle x, [A, B]y \rangle. \end{aligned}$$

When $\langle, \rangle$ is usual dot product on $\mathbb{F}^n$, the Lie algebra we obtain is none other than the orthogonal Lie algebra. To see this, recall that for the dot product form, we have $\langle x, y \rangle = x^T y$, and so $\langle Ax, y \rangle = x^T A^T y = \langle x, A^T y \rangle$. Thus $\langle Ax, y \rangle = \langle x, -Ay \rangle$ for all x, y is equivalent to $A = -A^T$. This hints at a “concrete” way to define our skew-symmetric Lie algebras.

If M is any $n \times n$ matrix, we can define $\mathfrak{gl}(n, M) = \{ A \in \mathfrak{gl}(n) \mid A^T M = -MA \}$. It's easy to check that the condition $A^T M = -MA$ is closed under taking commutators (that's why the $-$ is needed), and so $\mathfrak{gl}(n, M)$ is a subalgebra of $\mathfrak{gl}(n)$. For example, $\mathfrak{gl}(n, I) = \mathfrak{o}(n)$ and $\mathfrak{gl}(n, 0) = \mathfrak{gl}(n)$. If P is nonsingular, the matrices M and PMP^T are congruent and the Lie algebras $\mathfrak{gl}(n, M)$ and $\mathfrak{gl}(n, PMP^T)$ are isomorphic.

In general, a bilinear form $\langle, \rangle$ has a matrix M whose i, j entry is $\langle \mathbf{e}_i, \mathbf{e}_j \rangle$. In this case $\langle x, y \rangle = x^T M y$. The Lie algebras $\mathfrak{gl}(V, \langle, \rangle)$ and $\mathfrak{gl}(n, M)$ are isomorphic, so we can use either construction to describe them.

If $\mathbb{F}$ is algebraically closed, any two symmetric nondegenerate bilinear forms are isomorphic (i.e., any two symmetric nonsingular matrices are congruent). If $\mathbb{F} = \mathbb{R}$, there are $n + 1$ symmetric nondegenerate bilinear forms, given by the signature. One form is the positive definite form (the dot product), one is the negative definite form (the negative of the dot product), and the others are indefinite.

Suppose $n = 2\ell$ is even and $\mathbb{F}$ contains a square root of -1 (e.g., $\mathbb{F}$ is $\mathbb{C}$ but not $\mathbb{R}$). Then the dot product (matrix I) is isomorphic to a bilinear form with block matrix $M = \begin{pmatrix} 0 & I_\ell \\ I_\ell & 0 \end{pmatrix}$. [Humphreys, §1.2, p. 3] defines the orthogonal Lie algebra $\mathfrak{o}(n)$ via the form corresponding to M . Thus he defines $\mathfrak{o}(n)$ to be

$$\{ X \in \mathfrak{gl}(n) \mid X^T M = -MX \} = \left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \mid A, B, C, D \in M_\ell(\mathbb{F}), A^T = -D, B^T = -B, C^T = -C \right\}.$$

We call this the Lie algebra of type D_ℓ ; its dimension is $\ell(2\ell - 1) = \frac{12}{n}(n - 1)$.

If $n = 2\ell + 1$ is odd, then [Humphreys, §1.2, p. 3] uses the block matrix $M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & I_\ell \\ 0 & I_\ell & 0 \end{pmatrix}$ (where the 1 is 1×1) to define $\mathfrak{o}(n)$. If $\mathbb{F}$ contains a square root of -1 (e.g., $\mathbb{F}$ is $\mathbb{C}$ but not $\mathbb{R}$), then the form defined by M is isomorphic to the dot product form we have used to define $\mathfrak{o}(n)$. The result is that $\mathfrak{o}(2\ell + 1)$ à la [Humphreys] consists of

matrices $\begin{pmatrix} 0 & s & t \\ -t^T & A & B \\ -s^T & C & -A^T \end{pmatrix}$ with $B^T = -B, C^T = -C$, where A, B, C are $\ell \times \ell$ and s, t

are $1 \times \ell$. We call this the Lie algebra of type B_ℓ ; its dimension is $m(2m + 1) = \frac{12}{n}(n - 1)$.

A vector space has a nondegenerate alternating form if and only if its dimension is even. If the dimension is even, any two nondegenerate alternating forms are isomorphic (for any $\mathbb{F}$). We define the **symplectic Lie algebra** $\mathfrak{sp}(n)$ to be the skew-symmetric Lie algebra relative to this alternating form. To be concrete, let $n = 2\ell$. Then a nondegenerate alternating form can be give by the matrix $M = \begin{pmatrix} 0 & I_\ell \\ -I_\ell & 0 \end{pmatrix}$. [Humphreys, §1.2, pp. 2-3] then defines $\mathfrak{sp}(2\ell)$ to be

$$\{X \in \mathfrak{gl}(n) \mid X^T M = -MX\} = \left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \mid A, B, C, D \in M_\ell(\mathbb{F}), A^T = -D, B^T = B, C^T = C \right\}.$$

We call this the Lie algebra of type C_ℓ ; its dimension is $2\ell^2 + \ell = \frac{1}{2}n(n+1)$.

We want to discuss real Lie algebras, and we begin with an abstract construction. An **involution** on a field $\mathbb{E}$ is a field homomorphism $S : \mathbb{E} \rightarrow \mathbb{E}$ with the property that $S(S(a)) = a$ for all $a \in \mathbb{E}$. Common notations for $S(a)$ are $\bar{a}$ and a^* . Let $\mathbb{F} = \{a \in \mathbb{E} \mid S(a) = a\}$. We extend the involution to $\mathfrak{gl}(n)$ by defining A^* to be the matrix whose i, j entry is $S(a_{ji})$ (so we transpose A and apply S to the entries). We can similarly define A^* on $\mathfrak{gl}(V)$; we challenge the reader to carry this out. (One approach is to define a non-quite-bilinear dot product form on V as follows. If $v_1, \dots, v_n$ is a basis and $x = \sum_i a_i v_i$, $y = \sum_i b_i v_i$, define $\langle x, y \rangle = \sum_i a_i S(b_i)$. Then show A^* is defined by the rule $\langle Ax, y \rangle = \langle x, A^*y \rangle$.) The map $A \mapsto A^*$ is an operator on $\mathfrak{gl}(n, \mathbb{F})$ or $\mathfrak{gl}(V, \mathbb{F})$ that preserves addition but reverses multiplication: $(AB)^* = B^*A^*$. It still satisfies $(A^*)^* = A$ and it twists scalar multiplication: $(cA)^* = S(c)A^*$. (Thus $A \mapsto A^*$ is linear as an $\mathbb{F}$ -vector-space map. We define the **unitary Lie algebra** over $\mathbb{F}$ (*not* over $\mathbb{E}$) by $\mathfrak{u}(n) = \{A \in \mathfrak{gl}(n, \mathbb{F}) \mid A^* = -A\}$ and we define the **special unitary Lie algebra** by $\mathfrak{su}(n) = \{A \in \mathfrak{gl}(n, \mathbb{F}) \mid A^* = -A \text{ and } \text{tr } A = 0\}$.

The situation where the unitary Lie algebras will arise for us is where $\mathbb{E} = \mathbb{C}$, S is complex conjugation, and $\mathbb{F} = \mathbb{R}$. As noted above, $\mathfrak{u}(n)$ and $\mathfrak{su}(n)$ are real Lie algebras but not complex Lie algebras. Unlike some of the examples above, $\mathfrak{su}(n)$ is a proper subalgebra of $\mathfrak{u}(n)$. If $A \in \mathfrak{u}(n)$, then $\text{tr } S(A) = \overline{\text{tr } A}$, so we have $\overline{\text{tr } A} = -\text{tr } A$. This implies $\text{tr } A$ is purely imaginary, but not that it is 0. The reader should check the following equalities: $\mathfrak{u}(1) = i\mathbb{R}$, $\mathfrak{su}(1) = 0$, $\mathfrak{u}(2) = \left\{ \begin{pmatrix} ia & b+ic \\ -b+ic & id \end{pmatrix} \mid a, b, c, d \in \mathbb{R} \right\}$, and $\mathfrak{su}(2) = \left\{ \begin{pmatrix} ia & b+ic \\ -b+ic & -ia \end{pmatrix} \mid a, b, c \in \mathbb{R} \right\}$.

1.3. Lie algebras of low dimension. We now turn to abstract Lie algebras. For operators A, B , we have $AB = BA$ if and only if $[A, B] = 0$. Thus we say a Lie algebra is

Abelian if every bracket is 0. For each n , there is a unique n -dimensional Abelian Lie algebra.

The first non-Abelian Lie algebra occurs in dimension 2. It is a vector space L with basis x, y and bracket $[x, y] = ax + by$ for scalars a, b , not both zero. It is understood that $[x, x] = [y, y] = 0$ and $[y, x] = -ax - by$, and we won't mention these relations in future examples. Any choice of a, b satisfies the Jacobi identity and so gives a Lie algebra. However all choices of $(a, b) \neq (0, 0)$ yield isomorphic Lie algebras. To see this, assume that $a \neq 0$ and set $X = ax + by, Y = a^{-1}y$. Then $[X, Y] = [ax, a^{-1}y] + [by, a^{-1}y] = ax + by + 0 = X$. We call the Lie algebra with basis x, y and bracket $[x, y] = x$ the two-dimensional non-Abelian Lie algebra. (Some sources call this the two-dimensional solvable Lie algebra, by which they mean solvable but not Abelian.)

There are some interesting 3-dimensional Lie algebras. One is the **Heisenberg Lie algebra** with basis x, y, z and brackets $[x, y] = z, [x, z] = [y, z] = 0$.

Another example is the Lie algebra with basis x, y, z and brackets $[x, y] = z, [y, z] = x, [z, x] = y$. When $\mathbb{F} = \mathbb{R}$, this is the cross product Lie algebra structure on $\mathbb{R}^3$ (with $\mathbf{i} = x, \mathbf{j} = y, \mathbf{k} = z$). This Lie algebra is isomorphic to $\mathfrak{so}(3)$. To see this, let e_{ij} denote the matrix units. In $\mathfrak{gl}(3)$, let $x = e_{12} - e_{21}, y = e_{23} - e_{32}, z = e_{13} - e_{31}$. It's straightforward to check $[x, y] = z, [y, z] = x, [z, x] = y$.

We also have a Lie algebra with basis e, h, f and brackets $[e, f] = h, [h, e] = 2e, [h, f] = -2f$. This Lie algebra is isomorphic to $\mathfrak{sl}(2)$. To see this, set $e = e_{12}, h = e_{11} - e_{22}, f = e_{21}$. The set $\{e, h, f\}$ is clearly a basis for $\mathfrak{sl}(3)$ and it's straightforward to check $[e, f] = h, [h, e] = 2e, [h, f] = -2f$. If -1 has a square root in $\mathbb{F}$ (e.g., $\mathbb{F}$ is algebraically closed), then this Lie algebra is isomorphic to the Lie algebra of the last paragraph, i.e., $\mathfrak{sl}(2) \cong \mathfrak{so}(3)$, but otherwise they are not isomorphic. To show the isomorphism, suppose $i \in \mathbb{F}$ and $i^2 = -1$ and $\mathfrak{so}(3)$ is spanned by x, y, z as above. Set $e = x + iy, h = 2iz, f = -x + iy$. Then we have $[e, f] = -[x, x] + i[x, y] - i[y, x] - [y, y] = 2iz = h$ and $[h, e] = 2i[z, x] - 2[z, y] = 2iy + 2x = 2f$ and $[h, f] = -2i[z, x] - 2[z, y] = -2iy + 2x = 2(x - iy) = -2f$.

As to the non-isomorphism when $\mathbb{F} = \mathbb{R}$, we will see later that a special bilinear form, the Killing form, on $\mathfrak{so}(3)$ is negative definite, while the Killing form on $\mathfrak{sl}(2)$ is indefinite, so the Lie algebras cannot be isomorphic. Without relying on $\mathbb{F} = \mathbb{R}$, the eigenvalues of the Killing form of $\mathfrak{sl}(2)$ are $-4, 4, 8$ and of the Killing form of $\mathfrak{so}(3)$ are $-2, -2, -2$.

Note. $\mathfrak{sl}(2)$ is perhaps the most important Lie algebra, and it is ubiquitous in the theory of semisimple Lie algebra. We have used e, h, f for our standard basis of $\mathfrak{sl}(2)$, as in [Jacobson] and [Erdmann-Wildon]. [Humphreys] uses x, h, y for the standard basis

(our e is his x and our f is his y); [Bourbaki 7-9] and [Fulton-Harris] use X, H, Y . Everybody uses h or H for the element $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

As an extended exercise, consider the Lie algebra L with basis x, y, z and brackets $[x, y] = cz$, $[y, z] = ax$, $[z, x] = by$ for scalars a, b, c , and consider it as we move through these notes. We call this the a, b, c Lie algebra and denote it $L_{a,b,c}$. (This is not a standard name.) We will use some terminology from Section 3 (Types of Lie Algebras, p. 52) below. See Section 4 (The a, b, c Lie Algebras, p. 176) in Chapter 8 (Miscellaneous Calculations and Observations, p. 173) for details, worked-out calculations, and more information.

First show the Jacobi identity is satisfied. In checking the Jacobi identity, you might note several things. First, it is enough to check it on a basis. Second, if two of the variables are equal, the Jacobi identity will always hold (assuming bilinearity and skew-symmetry). Finally, if the identity holds for one order of entries, it holds for any permutation of that order.

The algebra breaks up into four cases, depending on how many of a, b, c are nonzero. Observe that L is Abelian if and only if $a = b = c = 0$.

If just one of a, b, c is nonzero, we can make an easy change of variables and show that L is the Heisenberg Lie algebra. This is the nilpotent non-Abelian case. If exactly one a, b, c is 0, then L is solvable but not nilpotent.

If all of a, b, c are nonzero, then L is simple (and so semisimple). If $\mathbb{F} = \mathbb{C}$, then L is isomorphic to $\mathfrak{sl}(2)$. If $\mathbb{F} = \mathbb{R}$, there are two possibilities for L , depending on the signs of a, b, c . If they are all of the same sign, L is isomorphic to the cross product algebra, $\mathfrak{so}(3)$. If they are of different signs, L is isomorphic to $\mathfrak{sl}(2)$.

For a complete classification of 3-dimensional Lie algebras, see [Jacobson, §I.4, pp. 11-14; Erdmann-Wildon, §3.2, pp. 20-25 when $\mathbb{F} = \mathbb{C}$].

2. Morphisms and Subobjects

2.1. Homomorphisms and Isomorphisms. We've defined an algebraic structure and now we need to define all the objects and machinery associated with it. If L, M are Lie algebras and $\phi : L \rightarrow M$ is a function, we say ϕ is a **Lie algebra homomorphism** or **map of Lie algebras** if ϕ is linear and $\phi([x, y]) = [\phi(x), \phi(y)]$ for all $x, y \in L$, where the bracket on the left is in L and the bracket on the right is in M . A homomorphism that is a bijection is called an **isomorphism**, and its inverse is also an isomorphism. We say L and M are **isomorphic** and write $L \cong M$ if there exists an isomorphism $\phi : L \rightarrow M$. Since ϕ must be linear, it's clear that isomorphic Lie algebras must have the same dimension.

It's easy to see that the composition of homomorphisms is a homomorphism and hence the composition of isomorphisms is an isomorphism. It follows that isomorphism of Lie algebras is an equivalence relation.

2.2. Subalgebras and Ideals. As a homomorphism $\phi : L \rightarrow M$ is a linear map, we have the usual objects $\ker \phi$ and $\operatorname{im} \phi$. What sort of objects are they? If $a, b \in \operatorname{im} \phi$, then $a = \phi(x), b = \phi(y)$ for some $x, y \in L$, so $[a, b] = [\phi(x), \phi(y)] = \phi([x, y]) \in \operatorname{im} \phi$. If $x \in L, y \in \ker \phi$, then $\phi([x, y]) = [\phi(x), 0] = 0$. We define a **subalgebra** of a Lie algebra L to be a subspace closed under the bracket and we define a **ideal** of L to be a subspace I that is “super-closed”, meaning $x \in L, y \in I$ implies $[x, y] \in I$. Note every ideal is a subalgebra. When I is an ideal of L , we write $I \triangleleft L$. The computations above show that $\operatorname{im} \phi$ is a subalgebra of M and $\ker \phi$ is an ideal of L .

If I is an ideal of L , we can define a Lie algebra structure on the vector space L/I in the usual way: $[x + I, y + I] = [x, y] + I$. The usual isomorphism and correspondence theorems hold for homomorphisms and ideals.

We can define the direct sum of Lie algebras in the usual way, with the Lie bracket defined coordinatewise. If L is a Lie algebra with subalgebras I, J and we want $L = I \oplus J$ as a Lie algebra, we need, as for any vector space, $I \cap J = 0$ and $I + J = L$. In addition, we need both I and J to be ideals of L . When one is an ideal and the other is (only) a subalgebra, we get a sort of semidirect product.

For subspaces $S, T \subseteq L$, we define $[S, T]$ to be the subspace of L spanned by all the $[s, t]$ with $s \in S, t \in T$. We call $[S, T]$ the **bracket** or **commutator** of S and T . It is not always the case that the elements of $[S, T]$ are Lie brackets $[s, t]$: they may be sums of Lie brackets. Suppose $s \in S, t \in T, x \in L$. Then $[x, [s, t]] = [[x, s], t] + [s, [x, t]]$. If S, T are ideals of L , the sum on the right hand side is in $[S, T]$. This shows $[S, T] \triangleleft L$ if $S, T \triangleleft L$ and clearly $[S, T] \subseteq S \cap T$ in this case. If $S = T$ is a subalgebra of L , then this calculation (taking $x \in S$) shows $[S, S]$ is a subalgebra of L . (Actually, it shows $[S, S]$ is an ideal of S .)

If I, J are ideals of L , all of $I + J, I \cap J, [I, J]$ are ideals of L . The corresponding result for subalgebras is not valid: if S, T are subalgebras of L , only $S \cap T$ is guaranteed to be a subalgebra. If one of S, T is an ideal or if $[S, T] = 0$, then $S + T$ is a subalgebra. (It's not enough to have $S \cap T = 0$.) We've just seen that $[S, S]$ is a subalgebra whenever S is.

If S is a subspace of the Lie algebra L , its **centralizer** is $C(S) = \{x \in L \mid [s, x] = 0 \text{ for all } s \in S\}$ and its **normalizer** is $N(S) = \{x \in L \mid [s, x] \in S \text{ for all } s \in S\}$. More generally, if T is a subspace of L , we say T **normalizes** S if $[S, T] \subseteq S$, i.e., $T \subseteq N(S)$, and we say T **centralizes** S if $[S, T] = 0$, i.e., $T \subseteq C(S)$. If S is a subalgebra, then $N(S)$

and $C(S)$ are subalgebras. Moreover, $N(S)$ is the largest subalgebra N of L such that S is an ideal in N . If S is an ideal of L , then $C(S)$ is an ideal. Finally, S is an ideal of L if and only if $N(S) = L$.

The centralizer and normalizer are special cases of the following notion. If A, B are subspaces of L , we define the **transporter** of B into A to be $(A : B) = \{x \in L \mid [x, B] \subseteq A\}$. This construction is discussed in Section 12 (The Transporter, p. 189) in Chapter 8 (Miscellaneous Calculations and Observations, p. 173), where other names for the same object are given. We note that $C(S) = (0 : S)$ and $N(S) = (S : S)$.

2.3. Derivations. There is another type of map that is important for Lie algebras. A function $\delta : L \rightarrow L$ (note domain=codomain) is called a **derivation** if it is linear and obeys the product rule $\delta([x, y]) = [x, \delta(y)] + [\delta(x), y]$ for all $x, y \in L$. If $\delta : L \rightarrow L$ is a derivation, then $\text{im } \delta$ is a subspace, but it need not be a subalgebra. Likewise, $\ker \delta$ is a subalgebra, but it need not be an ideal. (See the example below.) However, if I is an ideal of L , then $I + \delta(I)$ is an ideal of L .

If $x \in L$, we define $\text{ad } x : L \rightarrow L$ by $\text{ad } x(y) = [x, y]$. This is a derivation: it is clearly linear and one version of the Jacobi identity says precisely that $\text{ad } x([y, z]) = [\text{ad } x(y), z] + [y, \text{ad } x(z)]$. A derivation of the form $\text{ad } x$ is called an **inner derivation**.

If L is a Lie algebra, the set of all derivations on L is denoted $\text{Der } L$. It is a Lie subalgebra of $\mathfrak{gl}(L)$. To see this, suppose $\gamma, \delta \in \text{Der } L$ and let $x, y \in L$. Then

$$\begin{aligned} [\gamma, \delta]([x, y]) &= \gamma(\delta([x, y])) - \delta(\gamma([x, y])) = \gamma([x, \delta(y)] + [\delta(x), y]) - \delta([x, \gamma(y)] + [\gamma(x), y]) \\ &= [x, \gamma \circ \delta(y)] + [\gamma(x), \delta(y)] + [\delta(x), \gamma(y)] + [\gamma \circ \delta(x), y] \\ &\quad - [x, \delta \circ \gamma(y)] - [\delta(x), \gamma(y)] - [\gamma(x), \delta(y)] - [\delta \circ \gamma(x), y] \\ &= [x, \gamma \circ \delta(y) - \delta \circ \gamma(y)] + [\gamma \circ \delta(x) - \delta \circ \gamma(x), y] = [x, [\gamma, \delta](y)] + [[\gamma, \delta](x), y]. \end{aligned}$$

The following two formulas are very useful.

Proposition 2.2.1 (Operations on inner derivations).

Let L be a Lie algebra, let $x, y \in L$, let $\delta : L \rightarrow L$ be a derivation, and let ϕ be an automorphism of L . The brackets on the left hand side in (1) and (2) are in the Lie algebra $\text{Der } L \subseteq \mathfrak{gl}(L)$, while the left hand side of (3) is in $\mathfrak{gl}(L)$.

- (1) $[\text{ad } x, \text{ad } y] = \text{ad}[x, y]$.
- (2) $[\delta, \text{ad } x] = \text{ad } \delta(x)$.
- (3) $\phi \circ \text{ad } x \circ \phi^{-1} = \text{ad } \phi(x)$.

PROOF. It's easy to see that (1) follows from (2).

For (2), note $[\delta, \text{ad } x](y) = \delta([x, y]) - [x, \delta(y)] = [x, \delta(y)] + [\delta(x), y] - [x, \delta(y)] = \text{ad}(\delta(x))(y)$.

(3) We leave this to the reader. ■

The first formula above shows that the map $x \mapsto \text{ad } x$ from L to $\text{Der } L$ is a homomorphism of Lie algebras. The second formula shows that im ad is an ideal of $\text{Der } L$; it is sometimes denoted $\text{IDer } L$. It is easy to see that $\ker \text{ad} = Z(L) = \{x \in L \mid [x, y] = 0 \text{ for all } y \in L\}$. (We will discuss the ideal $Z(L)$, the center of L , later.)

If L is an Abelian Lie algebra, then any linear operator on L is a derivation, and only 0 is inner. We will show that if L is semisimple, the only derivations of L are inner. (See [Humphreys, §5.3, Theorem, p. 23; Jacobson, §III.6, Theorem 6, p. 74; Bourbaki 1-3, §I.6.2, Corollary 3, p. 51].) We will show in Section 8 (Derivations of the 2-dimensional and the Heisenberg Lie Algebras, p. 184) of Chapter 8 (Miscellaneous Calculations and Observations, p. 173) that all derivations of the two-dimensional non-Abelian Lie algebra are inner, even though it is not semisimple. On the other hand, a nonzero nilpotent Lie algebra always has a non-inner derivation. (See [Humphreys, §1, Exercise 9, p. 14; Jacobson, Chapter I, Exercise 4, p. 30; Bourbaki 1-3, Exercise I.4.8, p. 92]. I've seen this attributed to Schenkman, Jacobson, and Dixmier.)

Consider the following example. Let $L = \mathfrak{sl}(2)$ with the usual basis e, h, f discussed in Section 1 (Definition and Examples, p. 41). Consider the derivation $\delta = \text{ad } h$. Then $\text{im } \delta = \mathbb{F}e + \mathbb{F}f$ is not a subalgebra and $\ker \delta = \mathbb{F}h$ is a subalgebra but not an ideal.

Derivations are useful for many things. One of them is to generate automorphisms. Recalling our discussion of the matrix exponential, if δ is a derivation on L , we define $e^\delta = \exp \delta \in \mathfrak{gl}(L)$ by $\exp \delta(x) = \sum_{i=0}^{\infty} \frac{1}{i!} \delta^i(x)$ for $x \in L$. Does this definition make sense? If δ is nilpotent, say $\delta^n = 0$, then the seemingly infinite sum is actually a finite sum $\exp \delta(x) = \sum_{i=0}^{n-1} \frac{1}{i!} \delta^i(x)$. If $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$, then it is well-known that the exponential series converges absolutely for all matrices and so e^δ makes sense in this case for all derivations.

Proposition 2.2.2 (Exponential of a derivation).

Let L be a finite-dimensional Lie algebra over a field of characteristic 0 and let δ be a derivation on L . Then e^δ is an automorphism of L whenever it makes sense.

PROOF. It's easy to see that e^δ is linear when it's defined. To see that it is a Lie algebra homomorphism, we use the formula $\delta^n([x, y]) = \sum_{i=0}^n \binom{n}{i} [\delta^i(x), \delta^{n-i}(y)]$ from Subsection 9.1 (Powers of a derivation on a commutator, p. 185) in Chapter 8 (Miscellaneous Calculations and Observations, p. 173). Using this, we see

$$[e^\delta(x), e^\delta(y)] = \sum_{i,j} \left[\frac{1}{i!} \delta^i(x), \frac{1}{j!} \delta^j(y) \right] = \sum_n \frac{1}{n!} \sum_{i=0}^n \binom{n}{i} [\delta^i(x), \delta^{n-i}(y)] = \sum_n \frac{1}{n!} \delta^n([x, y]) = e^\delta([x, y]).$$

As we saw in the discussion of the matrix exponential, if δ, ε are commuting derivations, then $e^\delta \circ e^\varepsilon = e^{\delta+\varepsilon}$. In particular, $e^{-\delta}$ is the inverse of e^δ , so e^δ is an automorphism of L . $\blacksquare$

If ϕ is any automorphism of L , then $\phi \circ e^\delta \circ \phi^{-1} = e^{\phi \circ \delta \circ \phi^{-1}}$ and $\phi \circ e^{\text{ad } x} \circ \phi^{-1} = e^{\text{ad } \phi(x)}$.

[**Bourbaki 1-3**, §6.8, Definition 6, p. 63; **Bourbaki 7-9**, §VII.3, p. 20] calls automorphisms $e^{\text{ad } x}$ with $\text{ad } x$ nilpotent **special** and products of such automorphisms **elementary**. [**Humphreys**, §2.3, p. 9] calls automorphisms of the form $e^{\text{ad } x}$ with $\text{ad } x$ nilpotent **inner** and then extends the terminology to products of such automorphisms. The calculation in the last paragraph shows the set of elementary/inner automorphisms forms a normal subgroup of $\text{Aut } L$. [**Bourbaki 7-9**, §VII.3, Definition 1, p. 20] denotes this subgroup $\text{Aut}_e L$ and [**Humphreys**] denotes it $\text{Int } L$. [**Bourbaki 1-3**, §III.2, Definition 2, p. 310] uses $\text{Int } L$ for his subgroup of inner automorphisms, which is larger than the set we have described above. Perhaps other sources allow automorphisms of the form $e^{\text{ad } x}$, without the nilpotent requirement, or even e^δ , to be called inner. Other sources call an automorphism inner if it is induced by an inner automorphism of the Lie group that corresponds to L .

A derivation $\delta : L \rightarrow L$ is an operator in $\mathfrak{gl}(L)$, and so it has a Jordan-Chevalley decomposition $\delta = \delta_s + \delta_n$, where $\delta_s \in \mathfrak{gl}(L)$ is semisimple and $\delta_n \in \mathfrak{gl}(L)$ is nilpotent. We will now show δ_s, δ_n are derivations.

Proposition 2.2.3 (Jordan-Chevalley decomposition of derivations).

Let L be a finite-dimensional Lie algebra over a field of characteristic 0 and let δ be a derivation of L . Let $\delta = \delta_s + \delta_n$ be the Jordan-Chevalley decomposition of the operator $\delta \in \mathfrak{gl}(L)$. Then δ_s and δ_n are derivations.

REFERENCES AND REMARKS. See [**Humphreys**, §4.2, Lemma B, p. 18; **Bourbaki 7-9**, §VII.1.1, Proposition 4, p. 6; **Erdmann-Wildon**, Proposition 9.14, p. 86].

The proposition and the proof remain valid when L is any finite-dimensional algebra, associative or Lie or other type.

PROOF. Since linear combinations of derivations are derivations, it is enough to show either δ_s or δ_n is a derivation. The proof we give uses generalized eigenspaces and has similarities to the root space decomposition, except we do not assume δ_s is diagonalizable.

We assume δ is split. **General case?** For $\lambda \in \mathbb{F}$, let $V_{[\lambda]}$ be the generalized λ -eigenspace of δ , i.e., $V_{[\lambda]} = \{x \in L \mid (\delta - \lambda I)^k(x) = 0 \text{ for some } k\}$, where I is the identity map. By Proposition 1.1.4 (Eigenspace and generalized eigenspace decomposition, p. 22), $L = \bigoplus_{\lambda} V_{[\lambda]}$, where the sum is over the eigenvalues λ of δ . Each $V_{[\lambda]}$ is invariant under δ and hence under δ_s, δ_n (since they are polynomials in δ). The Jordan-Chevalley decomposition of $\delta|_{V_{[\lambda]}}$ is $\delta_s|_{V_{[\lambda]}} + \delta_n|_{V_{[\lambda]}}$, and so δ_s acts as scalar multiplication by λ

on $V_{[\lambda]}$. Let I denote the identity on L and let $x, y \in L$: we have the formula $(\delta - (\lambda + \mu)I)^k([x, y]) = \sum_{i=0}^k [(\delta - \lambda I)^{k-i}(x), (\delta - \mu I)^i(y)]$ from Subsection 9.2 (Powers of a derivation minus scalars on a commutator, p. 185) in Chapter 8 (Miscellaneous Calculations and Observations, p. 173). It follows that $[V_{[\lambda]}, V_{[\mu]}] \subseteq V_{[\lambda+\mu]}$ (just as for root space decompositions). If $x \in V_{[\lambda]}, y \in V_{[\mu]}$, then $\delta_s([x, y]) = (\lambda + \mu)[x, y]$ (since $[x, y] \in V_{[\lambda+\mu]}$). However, $[\delta_s(x), y] + [x, \delta_s(y)] = (\lambda + \mu)[x, y]$ as well. Linearity now shows δ_s is a derivation on $L = \oplus_{\lambda} V_{[\lambda]}$.

Here is an alternative approach that is unfinished. Decompose L as an $\mathbb{F}\delta$ -module, i.e., $L = \oplus_i L_i$ where each L_i is indecomposable with respect to δ -invariant submodule sums. Since δ_s, δ_n are polynomials in δ , each L_i is indecomposable with respect to δ_s and with respect to δ_n . Suppose **Of course, why can we?** that for each i , there is a scalar λ_i with $\delta_s(x) = \lambda_i x$ for $x \in L_i$. Then as above, can show $\delta_s([x, y]) = x\delta_s(y) + \delta_s(x)y$ for $x \in L_i, y \in L_j$. By linearity, we'll get δ_s is a derivation. ■

3. Types of Lie Algebras

3.1. Solvable and Nilpotent Lie algebras. An important concept in group theory and Galois theory is the notion of a solvable group. The general definition is as follows: the group G is solvable if there is a chain $1 = G_0 \triangleleft G_1 \triangleleft \cdots \triangleleft G_k = G$ such that each factor G_i/G_{i-1} is Abelian. When G is finite, we can replace “Abelian” with the stronger condition “cyclic”; in fact, that is the definition appropriate to solving equations by radicals. In general, we do not require each G_i to be normal in the whole group G , but we can find a chain with every $G_i \triangleleft G$ and G_i/G_{i-1} Abelian (use the derived series). If we can find a chain with every $G_i \triangleleft G$ and each G_i/G_{i-1} cyclic, we say G is **supersolvable**.

You may have seen a second definition of solvable via the derived series. There is another notion that is useful in the study of groups, namely nilpotent groups, which can also be defined via a certain type of series.

We will now define solvable and nilpotent Lie algebras, and it all starts with Abelian Lie algebras. Note that every 1-dimensional Lie algebra is Abelian and every Abelian Lie algebra is the direct sum of 1-dimensional Lie algebras. Note also that every subspace of an Abelian Lie algebra is an ideal (and hence a subalgebra). Suppose $\phi : L \rightarrow K$ is a linear map. If L is Abelian, then ϕ is a homomorphism of Lie algebras if and only if $\text{im } \phi$ is Abelian. If K is Abelian, then ϕ is always a homomorphism of Lie algebras.

We define the **center** of L to be $Z(L) = \{x \in L \mid [x, y] = 0 \text{ for all } y \in L\}$. Of course L is Abelian if and only if $Z(L) = L$. Obviously $Z(L)$ is an ideal of L .

First we give the chain definitions for solvable and nilpotent. We say a Lie algebra L is **solvable** if there is a chain $0 = L_0 \triangleleft L_1 \triangleleft \cdots \triangleleft L_k = L$ such that L_i/L_{i-1} is Abelian for all i . We do not require that each $L_i \triangleleft L$, but we can always find a series with this true

if L is solvable (as we will see below). If L is solvable, we can always find a chain with $\dim L_i/L_{i-1} = 1$ (but only $L_{i-1} \triangleleft L_i$). Such a chain is a full flag on L . If we can find a full flag where each $L_i \triangleleft L$, we say L is **completely solvable**. (This is the equivalent of supersolvable for groups. [Knapp] calls such Lie algebras **split solvable**.) We will show below that if $\mathbb{F}$ is algebraically closed, then every solvable Lie algebra is completely solvable.

We say a Lie algebra L is **nilpotent** if there is a chain $0 = L_0 \triangleleft L_1 \triangleleft \cdots \triangleleft L_k = L$ such that $L_i/L_{i-1} \subseteq Z(L/L_{i-1})$ for each i . This of course means each L_i/L_{i-1} is Abelian and so any nilpotent Lie algebra is solvable, but we can say more. If $x \in L_i, y \in L$, then $0 = [x + L_{i-1}, y + L_{i-1}] = [x, y] + L_{i-1}$, so we see that $[x, y] \in L_{i-1} \subseteq L_i$. Thus each L_i is automatically an ideal of L . In addition, we can insert links in the chain arbitrarily to get a full flag. Thus if L is nilpotent, there is a full flag $0 = L_0 \triangleleft L_1 \triangleleft \cdots \triangleleft L_n = L$ such that $L_i/L_{i-1} \subseteq Z(L/L_{i-1})$ for each i , and each $L_i \triangleleft L$. In particular, L is completely solvable.

We can also generalize in a different direction. Given a chain as in the definition of nilpotence, note that $L_1 \subseteq Z(L)$. If we replace L_1 by $Z(L)$ and adjust accordingly (meaning replace L_i by $L_i + Z(L)$), we can assume $L_1 = Z(L)$. We can now expand L_2 so that $L_2/L_1 = Z(L/L_1)$ and continue in this fashion. This leads to the following notion for any Lie algebra L : the **upper central series** or **ascending central series** of L is the chain $0 = Z_0 \triangleleft Z_1 \triangleleft \cdots \triangleleft Z_k \subseteq L$, where $Z_i/Z_{i-1} = Z(L/Z_{i-1})$ and $Z_k = Z_{k+1}$ (i.e., the center of L/Z_k is trivial). Our discussion above shows L is nilpotent if and only if $Z_k = L$. If L is not nilpotent, this series always terminates short of L . Thus we see that L is nilpotent if and only if the ascending central series reaches L . [Bourbaki 1-3, §I.1.6, p. 6] writes $\mathcal{C}_i L$ for what we have denoted Z_i . See Subsection 3.2 (Examples of derived and central series, p. 55) below for some examples.

For a simple example, consider the 2-dimensional non-Abelian Lie algebra, i.e., the Lie algebra L with basis x, y and $[x, y] = x$. The chain $0 \subseteq \mathbb{F}x \subseteq L$ is a full flag of ideals, so L is completely solvable. However, this is not a central chain. Indeed, the center of L is 0, and so the upper central series never gets off the ground. Hence L is not nilpotent.

As an exercise, consider the Heisenberg Lie algebra L with basis x, y, z and only nonzero bracket among these basis vectors $[x, y] = z$. Show that the ascending central series reaches L at step 2 (i.e., $Z_2 = L$).

We now define two descending series for L : the derived series and the lower or descending central series. See Subsection 3.2 (Examples of derived and central series, p. 55) below for some examples.

The subspace $[L, L]$ is called the **derived subalgebra** of L . It is an ideal, and it is denoted by L' or L^1 or $L^{(1)}$. (We'll explain the parentheses momentarily.) It's obvious

that L/L' is Abelian, and it's easy to see that $[L, L]$ is the largest ideal I with L/I Abelian. We say L is **perfect** if $L = [L, L]$. Proposition 2.7.12 (Semisimple Lie algebra equals its derived subalgebra, p. 78) will show that semisimple Lie algebras are perfect, but the example just before it shows that other Lie algebras may be perfect as well.

The **derived series** of L is defined as follows: we set $L^{(0)} = L$, $L^{(1)} = [L, L]$, $L^{(i+1)} = [L^{(i)}, L^{(i)}]$. It's not hard to see that $[L^{(i)}, L^{(j)}] = L^{(i+j)}$ for all i, j . [Bourbaki 1-3, §I.1.5, p. 6] writes $\mathcal{D}^i L$ for what we have denoted $L^{(i)}$ and writes $\mathcal{D}L$ for L^1 .

Each $L^{(i)}$ is an ideal of L . We will give an elegant proof of this in Section 5 (Characteristic Subspaces, p. 64), but for now we simply prove it by induction on i . We've observed this is true for $i = 0, 1$. Now suppose $L^{(i)}$ is an ideal of L and $x \in L, y, z \in L^{(i)}$. To show $L^{(i+1)}$ is an ideal, it is enough to show $[x, [y, z]] \in L^{(i+1)}$. We have $[x, [y, z]] = [[x, y], z] + [y, [x, z]]$. By induction, $[x, y], [x, z] \in L^{(i)}$ and so $[x, [y, z]] \in L^{(i+1)}$.

We can use the derived series to give another characterization of solvability.

Proposition 2.3.1 (Solvability and the derived series).

The Lie algebra L is solvable if and only if $L^{(k)} = 0$ for some k . If L is solvable and $\dim L = n$, then $L^{(n)} = 0$.

PROOF. The chain $L = L^{(0)} \supseteq L^{(1)} \supseteq L^{(2)} \supseteq \dots \supseteq L^{(k)} = 0$ demonstrates that L is solvable if $L^{(k)} = 0$. Suppose conversely that there is a chain $0 = L_0 \triangleleft L_1 \triangleleft \dots \triangleleft L_k = L$ such that L_i/L_{i-1} is Abelian for all i . We claim that $L^{(i)} \subseteq L_{k-i}$ for all i : then putting $i = k$ gives $L^{(k)} = 0$. To prove the claim, we note that it is true for $i = 0$. Suppose it is true for a given i . Then $L^{(i+1)} = [L^{(i)}, L^{(i)}] \subseteq [L_{k-i}, L_{k-i}]$. Since L_{k-i}/L_{k-i-1} is Abelian, we must have $[L_{k-i}, L_{k-i}] \subseteq L_{k-i-1} = L_{k-(i+1)}$. This proves the inductive step.

We leave the claim that $L^{(n)} = 0$ if $\dim L = n$ to the reader. ■

The **lower central series** or **descending central series** of L is defined as follows: we set $L^0 = L$, $L^1 = [L, L]$, $L^{i+1} = [L, L^i]$. Clearly $L^{i+1} \subseteq L^i$ for all i . As for the $L^{(i)}$, one can show each L^i is an ideal of L . [Bourbaki 1-3, §I.1.5, p. 6] writes $\mathcal{C}^i L$ for what we have denoted L^i .

Proposition 2.3.2 (Nilpotence and the central series).

Let L be a Lie algebra. Then the following conditions are equivalent.

- (1) L is nilpotent.
- (2) $Z_k = L$ for some k .
- (3) $L^k = 0$ for some k .

If L is nilpotent of dimension n , then $L^n = 0$ and $Z_n = L$.

PROOF. (1) $\iff$ (2) is discussed above.

(2) $\iff$ (3). We will show $Z_k = L$ if and only if $L^k = 0$. We do this by proving for any $i \geq 0, j \geq 1$, we have $L^i \subseteq Z_j \iff L^{i+1} \subseteq Z_{j-1}$. For $x, y \in L$, let $\bar{x}, \bar{y}$ denote the images of x, y in L/Z_{j-1} .

First suppose $L^i \subseteq Z_j$: to prove $L^{i+1} \subseteq Z_{j-1}$ it's enough to prove that for any $x \in L, y \in L^i$, we have $[x, y] \in Z_{j-1}$. Since $y \in L^i \subseteq Z_j$, we have $\bar{y} \in Z_j/Z_{j-1} = Z(L/Z_{j-1})$. This means $[\bar{x}, \bar{y}] = 0$, so $[x, y] \in Z_{j-1}$, as desired.

Suppose conversely that $L^{i+1} \subseteq Z_{j-1}$, let $y \in L^i$, and consider any $x \in L$. Then $[x, y] \in L^{i+1} \subseteq Z_{j-1}$, so $[\bar{x}, \bar{y}] = 0$. This proves $y \in Z_j$, as desired.

Fix a k and start with $i = 0, j = k$ and continue down to $i = k - 1, j = 1$. We get a string of equivalences that show $L^0 \subseteq Z_k \iff L^k \subseteq Z_0$, that is, $L = Z_k$ if and only if $L^k = 0$. ■

The next two results will be useful.

Proposition 2.3.3 (Ideals meet the center in nilpotent algebras).

Let L be a nilpotent Lie algebra. If I is a nonzero ideal of L , then $I \cap Z(L) \neq 0$.

PROOF. We define a sequence of ideals by $I_0 = I, I_1 = [I, L], \dots, I_{j+1} = [I_j, L], \dots$. Each I_j is an ideal contained in L^j and $I = I_0 \supseteq I_1 \supseteq \dots$. Since L is nilpotent, $L^j = 0$ for some j and so $I_j = 0$. Thus there must be a k with $I_k \neq 0 = I_{k+1}$. By definition, $[I_k, L] = 0$, so $I_k \subseteq I \cap Z(L)$. ■

Proposition 2.3.4 (Normalizers of ideals in nilpotent Lie algebras).

Let L be a nilpotent Lie algebra. If I is a proper ideal of L , the normalizer of I properly contains I .

PROOF. Since L is nilpotent and $I \subsetneq L$, there is a largest k with $Z_k \subseteq I$. Now $[Z_{k+1} + I, I] = [Z_{k+1}, I] + [I, I] \subseteq Z_k + I = I$. Thus $Z_{k+1} + I$ normalizes I and is not contained in I . ■

We will later give several tests that help us determine when a Lie algebra is nilpotent or solvable, such as Proposition 2.4.3 (Concrete Engel's theorem, p. 63), Proposition 2.4.4 (Engel's theorem for ad , p. 64), Proposition 3.2.6 (Concrete Lie's theorem, p. 100), Proposition 3.2.7 (Lie's theorem on complete solvability, p. 101), Proposition 2.6.2 (Concrete Cartan's criterion for solvability, p. 67).

3.2. Examples of derived and central series. We illustrate the above constructions in the case of some low-dimensional Lie algebras and then in the case of upper triangular matrices.

We noted above that if L is the 2-dimensional Lie algebra with basis x, y and $[x, y] = x$, then $L^{(1)} = \mathbb{F}x$, $L^{(2)} = 0$, and $L^i = \mathbb{F}x$ for all $i \geq 1$. Since $Z(L) = 0$, we have $Z_i(L) = 0$ for all i .

If $L = L_{a,b,c}$ is an a, b, c Lie algebra (with basis x, y, z and $[x, y] = cz$, $[y, z] = ax$, $[z, x] = by$), then L^1 is the span of ax, by, cz . When all of a, b, c are 0, we have L Abelian, whence $L^i = L^{(i)} = 0$ and $Z_i = L$ for $i > 0$.

If exactly one of a, b, c is nonzero, say $c \neq 0 = a = b$, then $L^1 = \mathbb{F}z$ and $L^i = L^{(i)} = 0$ for $i \geq 2$. Also, $Z_1 = \mathbb{F}z$ and $Z_i = L$ for $i \geq 2$.

If exactly two of a, b, c are nonzero, say $b, c \neq 0 = a$, then $L^i = \mathbb{F}y + \mathbb{F}z$ for $i \geq 1$ and $L^{(i)} = 0$ for $i \geq 2$. We have $Z_i = 0$ for all i .

If all three of a, b, c are nonzero, then $L^i = L^{(i)} = L$ and $Z_i = 0$ for all $i \geq 0$.

For further discussion of this example, see Section 4 (The a, b, c Lie Algebras, p. 176) in Chapter 8 (Miscellaneous Calculations and Observations, p. 173).

If $L = \mathfrak{sl}(2)$, then $L^i = L^{(i)} = L$ and $Z_i = 0$ for all $i \geq 0$. This follows from the facts that L is not Abelian and the only ideals are 0 and L .

In Proposition 8.2.1 (Multiplication of triangular matrices, p. 174), we define $T_0 = \mathfrak{t}(n)$ to be the space of upper triangular matrices and T_k for $k > 0$ to be the space of upper triangular matrices such that all entries of the main diagonal and the first $k - 1$ diagonals above the main diagonal are all 0. We define $T_k = 0$ for $k \geq n$. That is, $A \in T_k$ if and only if $a_{ij} = 0$ whenever $j < i + k$. For example, $T_1 = \mathfrak{n}(n)$.

Proposition 2.3.5 (Derived and central series of triangular matrices).

Let n be a positive integer and let T_i be as defined just before the proposition.

- (1) *If $i = j = 0$, then $[T_0, T_0] = T_1$. Otherwise, $[T_i, T_j] = T_{i+j}$.*
- (2) *Let $L = \mathfrak{t}(n)$: then $L^i = \mathfrak{n}(n)$ and $L^{(i)} = T_{2i-1}$ for all $i > 0$. Moreover, $Z_i = 0$ for all $i \geq 0$. In particular, $\mathfrak{t}(n)$ is solvable but not nilpotent.*
- (3) *Let $L = \mathfrak{n}(n)$: then $L^i = T_{i+1}$ and $L^{(i)} = T_{2i}$ for $i \geq 0$. Moreover, $Z_i = T_{n-i}$ for $0 \leq i \leq n - 1$. In particular, $\mathfrak{n}(n)$ is nilpotent.*

PROOF. Exercise. See Proposition 8.2.1 (Multiplication of triangular matrices, p. 174) and Proposition 8.2.2 (Centralizers and normalizers for triangular matrices, p. 175) for details. ■

3.3. The radical. Every Lie algebra has a largest solvable ideal and a largest nilpotent ideal, and a Lie algebra is semisimple if this largest solvable ideal is 0. We now state some results that enable us to prove the existence of these ideals. The first result has independent interest and will be used several times.

Proposition 2.3.6 (Solvable and nilpotent for ideals and quotients).

Let L be a Lie algebra, I an ideal of L , and S a subalgebra of L .

- (1) If L is solvable, then S and L/I are solvable.
- (2) If I and L/I are solvable, then L is solvable.
- (3) If L is nilpotent, then S and L/I are nilpotent.
- (4) If $I \subseteq Z(L)$ and L/I is nilpotent, then L is nilpotent.

REFERENCES AND REMARKS. See [Humphreys, §3.1, Proposition, p. 11; §3.2, Proposition, p. 12; Jacobson, §I.7, Lemma, p. 24; Bourbaki 1-3, §I.5.1, Proposition 1, p. 43; §I.4.1, Proposition 2, p. 39; Erdmann-Wildon, Lemma 4.4, p. 29; Lemma 4.9, p. 31].

The proof of (1), (3) at the start uses the following general fact: if $\phi : L \rightarrow M$ is a homomorphism of Lie algebras, then $\phi(L^{(i)} \subseteq M^{(i)}$ and $\phi(L^i \subseteq M^i$, with equality in both cases if ϕ is surjective.

PROOF. Clearly $S^{(i)} \subseteq K \cap L^{(i)}$ and $S^i \subseteq K \cap S^i$. If $\pi : L \rightarrow \bar{L} = L/I$ is the natural map, it's easy to see $\bar{L}^{(i)} \subseteq \pi(L^{(i)})$ and $\bar{L}^i \subseteq \pi(L^i)$. These observations prove (1) and (3).

(2) Since I is solvable, there is a chain $0 = I_0 \triangleleft I_1 \triangleleft \cdots \triangleleft I_k = I$ such that each I_j/I_{j-1} is Abelian. Since L/I is solvable, there is a chain $I = I_k \triangleleft I_{k+1} \triangleleft \cdots \triangleleft I_m = L$ such that each $(I_j/I)/(I_{j-1}/I)$ is Abelian. But $(I_j/I)/(I_{j-1}/I) \cong I_j/I_{j-1}$, so the chain $0 = I_0 \triangleleft I_1 \triangleleft \cdots \triangleleft I_m = L$ shows that L is solvable.

(4) Since L/I is nilpotent, there is a chain $I = I_1 \triangleleft I_2 \triangleleft \cdots \triangleleft I_m = L$ such that each $(I_j/I)/(I_{j-1}/I) \subseteq Z((L/I)/(I_{j-1}/I))$. This containment is the same as $I_j/I_{j-1} \subseteq Z(L/I_{j-1})$. Since $I/I_0 = I \subseteq Z(L) = Z(L/I_0)$, the chain $0 = I_0 \triangleleft I_1 = I \triangleleft I_2 \triangleleft \cdots \triangleleft I_k = I$ shows that L is nilpotent. ■

Note that (4) does not hold without the assumption that $I \subseteq Z(L)$. For example, the 2-dimensional non-Abelian Lie algebra is not nilpotent, but it has a 1-dimensional ideal with 1-dimensional quotient, so both ideal and factor are Abelian and hence nilpotent.

The following result shows that taking (finite) direct sums preserves solvability and nilpotence.

Proposition 2.3.7 (Direct sums of solvable and nilpotent algebras).

Let $L_1, \dots, L_n$ be Lie algebras and set $L = L_1 \oplus \cdots \oplus L_n$.

- (1) If $L_1, \dots, L_n$ are solvable, then L is solvable.
- (2) If $L_1, \dots, L_n$ are nilpotent, then L is nilpotent.

PROOF. We leave this to the reader. ■

The next result shows that sums of ideals preserve nilpotence or solvability. This is not true of sums of subalgebras.

Proposition 2.3.8 (Sums of solvable and nilpotent ideals).

Let L be a finite-dimensional Lie algebra. Any sum of solvable ideals of L is solvable. Any sum of nilpotent ideals of L is nilpotent.

REFERENCES AND REMARKS. See [Humphreys, §3.2, Proposition, p. 11 & §3, Exercise 6, p. 14; Jacobson, §I.7, Proposition 4, p. 24; §I.7, Proposition 6, p. 25; Bourbaki 1-3, §I.5.2, p. 44; Erdmann-Wildon, Lemma 4.4(c), p. 29].

PROOF. Since we deal with finite-dimensional vector spaces, any (possibly infinite) sum of subspaces is a sum of finitely many of them. Thus it suffices to prove the claims for the sum of two ideals I, J .

We have $(I + J)/J \cong I/(I \cap J)$. We now apply Proposition 2.3.6 (Solvable and nilpotent for ideals and quotients, p. 57). If I is solvable, then $I/(I \cap J)$ is solvable and hence $(I + J)/J$ is solvable. If also J is solvable, this implies $I + J$ is solvable.

The same proof for nilpotent does not quite work, since we can't conclude that $I + J$ is nilpotent given that $(I + J)/J$ is, without an extra hypothesis. We use an alternate approach. To illustrate it, let us compute $(I + J)^2$. Of course $(I + J)^1 = [I, I] + [J, J] + [I, J]$ and

$$\begin{aligned} (I + J)^2 &= [I + J, (I + J)^1] = [I, [I, I]] + [I, [J, J]] + [I, [I, J]] + [J, [I, I]] + [J, [J, J]] + [J, [I, J]] \\ &\subseteq [I, I] + [J, J] + [I, I] + [I, I] + [J, J] + [J, J] = I^1 + J^1. \end{aligned}$$

In general, $(I + J)^{2k} \subseteq I^k + J^k$. The reason for this is the following general statement. Suppose $I_1, \dots, I_m$ are ideals of L and at least $k + 1$ of them equal I . Then the bracket of these ideals, in any grouping, is contained in I^k : this can be proved by induction on k or m . **give details?** This implies the claim $(I + J)^{2k} \subseteq I^k + J^k$. Since I, J are solvable, we can choose k so that $I^k = J^k = 0$ and hence $(I + J)^{2k} = 0$. ■

In particular, the sum of all solvable ideals of L is a finite sum, and so by Proposition 2.3.8 (Sums of solvable and nilpotent ideals, p. 58), it is solvable. Said another way, every Lie algebra L has a solvable ideal that contains all solvable ideals; we call this ideal the **solvable radical** or just the **radical** of L . We denote it $\text{Rad } L$ or sometimes just R .

It is important to note that the radical of L only contains all solvable ideals; it need not contain all solvable subalgebras. As a trivial example, any 1-dimensional subspace of L is an Abelian subalgebra but is not generally an ideal, and hence need not be contained in the solvable radical.

A similar argument shows every Lie algebra L contains a **nil radical**, a nilpotent ideal that contains all nilpotent ideals. This is sometimes denoted $\text{Nil } L$. However, if N is the nil radical of L , it is still possible for L/N to have nonzero nilpotent ideals. We will eventually show L/N is always a reductive Lie algebra. As an example, consider again the

2-dimensional non-Abelian Lie algebra L , defined by $[x, y] = x$. The subspace $N = \mathbb{F}x$ is a nilpotent ideal, and it equals the nil radical (since L is not nilpotent). However, L/N is Abelian and not zero. As with the solvable radical, the nil radical need not contain all nilpotent subalgebras.

Warning: [Bourbaki 1-3, §I.5.3, Definition 3, p. 44] defines the **nilpotent radical** of a Lie algebra L to be the intersection of the kernels of all the finite-dimensional irreducible L -modules, that is, the largest ideal I such that I acts as 0 on each finite-dimensional irreducible representation of L . This nilpotent radical is contained in the nil radical of L , but they need not be equal.

3.4. Semisimple and reductive Lie algebras. We now consider algebras that are opposite to solvable in some sense. A Lie algebra L is said to be **simple** if L is not Abelian (hence nonzero) and the only ideals of L are 0 and L . We say a Lie algebra L is **semisimple** if has no nonzero Abelian ideals. (Thus we call the zero Lie algebra semisimple; some sources may exclude this case.) There are several equivalent definitions of semisimple; we list them just below. See [Humphreys, §3.2, p. 11; Jacobson, §I.7, p. 24; Bourbaki 1-3, §I.6.1, p. 50; Erdmann-Wildon, Definition 4.6, p. 31; Wikipedia] for the definitions of semisimple given in our references.

- L has no nonzero Abelian ideals (i.e., our definition of semisimple).
- The solvable radical of L is 0.
- L is a direct sum of simple Lie algebras. (Proposition 2.7.11 (Semisimple is equivalent to sum of simples, p. 77))
- The Killing form on L is nondegenerate. (Proposition 2.7.6 (Cartan's criterion for semisimplicity, p. 74))
- Every (finite-dimensional) L -module is completely reducible. (One direction is Proposition 3.5.3 (Weyl's theorem on complete reducibility, p. 118).)

We will prove the equivalence of the first two now and of the others later. The next result relies on an easy result from Section 5 (Characteristic Subspaces, p. 64); the reader may try a direct proof of that part.

Proposition 2.3.9 (Semisimple is equivalent to zero solvable radical).

A finite-dimensional Lie algebra is semisimple if and only if its solvable radical is 0.

PROOF. If L is not semisimple, it has a nonzero Abelian ideal and so its solvable radical is nonzero.

Suppose that L has nonzero solvable radical R . Then there is a positive integer k with $R^{(k)} = 0 \neq R^{(k-1)}$. Thus $R^{(k-1)}$ is a nonzero Abelian subalgebra; by Proposition 2.5.7 (Derived and central series of ideals are ideals, p. 66), it is an ideal of L . ■

By Proposition 2.3.6 (Solvable and nilpotent for ideals and quotients, p. 57), $L/\text{Rad } L$ has no nonzero solvable ideals and hence is semisimple for any Lie algebra L . We will see later (Levi's theorem) that there is always a semisimple subalgebra K of L with $L = \text{Rad } L \oplus K$ as a vector space. However, K need not be an ideal, so this is not a direct sum decomposition as a Lie algebra. (It is a semidirect product.)

We say L is **reductive** if every solvable ideal of L is contained in the center. This is plainly equivalent to saying that the solvable radical is the center or that $L/Z(L)$ is semisimple. Here are other equivalent characterizations that we will prove in Proposition 3.5.5 (Reductive algebras as direct sums, p. 121).

- L is the direct sum of a semisimple and an Abelian Lie algebra.
- L is the direct sum of simple Lie algebras and 1-dimensional Lie algebras.
- L is completely reducible as an $\text{ad } L$ -module.

Characterization (2) says roughly that reductive Lie algebras are what semisimple Lie algebras would be if we counted the 1-dimensional Lie algebra as simple.

4. Adjoint

The topic of this section is extremely important in the study of Lie algebras, but it introduces a level of abstraction that can sometimes be confusing. For linear Lie algebras, we will henceforth have to distinguish between the actual elements of the algebra and their adjoints, which are also linear operators.

4.1. Basics of adjoints. We repeat some of the material from Subsection 2.3 (Derivations, p. 49). If $x \in L$, we define $\text{ad } x : L \rightarrow L$ by $\text{ad } x(y) = [x, y]$. This is a derivation: it is clearly linear and one version of the Jacobi identity says precisely that $\text{ad } x([y, z]) = [y, \text{ad } x(z)] + [\text{ad } x(y), z]$. A derivation of the form $\text{ad } x$ is called an **inner derivation**. We can regard ad as a map $x \mapsto \text{ad } x$ from L to $\text{Der } L$. It's not hard to see that ad is a Lie algebra map with kernel $Z(L)$, i.e., we have $\text{ad}[x, y] = \text{ad } x \circ \text{ad } y - \text{ad } y \circ \text{ad } x$. (This is a good exercise.) Thus the image of ad is a subalgebra of $\text{Der } L$ and thus a subalgebra of $\mathfrak{gl}(L)$. In fact the image of ad is an ideal of $\text{Der } L$ (though not of $\mathfrak{gl}(L)$); this ideal is sometimes denoted $\text{IDer } L$. To see this, let $x \in L$ and let $\delta : L \rightarrow L$ be a derivation on L . For $y \in L$, we have $[\delta, \text{ad } x](y) = \delta[x, y] - [x, \delta(y)] = [x, \delta(y)] + [\delta(x), y] - [x, \delta(y)] = \text{ad } \delta(x)(y)$. Thus $[\delta, \text{ad } x] = \text{ad } \delta(x)$, and this shows $\text{im } \text{ad}$ is an ideal of $\text{Der } L$. (This gives another proof of the formula $\text{ad}[x, y] = \text{ad } x \circ \text{ad } y - \text{ad } y \circ \text{ad } x$.)

Warning: If S is a subalgebra of L and $x \in S$, then $\text{ad } x$ could have two meanings: either the adjoint action on L or the adjoint action on S . We will denote the former by $\text{ad}_L x$ and the latter by $\text{ad}_S x$. These are not the same; indeed they are not really comparable, since $\text{ad}_L x \in \mathfrak{gl}(L)$ and $\text{ad}_S(x) \in \mathfrak{gl}(S)$. However, we may convert them to

numbers; for example we may take the trace of $\text{ad } x$. In this case, it is not always obvious whether the numbers are the same. If S is an ideal of L , the $\text{tr ad}_L x = \text{tr ad}_S x$, but this is not true for general subalgebras.

We sometimes are interested in finding a linear realization (isomorphic representation) of a Lie algebra L . When $Z(L) = 0$, the adjoint can be used to accomplish this goal. In this case, if $x_1, \dots, x_n$ is a basis for L , then $\text{ad } x_1, \dots, \text{ad } x_n$, written as matrices relative to the basis $x_1, \dots, x_n$, provide an isomorphic representation of L as a subalgebra of $\mathfrak{gl}(n)$.

The subalgebra $\text{ad } L$ of $\mathfrak{gl}(L)$ can tell us a lot about L .

Proposition 2.4.1 (Properties of L from $\text{ad } L$).

Let L be a Lie algebra. Then the following statements hold.

- (1) L is Abelian if and only if $\text{ad } L = 0$.
- (2) L is nilpotent if and only if $\text{ad } L$ is nilpotent.
- (3) L is solvable if and only if $\text{ad } L$ is solvable.
- (4) L is reductive if and only if $\text{ad } L$ is semisimple.
- (5) L is semisimple if and only if $\text{ad } L$ is semisimple and $Z(L) = 0$.
- (6) L is simple if and only if $\text{ad } L$ is simple and $Z(L) = 0$.

PROOF. (1) is obvious.

(2), (3) This follows from Proposition 2.3.6 (Solvable and nilpotent for ideals and quotients, p. 57) and the fact that the kernel of ad is $Z(L)$.

(4) $\text{ad } L$ is semisimple if and only if $L/Z(L)$ is semisimple. Thus we need to show L is reductive if and only if $L/Z(L)$ is semisimple.

If $L/Z(L)$ contains a solvable ideal, it has the form $I/Z(L)$ where I contains $Z(L)$. Since $I/Z(L)$ and $Z(L)$ are solvable, we see that I is solvable. If L is reductive, then we must have $I \subseteq Z(L)$, i.e., $I/Z(L) = 0$.

Suppose conversely that $L/Z(L)$ is semisimple and I is a solvable ideal of L . Then $I + Z(L)$ is solvable, so the ideal $(I + Z(L))/Z(L)$ must be 0, i.e., we must have $I \subseteq Z(L)$. This proves L is reductive.

(5) This follows from (4).

(6) This is obvious, since $\text{ad } L \cong L$ when $Z(L) = 0$. ■

If $L \subseteq \mathfrak{gl}(V)$ is a linear Lie algebra and $x \in L$, we can ask about the relationship between properties of x as an operator and $\text{ad } x \in \mathfrak{gl}(L)$ as an operator. The next proposition gives us some information along those lines.

Proposition 2.4.2 (Semisimple/nilpotent and ad).

Let V be a finite-dimensional vector space, let $L \subseteq \mathfrak{gl}(V)$ be a linear Lie algebra, and let $x \in L$. Set $G = \mathfrak{gl}(V)$. Then the following statements hold.

- (1) If x is nilpotent, then $\text{ad}_L x$ and $\text{ad}_G x$ are nilpotent.
- (2) If x is diagonalizable, then $\text{ad}_L x$ and $\text{ad}_G x$ are diagonalizable.
- (3) If x is semisimple, then $\text{ad}_L x$ and $\text{ad}_G x$ are semisimple.

REFERENCES AND REMARKS. See [Humphreys, §3.2, Lemma, p. 12; §4.2, Lemma A & comments before it, p. 18; **Bourbaki 1-3**, §I.4.2, Lemma 1, p. 39; §I.5.4, Lemma 2, p. 47; **Erdmann-Wildon**, Exercise 1.17, p. 8; Lemma 5.1, p. 37].

The converse to the proposition is at least partially false, since the identity I is not nilpotent, but $\text{ad } I = 0$ is. For $\dim V = 1, 2$, the other converses are true and there is a characterization of when $\text{ad } x$ is nilpotent: see Section 10 (Jordan-Chevalley Computations, p. 186) in Chapter 8 (Miscellaneous Calculations and Observations, p. 173).

What about the converses in general?

PROOF. Since $\text{ad}_L x = (\text{ad}_G x)|_L$ and since all of these properties carry over to restrictions of operators, it is enough to prove the claims in the case $L = G$. Thus $\text{ad } x$ will mean $\text{ad}_G x$ below.

(1) Assume x is nilpotent. We give two proofs that $\text{ad } x$ is nilpotent. The proofs are essentially the same, but look different and are both appealing. First, define $\phi, \psi : \mathfrak{gl}(V) \rightarrow \mathfrak{gl}(V)$ by $\phi(y) = x \circ y$ and $\psi(y) = y \circ x$. Note that ϕ, ψ are nilpotent, since for example, $\phi^i(y) = x^i \circ y$ and x is nilpotent. Then $\phi \circ \psi(y) = (y \circ x) \circ y = y \circ (x \circ y) = \psi \circ \phi(y)$. For $y \in \mathfrak{gl}(V)$, we have $\text{ad } x(y) = \phi(y) - \psi(y)$, so $\text{ad } x = \phi - \psi$. It is easy to prove using the Binomial Theorem that the sum of commuting nilpotent operators is nilpotent. Since ϕ, ψ commute and are each nilpotent, it follows that $\text{ad } x$ is nilpotent.

For our second proof, we use the formula $(\text{ad } x)^n(y) = \sum_{i=0}^n (-1)^{n-i} \binom{n}{i} x^i y x^{n-i}$ from Subsection 9.3 (Powers of ad in $\mathfrak{gl}(V)$, p. 185) in Chapter 8 (Miscellaneous Calculations and Observations, p. 173). If $x^m = 0$ and we choose $n \geq 2m - 1$, then each term has either $i \geq m$ or $n - i \geq m$, so either x^i or x^{n-i} is 0. It follows that $(\text{ad } x)^n(y) = 0$.

(2) Assume x is diagonalizable, with $v_1, \dots, v_n$ a basis of eigenvectors, where $x(v_i) = \lambda_i v_i$. Let e_{ij} be the usual matrix units basis of G derived from $v_1, \dots, v_n$, that is, $e_{ij}(v_k) = \delta_{ik} v_j$. Then $\text{ad } x(e_{ij})(v_k) = x(e_{ij}(v_k)) - e_{ij}(x(v_k)) = \delta_{ik} x(v_j) - e_{ij}(\lambda_k v_k) = \delta_{ik}(\lambda_j - \lambda_k) v_j = \delta_{ik}(\lambda_j - \lambda_i) v_j$. This shows $\text{ad } x(e_{ij}) = (\lambda_j - \lambda_i) e_{ij}$, and this proves $\text{ad } x$ is diagonalizable.

(3) Suppose x is semisimple and set $\mathbb{E} = \overline{\mathbb{F}}$. Then the extension of x to $G_{\mathbb{E}}$ is diagonalizable. Thus $(\text{ad}_G x)_{\mathbb{E}} = \text{ad}_{G_{\mathbb{E}}} x$ is diagonalizable by (2). This shows that $\text{ad}_G x$ is semisimple. ■

Suppose $L \subseteq \mathfrak{gl}(V)$ is a linear Lie algebra and $x \in L$. We can write $x = s + n$ for $s, n \in \mathfrak{gl}(V)$ with s semisimple, n nilpotent, and with the rest of the properties of the Jordan-Chevalley decomposition. Of course, $\text{ad } x = \text{ad } s + \text{ad } n$ and by the above

proposition, $\text{ad } s$ is semisimple and $\text{ad } n$ is nilpotent. Moreover, $[\text{ad } s, \text{ad } n] = \text{ad}[s, n] = 0$, so this is the Jordan-Chevalley decomposition of $\text{ad } x$ in $\mathfrak{gl}(\mathfrak{gl}(V))$. But it is not clear (and may not be true **example**) that $s, n \in L$ — we only know they are in $\mathfrak{gl}(V)$. We will see in Proposition 3.5.4 (Abstract Jordan-Chevalley decomposition, p. 120) that when L is semisimple, then $s, n \in L$. The proof is not easy.

4.2. Engel's theorem. Although it is a consequence of Proposition 1.3.7 (Generalized Jacobson's theorem, p. 31), we state the usual concrete version of Engel's theorem for reference, and give a direct proof.

Proposition 2.4.3 (Concrete Engel's theorem).

Let V be a finite-dimensional vector space. If $L \subseteq \mathfrak{gl}(V)$ is a linear Lie algebra and each element of L is a nilpotent operator, then L is a nilpotent Lie algebra and L is strictly triangularizable.

REFERENCES AND REMARKS. See [Humphreys, §3.3, pp. 12-13; Jacobson, §II.3, p. 36; Bourbaki 1-3, §4.2, pp. 39-40; Erdmann-Wildon, §6.1, pp. 46-48; Radjavi-Rosenthal, Corollary 1.7.6, p. 23].

PROOF. If we can prove that L is strictly triangularizable, then L is isomorphic to a subalgebra of $\mathfrak{n}(n)$ and hence is nilpotent. One can then also use flags to directly show that L is nilpotent. The fact that L is strictly upper triangularizable follows from Proposition 1.3.7 (Generalized Jacobson's theorem, p. 31), but we give here the proof from [Humphreys, §3.3, pp. 12-13]. The proof below uses the adjoint action — can we eliminate the use of the adjoint?

Our goal is to find a nonzero $v \in V$ with $xv = 0$ for every $x \in L$. If we can do this, then we can define $\bar{V} = V/\mathbb{F}v$ and $\bar{L} = \{\bar{x} \mid x \in L\}$. The latter Lie algebra is a subalgebra of $\mathfrak{gl}(\bar{V})$ and each $\bar{x}$ is nilpotent, so we may apply induction to get a basis $v_2, \dots, v_n$ of $\bar{V}$ such that the matrices of the $\bar{x} \in \bar{L}$ are strictly upper triangular with respect to this basis. It then follows that the matrices of the $x \in L$ are strictly upper triangular with respect to the basis $v, v_2, \dots, v_n$.

Thus we assume that for any linear Lie algebra of dimension less than $n = \dim L$ has a common null vector. In particular, if K is a proper subalgebra of L , we can assume the space $W = \{v \in V \mid Kv = 0\}$ is nonzero. We will return to this set W later.

Again, let K be a proper subalgebra of L and consider $\phi : K \rightarrow \mathfrak{gl}(L/K)$ given by $\phi(y)(\bar{x}) = \overline{\text{ad } y(x)}$. This is well defined since K is a subalgebra, and by Proposition 2.4.2 (Semisimple/nilpotent and ad , p. 61), each $\text{ad}_L x$ is nilpotent, whence each $\phi(x)$ is nilpotent. Since $\dim \text{im } \phi < \dim L$, induction gives us an $x \in L \setminus K$ with $[y, x] + K = \phi(y)(\bar{x}) = 0$ for all $y \in K$. This says that $x \in N_L(K)$ and so we have proved $K \subsetneq N(K)$.

If we take K to be a maximal proper subalgebra of L , then we must have $N(K) = L$, i.e., K must be an ideal of L . Moreover, for any $z \in L \setminus K$, we see that $K + \mathbb{F}z$ is a subalgebra of L properly containing K . Thus $L = K + \mathbb{F}z$ and $\dim L/K = 1$.

We noted above that $W = \{v \in V \mid Kv = 0\}$ is nonzero. If $x \in L, y \in K, w \in W$, then $[x, y] \in K$ and so $y(xw) = x(yw) + [x, y]w = 0$. This shows W is L -invariant. Now take $z \in L$ with $L = K + \mathbb{F}z$, as in the previous paragraph. Since W is invariant under z , it makes sense to talk about $z|_W$. Furthermore, $z|_W$ is a nilpotent operator since z is a nilpotent operator. It follows that there is a nonzero $w \in W$ with $zw = 0$. By definition of W , we have $Kw = 0$ and so $Lw = 0$. This proves L has a common null vector. ■

We now give the “abstract” version of Engel’s theorem.

Proposition 2.4.4 (Engel’s theorem for ad).

Let L be a finite-dimensional Lie algebra. Then L is nilpotent if and only if $\text{ad } x$ is a nilpotent operator for every $x \in L$.

PROOF. If $\text{ad } x$ is nilpotent for each $x \in L$, then the set $\text{ad } L$ is nil and closed under the Lie bracket. By Proposition 1.3.8 (Strong Engel’s theorem, p. 33), the set $\text{ad } L$ is strictly upper triangularizable and so embeds in $\mathfrak{n}(n)$ for $n = \dim L$. Since $\mathfrak{n}(n)$ is a nilpotent Lie algebra, this implies $\text{ad } L$ is nilpotent as a Lie algebra. Thus L is nilpotent by Proposition 2.4.1 (Properties of L from $\text{ad } L$, p. 61)(2).

Suppose L is nilpotent. Note that $\text{ad } x(L^i) = [x, L^i] \subseteq L^{i+1}$. By induction, we see that $(\text{ad } x)^j(L^i) \subseteq L^{i+j}$. In particular $(\text{ad } x)^n(L) = (\text{ad } x)^n(L^0) \subseteq L^n = 0$. Thus $\text{ad } x$ is nilpotent. ■

5. Characteristic Subspaces

If G is a group and H is a subset (usually a subgroup), we say H is a characteristic subset if $\phi(H) \subseteq H$ for all automorphisms ϕ of G . If we apply this test with both ϕ and ϕ^{-1} , we see that in fact $\phi(H) = H$ for any characteristic subset. Since conjugation is an automorphism, any characteristic subgroup is normal. An example of a characteristic subgroup is the center.

When L is a Lie algebra, there are two notions of characteristic. We say a subspace $S \subseteq L$ is an **automorphism characteristic** subspace if $\phi(S) \subseteq S$ for any automorphism ϕ of L (and we get $\phi(S) = S$ for all ϕ , as for groups). We say a subspace $S \subseteq L$ is a **derivation characteristic** subspace if $\delta(S) \subseteq S$ for any derivation δ of L . Most sources simply use the name **characteristic** for “derivation characteristic”. Since $\text{ad } x$ is a derivation, any derivation characteristic subspace of L is an ideal. This example hints at why derivation characteristic for Lie algebras is like automorphism characteristic for groups.

One of the reasons derivation characteristic subspaces are so useful is the following observation, particular part (3).

Proposition 2.5.1 (Transitivity of characteristic).

Let L be a Lie algebra, let T be a subalgebra of L , and let S be a subspace of T .

- (1) *If S is a derivation characteristic subspace of T and T is a derivation characteristic subspace of L , then S is a derivation characteristic subspace of L .*
- (2) *If S is an automorphism characteristic subspace of T and T is an automorphism characteristic subspace of L , then S is an automorphism characteristic subspace of L .*
- (3) *If S is a derivation characteristic subspace of T and T is an ideal of L , then S is an ideal of L .*

PROOF. (1) and (2) are easy.

(3) Let $x \in L$. Then $\text{ad}_T x = (\text{ad } x)|_T$ is a derivation of T , whence S is invariant under $\text{ad}_T x$. This just says $[x, S] \subseteq L$ for all $x \in L$, and so S is an ideal of L . ■

We now give a second proof of the fact that $[S, T]$ is an ideal if S, T are, using the characteristic notion.

Proposition 2.5.2 (Operations on invariant subspaces).

Let S, T be subspaces of the Lie algebra L , let δ be a derivation on L , and let ϕ be an automorphism of L .

- (1) *If S, T are invariant under δ , then so are $S + T, S \cap T, [S, T]$.*
- (2) *If S, T are invariant under ϕ , then so are $S + T, S \cap T, [S, T]$.*

PROOF. We prove this just for $[S, T]$. Take $s \in S, t \in T$.

- (i) $\delta([s, t]) = [\delta(s), t] + [s, \delta(t)]$ and both summands are in $[S, T]$.
- (ii) $\phi([s, t]) = [\phi(s), \phi(t)] \in [S, T]$. ■

To see that $[S, T]$ is an ideal, apply the proposition to $\delta = \text{ad } x$ for all $x \in L$.

Proposition 2.5.3 (Operations on characteristic subspaces).

Let S, T be subspaces of the Lie algebra L .

- (1) *If S, T are derivation characteristic, then so are $S + T, S \cap T, [S, T]$.*
- (2) *If S, T are automorphism characteristic, then so are $S + T, S \cap T, [S, T]$.*

PROOF. This follows immediately from the previous proposition. ■

We now turn our attention to the derived series and the upper and lower central series.

Proposition 2.5.4 (The center is characteristic).

If L is a Lie algebra, then the center $Z(L)$ is derivation and automorphism characteristic.

PROOF. This is quite clear for automorphisms. Suppose δ is a derivation on L and $z \in Z(L), x \in L$. Then $0 = \delta([x, z]) = [\delta(x), z] + [x, \delta(z)] = [x, \delta(z)]$. This shows $\delta(z) \in Z(L)$. ■

Proposition 2.5.5 (Derived and central series are invariant).

Let L be a Lie algebra, let δ be a derivation on L , and let ϕ be an automorphism of L . Then the derived series and the descending and ascending central series are invariant under δ and ϕ . Explicitly, for every i we have $\delta(L^{(i)}) \subseteq L^{(i)}$, $\phi(L^{(i)}) = L^{(i)}$, $\delta(L^i) \subseteq L^i$, $\phi(L^i) = L^i$, $\delta(Z_i) \subseteq Z_i$, and $\phi(Z_i) = Z_i$.

PROOF. The statements about the derived series and the descending central series follow from Proposition 2.5.3 (Operations on characteristic subspaces, p. 65) and induction, and the statement about the ascending central series follows from Proposition 2.5.4 (The center is characteristic, p. 65) and induction. ■

Proposition 2.5.6 (Derived and central series are characteristic).

All terms in the derived series, ascending central series, and descending central series of a Lie algebra are derivation and automorphism characteristic.

PROOF. This follows immediately from Proposition 2.5.5 (Derived and central series are invariant, p. 66). ■

An even more general result is true, and this shows the power of the “characteristic” method.

Proposition 2.5.7 (Derived and central series of ideals are ideals).

Let I be an ideal of the Lie algebra L and let $I^k, I^{(k)}, Z_k(I)$ be the terms of the derived and upper and lower central series computed for the Lie algebra I . Then all of these spaces are ideals of L .

PROOF. By Proposition 2.5.6 (Derived and central series are characteristic, p. 66), these subspaces are characteristic subspaces of I . Since I is an ideal of L , we can now apply Proposition 2.5.1 (Transitivity of characteristic, p. 65) to conclude that they are all ideals of L . ■

6. Cartan’s Criterion for solvability

The following result, which has a complicated and interesting proof, is used in one proof of Cartan’s Criterion for solvability. (We will also give a proof for the real or complex numbers that does not use this result.)

Proposition 2.6.1 (Operators taking a space into a smaller space).

Let V be a finite-dimensional vector space over a field of characteristic 0. Let $A \subseteq B$ be subspaces of $\mathfrak{gl}(V)$ and set $M = (A : B) = \{x \in \mathfrak{gl}(V) \mid [x, B] \subseteq A\}$, a subalgebra of $\mathfrak{gl}(V)$. If $x \in M$ satisfies $\text{tr } xy = 0$ for all $y \in M$, then x is nilpotent.

REFERENCES AND REMARKS. See [Humphreys, §4.3, Lemma, p. 19; Bourbaki 1-3, §I.5.4, Lemma 3, p. 47].

See Section 12 (The Transporter, p. 189) in Chapter 8 (Miscellaneous Calculations and Observations, p. 173) for more on the $(A : B)$ construction.

PROOF. Set $G = \mathfrak{gl}(V)$. Assume $x \in M$ is split and satisfies $\text{tr } xy = 0$ for all $y \in M$. Let $x = x_s + x_n$ be the Jordan-Chevalley decomposition of x , so x_s is diagonalizable. Let $v_1, \dots, v_n$ be a basis of eigenvectors of V with corresponding eigenvalues $\lambda_1, \dots, \lambda_n$. Note that with respect to this basis, x_n is strictly upper triangular. **check** Let E be the $\mathbb{Q}$ -subspace of $\mathbb{F}$ spanned by $\lambda_1, \dots, \lambda_n$. Our goal is to show each $\lambda_i = 0$ and hence $E = 0$. Suppose this is not the case: then there is a nonzero $\mathbb{Q}$ -linear functional $\phi : E \rightarrow \mathbb{Q}$. We will derive a contradiction from the existence of ϕ .

Let $y \in G$ be defined by $yv_i = \phi(\lambda_i)v_i$ for $i = 1, \dots, n$. Let e_{ij} be the usual basis of G with respect to $v_1, \dots, v_n$, i.e., $e_{ij}(v_k) = \delta_{ik}v_j$. As in the proof of Proposition 2.4.2 (Semisimple/nilpotent and ad , p. 61), we see that e_{ij} is an eigenvector for $\text{ad } x_s$ with eigenvalue $\lambda_i - \lambda_j$ and also an eigenvector for $\text{ad } y$ with eigenvalue $\phi(\lambda_i) - \phi(\lambda_j) = \phi(\lambda_i - \lambda_j)$.

Proposition 1.4.5 (The Chinese remainder theorem and Lagrange interpolation, p. 38) tells us there is a polynomial $h \in \mathbb{F}[t]$ with $h(\lambda_i - \lambda_j) = \phi(\lambda_i - \lambda_j)$ for all i, j . In particular, $h(0) = \phi(\lambda_i - \lambda_i) = 0$. Since $\text{ad } x_s, \text{ad } y$ are diagonal, we see that $\text{ad } y = h(\text{ad } x_s)$. The discussion after Proposition 2.4.2 (Semisimple/nilpotent and ad , p. 61) tells us that $\text{ad } x = \text{ad } x_s + \text{ad } x_n$ is the Jordan-Chevalley decomposition of $\text{ad } x$, and so in particular, $\text{ad } x_s = g(\text{ad } x)$ for some $g \in \mathbb{F}[t]$ with $g(0) = 0$. Thus $\text{ad } y = (h \circ g)(\text{ad } x)$ and $f = h \circ g$ has zero constant term. We know $[x, b] \in A$, so $(\text{ad } x)^2(b) = [x, [x, b]] \in [x, A] \subseteq A$. By induction, we see each $(\text{ad } x)^k(b) \in A$. Thus $[y, b] = \text{ad } y(b) = (f(\text{ad } x))(b) \in A$, and so $y \in M$. This means $\text{tr } xy = 0$.

Since x_n is strictly upper triangular and y is diagonal with respect to a basis, $x_n y$ is still strictly upper triangular, and so has trace 0. Thus $0 = \text{tr } xy = \text{tr } x_s y + \text{tr } x_n y = \text{tr } x_s y = \sum_{i=1}^n \lambda_i \phi(\lambda_i)$. This is an equation in the vector space E , and we can apply ϕ to it to obtain $0 = \phi(\sum_{i=1}^n \lambda_i \phi(\lambda_i)) = \sum_{i=1}^n \phi(\lambda_i)^2$. The $\phi(\lambda_i)^2$ are nonnegative rational numbers, so they must all be 0. But we chose ϕ not identically 0, so this is a contradiction.

Thus all eigenvalues of x are 0 and so x is nilpotent.

We can now use extension of scalars for the general case. **Details?** ■

Proposition 2.6.2 (Concrete Cartan's criterion for solvability).

Let V be a finite-dimensional vector space over a field of characteristic 0 and let $L \subseteq \mathfrak{gl}(V)$ be a linear Lie algebra. If either of the following conditions holds, then L is solvable.

- (1) $\operatorname{tr} x^2 = 0$ for all $x \in [L, L]$.
- (2) $\operatorname{tr} xy = 0$ for all $x \in L, y \in [L, L]$.

REFERENCES AND REMARKS. See [Humphreys, §4.3, Theorem, p. 20; Jacobson, Section III.4, Corollary, pp. 68-69; Bourbaki 1-3, §I.5.4, Theorem 2, p. 48; Erdmann-Wildon, Proposition 9.3, p. 79; Fulton-Harris, Appendix §C.1, Theorem C.5, p. 479; Wikipedia].

Since condition (1) in the proposition is a special case of condition (2), the reader may wonder why we have included condition (2). When checking these conditions, it is enough to check condition (2) for x in a basis of $[L, L]$ and y in a basis of L , but condition (1) must be checked for all $x \in [L, L]$. Moreover, this proposition is normally stated using condition (2). The sufficiency of condition (1) follows from [Jacobson, Section III.4, p. 68], which has a quite different proof from the one we give.

These conditions can be phrased in terms of the concrete trace form τ . (1) says $\tau(x, x) = 0$ for all $x \in [L, L]$ and (2) says that $[L, L]$ is contained in the radical of the form τ .

We will give two proofs of Cartan's criterion; the second one does not use Proposition 2.6.1 (Operators taking a space into a smaller space, p. 67), but it assumes $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$. In both cases, we prove $L^{(2)}$ is nilpotent. Could it be the case that L^1 is actually nilpotent under these hypotheses?

PROOF. We will show $L^{(2)}$ is nilpotent, which certainly implies L is solvable. We will do this by showing every $x \in L^{(2)}$

To show L is solvable, it is enough to prove $L^{(1)} = [L, L]$ is solvable. We will actually show $L^{(2)}$ is nilpotent, which Our goal below is to show $[L, L]$ is nilpotent (and hence solvable) by showing every $x \in [L, L]$ is nilpotent and applying Engel's Theorem.

(1) Replacing L by $[L, L]$, we can assume $\operatorname{tr} x^2 = 0$ for all $x \in L$. This in turn implies that for any $x, y \in L$, we have $0 = \operatorname{tr}(x + y)^2 = \operatorname{tr} x^2 + \operatorname{tr} xy + \operatorname{tr} yx + \operatorname{tr} y^2 = 2 \operatorname{tr} xy$. Thus we can assume $\operatorname{tr} xy = 0$ for all $x, y \in L$, which means (2) holds. We will show every $x \in [L, L]$ (which is the original $L^{(2)}$) is nilpotent. Proposition 2.4.3 (Concrete Engel's theorem, p. 63) then shows $[L, L]$ is nilpotent and hence L is solvable.

We will apply Proposition 2.6.1 (Operators taking a space into a smaller space, p. 67) with $A = [L, L]$, $B = L$, $M = \{x \in \mathfrak{gl}(V) \mid [x, L] \subseteq [L, L]\}$. Clearly $L \subseteq M$ but M could be bigger. Suppose $y, w \in L, z \in M$. Then $[w, z] \in [L, L]$ by definition of M , so by condition (2), $0 = \operatorname{tr} y[w, z] = \operatorname{tr}[y, w]z$. It follows that $\operatorname{tr} xz = 0$ for every $x \in [L, L]$. This

is true for all $z \in M$, so by Proposition 2.6.1 (Operators taking a space into a smaller space, p. 67), every $x \in [L, L]$ is nilpotent.

Here is another proof of (2), for $\mathbb{F} = \mathbb{C}$, that doesn't use Proposition 2.6.1 (Operators taking a space into a smaller space, p. 67). It is taken from [Fulton-Harris, Appendix §C.1, Theorem C.5, p. 479] and also appears in [Erdmann-Wildon, Proposition 9.3, p. 79]. It also applies when $\mathbb{F} = \mathbb{R}$ by extension of scalars. Let $x \in [L, L]$; as in the first proof above, we wish to show x is nilpotent. Choose a basis of V so that the Jordan-Chevalley decomposition $x = x_s + x_n$ is the usual Jordan form, so $x_s = \text{diag}(\lambda_1, \dots, \lambda_n)$ and x_n is strictly upper triangular. Set $w = \text{diag}(\bar{\lambda}_1, \dots, \bar{\lambda}_n)$, where $\bar{\lambda}$ is the complex conjugate of λ . Then $xw = x_s w + x_n w$ is an upper triangular matrix with diagonal entries $\lambda_i \bar{\lambda}_i = |\lambda_i|^2$. Thus $\text{tr } xw = \sum_i |\lambda_i|^2$. We will show $\text{tr } xw = 0$, and this implies each $\lambda_i = 0$. This in turn implies $x = x_n$ is nilpotent, as desired.

If $w \in L$, our hypothesis immediately gives us $\text{tr } xw = 0$. However, we don't know $w \in L$, so we have to do some work. Let us say $x = \sum_j [y_j, z_j]$ for $y_j, z_j \in L$. Then $\text{tr } xw = \sum_j \text{tr } [y_j, z_j]w = \sum_j \text{tr } y_j [z_j, w]$ by the invariance of the concrete trace form. We will show $[z_j, w] \in [L, L]$ and our hypothesis then implies $\text{tr } y_j [z_j, w] = 0$ and hence $\text{tr } xw = 0$ as desired.

To see $[z_j, w] \in [L, L]$, first note that $\text{ad } x = \text{ad } x_s + \text{ad } x_n$ is the Jordan-Chevalley decomposition of $\text{ad } x$ by Proposition 2.4.2 (Semisimple/nilpotent and ad , p. 61) and the remarks after the proof. Thus $\text{ad } x_s = f(\text{ad } x)$ for some polynomial $f \in \mathbb{F}[t]$ with $f(0) = 0$. The proof of Proposition 2.4.2 (Semisimple/nilpotent and ad , p. 61) also shows that $\text{ad } x$ and $\text{ad } w$ are diagonalizable with the same bases e_{ij} in $\mathfrak{gl}(\mathfrak{gl}(V))$ and with eigenvalues $\lambda_i - \lambda_j$ and $\bar{\lambda}_i - \bar{\lambda}_j$. Furthermore, Proposition 1.4.5 (The Chinese remainder theorem and Lagrange interpolation, p. 38) gives us a polynomial $g \in \mathbb{F}[t]$ with $g(\lambda_i - \lambda_j) = \bar{\lambda}_i - \bar{\lambda}_j$ for each i , and with $g(0) = 0$. Thus $g(\text{ad } x_s) = \text{ad } w$, and so $h(\text{ad } x) = \text{ad } w$, where $h = g \circ f$ is a polynomial with $h(0) = 0$. Thus $[z_j, w] = -\text{ad } w(z_j) = -h(\text{ad } x)(z_j)$. Since h has no constant term, $h(\text{ad } x)(z_j)$ is a linear combination of elements $(\text{ad } x)^i(z_j)$ for $i > 0$. Each such term is a commutator and so is in $[L, L]$. Thus $[z_j, w] = -h(\text{ad } x)(z_j) \in [L, L]$. This completes the proof. ■

We now use ad to turn Cartan's criterion for linear Lie algebras until a criterion for all Lie algebras.

Proposition 2.6.3 (Cartan's criterion for solvability).

Let L be a finite-dimensional Lie algebra over a field of characteristic 0. Then the following conditions are equivalent.

- (1) L is solvable.

- (2) $\text{tr}(\text{ad } x)^2 = 0$ for all $x \in [L, L]$.
 (3) $\text{tr } \text{ad } x \text{ ad } y = 0$ for all $x \in L, y \in [L, L]$.

REFERENCES AND REMARKS. See [Humphreys, §4.3, Corollary, p. 20; Jacobson, Section III.4, Corollary, pp. 68-69; Bourbaki 1-3, §I.5.4, Theorem 2, p. 48; Erdmann-Wildon, Theorem 9.4, p. 80; Wikipedia].

See our remarks on Proposition 2.6.2 (Concrete Cartan's criterion for solvability, p. 67). There is an alternative statement of (3), which many sources use: $[L, L]$ is contained in the radical of the Killing form (see Section 7 (Invariant Forms, p. 70)).

PROOF. Clearly (3) $\Rightarrow$ (2). By Proposition 2.6.2 (Concrete Cartan's criterion for solvability, p. 67), (2) implies $\text{ad } L$ is solvable. Thus L is solvable by Proposition 2.4.1 (Properties of L from $\text{ad } L$, p. 61).

Suppose (1) holds and $\mathbb{F}$ is algebraically closed. Then $\text{ad } L$ is solvable, so by Proposition 3.2.6 (Concrete Lie's theorem, p. 100) (in Chapter 3 (Representations, p. 83)), $\text{ad } L$ is isomorphic to a subalgebra of $\mathfrak{t}(n)$ for $n = \dim V$. Thus $\text{ad}[L, L] = [\text{ad } L, \text{ad } L]$ is isomorphic to a subalgebra of $\mathfrak{n}(n)$. This implies $\text{ad } x \text{ ad } y$ is nilpotent for all $x, y \in L$, and so $\text{tr } \text{ad } x \text{ ad } y = 0$. Thus (3) holds. ■

7. Invariant Forms

If $\langle, \rangle$ is a bilinear form on a Lie algebra L , we say $\langle, \rangle$ is **associative** or **invariant** if $\langle[x, y], z\rangle = \langle x, [y, z]\rangle$ for all $x, y, z \in L$. ([Humphreys] and [Erdmann-Wildon] use “associative”, [Bourbaki 1-3] uses “invariant”, and [Jacobson] uses both.) The existence of such forms, particularly nondegenerate ones, can have consequences for the structure of L . If $\langle\delta(x), y\rangle + \langle x, \delta(y)\rangle = 0$ for all $x, y \in L$ and all derivations δ on L , we say $\langle, \rangle$ is **completely invariant** (see [Bourbaki 1-3, §I.3.6, p. 33]). Completely invariant forms are invariant. For example, the Killing form is completely invariant: see Proposition 2.7.3 (Killing form is completely invariant, p. 72).

If L is a simple Lie algebra and $\mathbb{F}$ is algebraically closed, then any two nondegenerate invariant symmetric bilinear forms are scalar multiples of each other. See [Humphreys, §6, Exercise 6, p. 31; Jacobson, Chapter III, Exercise 9, p. 104; Bourbaki 1-3, Exercise I.6.18, p. 105; Erdmann-Wildon, Chapter 9, Exercise 9.11, p. 89; Wikipedia]. Extended hint: If $\langle, \rangle, \langle, \rangle'$ are such forms, define $\theta, \theta' : L \rightarrow L^*$ by $\theta(x)(y) = \langle x, y\rangle$ and $\theta'(x)(y) = \langle x, y\rangle'$. The maps θ, θ' are linear isomorphisms since the forms $\langle, \rangle, \langle, \rangle'$ are nondegenerate; check that they are module isomorphisms. Consider L as an irreducible module over itself and apply Schur's Lemma to get a scalar λ such that $\theta^{-1} \circ \theta'$ acts like multiplication by λ . Now show $\langle, \rangle = \lambda \langle, \rangle'$.

This result may fail if $\mathbb{F}$ is not algebraically closed or if L is not simple. It always fails if L is semisimple but not simple.

7.1. The Killing form. We are now going to introduce a very important invariant form on linear Lie algebras and a related form on any Lie algebra. These forms will be used heavily throughout the rest of these notes. If L is a linear Lie algebra, the **concrete trace form** on L defined by $\tau(x, y) = \text{tr } xy$ is a symmetric invariant form. If L is any Lie algebra, one example of a symmetric invariant form on L is the **Killing form** or **Killing-Cartan form**, defined by $\kappa(x, y) = \text{tr}(\text{ad } x \circ \text{ad } y)$. To see that it is invariant, imitate the proof for the concrete trace form (or use the concrete trace form for $\text{ad } L$). We will show in Proposition 2.7.6 (Cartan's criterion for semisimplicity, p. 74) that L is semisimple if and only if the Killing form on L is nondegenerate.

Some examples of the trace and Killing forms can be found in Section 5 (Killing and Trace Form Computations, p. 181) and Subsection 4.4 (The Killing form and trace form, p. 179) in Chapter 8 (Miscellaneous Calculations and Observations, p. 173). For

$\mathfrak{sl}(2)$, the matrix of the trace form is $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ and the matrix of the Killing form is

$\begin{pmatrix} 0 & 0 & 4 \\ 0 & 8 & 0 \\ 4 & 0 & 0 \end{pmatrix}$. When L is the Heisenberg Lie algebra defined by $[x, y] = z$, $[x, z] = [y, z] = 0$, the Killing form is identically 0.

Here is a result relating the Killing and trace forms for several series of Lie algebras. We will not prove it.

Proposition 2.7.1 (Killing form for classical Lie algebras).

Let L be a Lie algebra over a field $\mathbb{F}$ of characteristic 0 with Killing form κ and trace form τ .

- (1) *If $L = \mathfrak{sl}(n)$, then $\kappa = 2n\tau$.*
- (2) *If $L = \mathfrak{so}(n)$, then $\kappa = (n - 2)\tau$.*
- (3) *If $L = \mathfrak{sp}(n)$ (n even), then $\kappa = (n + 2)\tau$.*

REFERENCES AND REMARKS. See [Humphreys, §6, Exercise 7, p. 31; Bourbaki 7-9, §VIII.Exercises.13, Exercise 12, p. 268; Erdmann-Wildon, §12.6, p. 137 and Exercise 12.3, p. 13; Wikipedia].

PROOF. Omitted. ■

We will often need to consider the Killing form on a subspace S of L . There are two possibilities here. First, we can take the original Killing form $\kappa = \kappa_L$ and simply restrict

its inputs to S : we denote this $\kappa|_S$. If S is a subalgebra, we can also consider the Killing form on the Lie algebra S : we denote this κ_S . The two forms need not be equal. For example, if L is the 2-dimensional non-Abelian Lie algebra with basis x, y and $[x, y] = x$, then the Killing form for L (relative to the basis x, y) is $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. Let S be the subalgebra spanned by y . Then $\kappa|_S = (1)$. On the other hand, S is an Abelian subalgebra, so $\kappa_S = 0$. If we choose T to be the span of x , however, we'd get $\kappa|_T = \kappa_T = 0$. This is explained by the following result.

Proposition 2.7.2 (Killing form on an ideal).

Let L be a finite-dimensional Lie algebra with Killing form κ and let I be an ideal of L .

- (1) $\kappa|_I = \kappa_I$.
- (2) *The radical of $\kappa|_I$ is $I \cap I^\perp$, where $I^\perp$ is the orthogonal complement of I with respect to κ .*

PROOF. (1) Suppose $x, y \in I$. Then $\kappa_I(x, y) = \text{tr ad}_L x \text{ad}_L y$, while $\kappa|_I(x, y) = \text{tr ad}_I x \text{ad}_I y$. Set $\phi = \text{ad}_L x \circ \text{ad}_L y$ and note that $\phi : L \rightarrow I$. It follows from Proposition 1.4.3 (Trace of a restriction, p. 36) that $\text{tr } \phi = \text{tr } \phi|_I$ in this case. In our case, $\phi|_I = \text{ad}_I x \circ \text{ad}_I y$, so this finishes the proof.

- (2) Let K be the radical of $\kappa|_I$. Then

$$x \in K \iff x \in I \text{ and } \kappa(x, y) = 0 \text{ for all } y \in I \iff x \in I \cap I^\perp.$$

■

Proposition 2.7.3 (Killing form is completely invariant).

Let L be a finite-dimensional Lie algebra. Then its Killing κ form is completely invariant, that is, for any derivation δ on L and $x, y \in L$, we have $\kappa(\delta(x), y) + \kappa(x, \delta(y)) = 0$.

REFERENCES AND REMARKS. See [Bourbaki 1-3, §I.3.6, Proposition 10, p. 35].

PROOF. We can regard $\mathbb{F}\delta$ as a 1-dimensional Lie algebra with Lie bracket zero. We can thus define a Lie algebra structure on $M = \mathbb{F}\delta \oplus L$ as in Proposition 3.1.12 (Semidirect product of Lie algebras, p. 96) in Chapter 3 (Representations, p. 83), with $[(\lambda\delta, x), (\mu\delta, y)] = (0, \delta(\lambda y - \mu x) + [x, y])$. Let κ_M denote the Killing form on M . Since κ_M is invariant, we have

$$\kappa_M(\delta(x), y) + \kappa_M(x, \delta(y)) = \kappa_M([\delta, x], y) + \kappa_M(x, [\delta, y]) = \kappa_M([\delta, x], y) + \kappa_M([x, \delta], y) = 0.$$

Since $L \oplus 0$ is an ideal of M , Proposition 2.7.2 (Killing form on an ideal, p. 72) tells us $\kappa_M|_L = \kappa$. Thus $\kappa(\delta(x), y) + \kappa(x, \delta(y)) = 0$. ■

Proposition 2.7.4 (Orthogonal complement is an ideal/subalgebra).

Let L be a finite-dimensional Lie algebra, let $\langle, \rangle$ be an invariant bilinear form on L , and let S be a subspace of L . Recall that $S^\perp = \{x \in L \mid \langle x, S \rangle = 0\}$ is always a subspace of L .

- (1) If S is a subalgebra of L , then $S^\perp$ is a subalgebra of L .
- (2) If S is an ideal of L , then $S^\perp$ is an ideal of L .

PROOF. (2) Suppose $y \in L, z \in S^\perp$ and take $x \in S$. Since S is an ideal, $[x, y] \in S$. Thus $0 = \langle [x, y], z \rangle = \langle x, [y, z] \rangle$, and so $[y, z] \in S^\perp$.

- (1) Use the proof in (2), but take $y \in S$. ■

In particular, the radical of $\langle, \rangle$, which is $L^\perp$, is always an ideal of L . If I is a derivation characteristic ideal of L and $\langle, \rangle$ is a completely invariant form, then $I^\perp$ is a derivation characteristic ideal of L . Thus the radical of a completely invariant form is a derivation characteristic ideal.

For some forms, there is a connection between the orthogonal complement of the derived subalgebra and the solvable radical of L , as the next result shows. It implies that for the concrete trace form and the Killing form, the radical of the form is contained in the solvable radical. The last part is essentially Cartan's criterion for semisimplicity.

Proposition 2.7.5 (Invariant forms and the solvable radical).

Let L be a finite-dimensional Lie algebra over a field of characteristic 0 with solvable radical R . Let $\langle, \rangle$ be a symmetric invariant bilinear form on L .

- (1) If L is a linear Lie algebra and $\langle, \rangle = \tau$ is the concrete trace form, then $[L, L]^\perp \subseteq R$.
- (2) If $\langle, \rangle = \kappa$ is the Killing form, then $[L, L]^\perp = R$.
- (3) If $\langle, \rangle = \kappa$ is the Killing form, then every Abelian ideal of L is contained in the radical $L^\perp$.

REFERENCES AND REMARKS. See [Jacobson, §III.5, Theorem 5, p. 73; Bourbaki 1-3, §I.5.5, Proposition 5, p. 48].

I don't know if equality holds in (1).

PROOF. Let $K = [L, L]^\perp$ and let $y \in [K, K]$. Then $y \in [L, L] \cap K$, and so $\langle y, y \rangle = 0$.

(1) We have $0 = \tau(y, y) = \text{tr } y^2$ for all $y \in [K, K]$. This shows K is solvable by Proposition 2.6.2 (Concrete Cartan's criterion for solvability, p. 67).

(2) We have $0 = \kappa(y, y) = \text{tr}(\text{ad } y)^2$ for all $y \in [K, K]$. This shows K is solvable by Proposition 2.6.3 (Cartan's criterion for solvability, p. 69). Thus $K \subseteq R$.

For the reverse inclusion, let $x \in R$ and $y, z \in L$. Then $\kappa(x, [y, z]) = \kappa([x, y], z)$. By Proposition 3.2.9 (Derivations take solvable to nilpotent, p. 103) (in Chapter 3 (Representations, p. 83)), $[R, L]$ is contained in the nil radical of L . Thus $\text{ad}[x, y]$ is nilpotent. **What now?** Somehow we get $\text{ad}[x, y] \circ \text{ad} z$ nilpotent, and so $\kappa([x, y], z) = \text{tr}(\text{ad}[x, y] \circ \text{ad} z) = 0$. This shows $\kappa(R, [L, L]) = 0$, which completes the proof.

(3) Suppose I is an Abelian ideal of L . Let $x \in I, y, z \in L$. Then $[x, [y, z]], [y, [x, [y, z]]] \in I$. Thus $(\text{ad} x \circ \text{ad} y)^2(z) = [x, [y, [x, [y, z]]]] = 0$ since I is Abelian. This shows that $\text{ad} x \text{ad} y$ is nilpotent, which in turn implies $\kappa(x, y) = \text{tr} \text{ad} x \text{ad} y = 0$. Thus $x \in L^\perp$. ■

7.2. Cartan's criterion for semisimplicity. For general forms, the radical of the form need not be solvable. As a trivial example, consider the zero form on a semisimple Lie algebra. The other inclusion, that the solvable radical is contained in the radical of the form, is not true in general even for the Killing form. We can see this in case 2 of the a, b, c Lie algebras, since the algebra is solvable but the Killing form is nonzero. (See Subsection 4.4 (The Killing form and trace form, p. 179) in Chapter 8 (Miscellaneous Calculations and Observations, p. 173).) However, we have the following result.

Proposition 2.7.6 (Cartan's criterion for semisimplicity).

Let L be a finite-dimensional Lie algebra over a field of characteristic 0. Then L is semisimple if and only if its Killing form is nondegenerate.

REFERENCES AND REMARKS. See [Humphreys, §5.1, Theorem, p. 22; Jacobson, §III.4, pp. 69-70; Bourbaki 1-3, §I.6.1, Theorem 1, p. 50; Erdmann-Wildon, Theorem 9.9, p. 82; Wikipedia].

Let us make a rare remark about positive characteristic. In that case, having a nondegenerate Killing form implies the Lie algebra is semisimple but the converse does not hold.

PROOF. Suppose first that L is semisimple, and so its solvable radical R is 0. By Proposition 2.7.5 (Invariant forms and the solvable radical, p. 73)(2), $L^\perp \subseteq [L, L]^\perp = 0$, so κ is nondegenerate.

Next suppose κ is nondegenerate. By Proposition 2.7.5 (Invariant forms and the solvable radical, p. 73)(3), every Abelian ideal is 0. Thus L is semisimple. ■

Here is a counterpoint to the last result. These are mostly invariant form restatements of previous results.

Proposition 2.7.7 (Zero Killing form).

Let L be a finite-dimensional Lie algebra over a field of characteristic 0.

(1) *If L is nilpotent, the Killing form on L is identically zero.*

- (2) *If the Killing form on L is identically zero, then L is solvable.*
 (3) *If L is a linear Lie algebra and the concrete trace form on L is identically 0, then L is solvable.*

REFERENCES AND REMARKS. See [Humphreys, §5, Exercise 1, p. 24; Bourbaki 1-3, §I.4.1, Corollary 2, p. 39; Erdmann-Wildon, Exercise 9.5, p. 88].

None of the converses to the statements in the proposition are true. For (2), we've seen that the Killing form of the 2-dimensional non-Abelian Lie algebra is nonzero. For (3), the concrete trace form on the (Abelian) Lie algebra $\mathfrak{d}(n)$ of diagonal matrices is actually nondegenerate.

Here are two counterexamples to the converse of (1), with details left to the reader. (See Subsection 5.2 (Examples where the Killing form is zero, p. 181) in Chapter 8 (Miscellaneous Calculations and Observations, p. 173) if you get stuck.) First, let a be a nonzero scalar and let L be the Lie algebra spanned by x, y, z with brackets $[x, y] = x$, $[y, z] = az$, $[z, x] = 0$. Then L^1 is spanned by x, z , whence $L^2 = L^1$ and $L^{(2)} = 0$, so L is solvable but not nilpotent. We have $\text{tr}(\text{ad } y)^2 = a^2 + 1$ and in fact, if $a^2 = -1$, then the Killing form on L is identically zero. Since $Z(L) = 0$, we have $L \cong \text{ad } L$ and the Killing form on L equals the concrete trace form on $\text{ad } L$. Thus the concrete trace form on $\text{ad } L$ is identically 0 and L is solvable but not nilpotent.

For an example over an arbitrary field (of characteristic 0), let b, c be nonzero scalars and let L be the Lie algebra spanned by x, y, z, w with brackets $[x, y] = x$, $[y, z] = bw$, $[y, w] = cz$, $[z, x] = [w, x] = [z, w] = 0$. Then L^1 is spanned by x, z, w , whence $L^2 = L^1$ and $L^{(2)} = 0$, so L is solvable but not nilpotent. We have $\text{tr}(\text{ad } y)^2 = 2bc + 1$ and in fact, if $2bc = -1$, then the Killing form on L is identically zero. We can always choose such b, c (e.g., $b = 1, c = -\frac{1}{2}$), so we can create a non-nilpotent Lie algebra with zero Killing form.

PROOF. (1) Proposition 2.4.4 (Engel's theorem for ad , p. 64) implies $\text{tr}(\text{ad } x)^k = 0$ for all k and all $x \in L$ if L is a nilpotent Lie algebra. In that case, $0 = \text{tr}(\text{ad } x + \text{ad } y)^2 = \text{tr}(\text{ad } x)^2 + \text{tr } \text{ad } x \text{ ad } y + \text{tr } \text{ad } y \text{ ad } x + \text{tr}(\text{ad } y)^2 = 2 \text{tr } \text{ad } x \text{ ad } y$ for all $x, y \in L$. This shows that the Killing form of a nilpotent Lie algebra is identically 0.

(2) This follows immediately from Proposition 2.6.3 (Cartan's criterion for solvability, p. 69).

(3) The proof is just like that of (2), using Proposition 2.6.2 (Concrete Cartan's criterion for solvability, p. 67). ■

The following observation is useful.

Proposition 2.7.8 (Bracket and orthogonal complement).

Let L be a finite-dimensional Lie algebra and let $\langle, \rangle$ be a symmetric invariant bilinear form on L . Let K be the radical of $\langle, \rangle$ and let I be an ideal of L .

- (1) $[I, I^\perp] \subseteq K$.
- (2) If $\langle, \rangle$ is nondegenerate, then $I \cap I^\perp$ is Abelian.
- (3) If L is semisimple and $\langle, \rangle$ is nondegenerate, then $I \oplus I^\perp = L$.

PROOF. (1) Let $x \in L, y \in I, z \in I^\perp$. Then $[x, y] \in I$, so $[x, [y, z]] = [[x, y], z] = 0$. This proves $[y, z] \in K$.

(2) This is immediate from (1).

(3) As L is semisimple, (2) implies $I \cap I^\perp = 0$, and this in turn implies $L = I \oplus I^\perp$ by Proposition 1.4.1 (Orthogonal complements, p. 35). ■

Proposition 2.7.9 (Semisimple ideal is summand).

Let L be a finite-dimensional Lie algebra over a field of characteristic 0 and let I be an ideal of L that is semisimple as a Lie algebra. Then $L = I \oplus K$ for some ideal K . In fact, if $I^\perp$ is orthogonal complement of I with respect to the Killing form on L , we may take $K = I^\perp$.

REFERENCES AND REMARKS. See [Humphreys, §5.3, proof of Theorem, p. 23; Jacobson, §I.3, Proposition 3, p. 11; Bourbaki 1-3, §I.6.1, Corollary 1, p. 51; Erdmann-Wildon, proof of Proposition 9.13, p. 85].

The proof in [Jacobson] applies to a larger class of Lie algebras, namely those that are **complete**, meaning the center is zero and all derivations are inner.

PROOF. Let κ be the Killing form on L and let $I^\perp$ be the orthogonal complement of I with respect to κ . By Proposition 2.7.2 (Killing form on an ideal, p. 72), the radical of the Killing form κ_I on I is $I \cap I^\perp$. Thus $I \cap I^\perp = 0$ by Proposition 2.7.6 (Cartan's criterion for semisimplicity, p. 74). Thus by Proposition 1.4.1 (Orthogonal complements, p. 35), we have $L = I \oplus I^\perp$. ■

A consequence of this is some results for semisimple Lie algebras that parallel those for solvable and nilpotent Lie algebras in Proposition 2.3.6 (Solvable and nilpotent for ideals and quotients, p. 57).

Proposition 2.7.10 (Semisimple for ideals and quotients).

Let L be a finite-dimensional Lie algebra over a field of characteristic 0 and let I be an ideal of L .

- (1) I and L/I are semisimple Lie algebras.
- (2) If I and L/I are semisimple Lie algebras, then L is semisimple.

(3) *A direct sum of Lie algebras is semisimple if and only if each summand is semisimple.*

REFERENCES AND REMARKS. See [Humphreys, §5.2, Corollary, p. 23; Jacobson, §III.5, Corollary 1, p. 72; Bourbaki 1-3, §I.6.1, Corollary 2, p. 51; §I.6.1, Remark 3 after Definition 1, p. 50; Erdmann-Wildon, Lemma 9.10, p. 83; Wikipedia].

Is there a more direct proof for (1), i.e., one not relying on ideals being summands?

PROOF. We first give direct proofs and then give proofs using nondegeneracy of the Killing form.

(1) By Proposition 2.7.8 (Bracket and orthogonal complement, p. 75), $L = I \oplus I^\perp$. This implies (why?) that any ideal of I is an ideal of L . Since L has no nonzero Abelian ideals, I cannot have such ideals either. Thus I is semisimple. Moreover, $L/I \cong I^\perp$ and $I^\perp$ is semisimple since it is an ideal of L .

(2) By Proposition 2.7.9 (Semisimple ideal is summand, p. 76), we have $L = I \oplus K$ for an ideal K of L . Moreover, $K \cong L/I$, so I, K are both semisimple and hence L is semisimple by (3).

(3) Suppose $L = \bigoplus_{j=1}^n L_j$. (We can actually allow L to be infinite dimensional and the sum to be infinite.) Suppose each L_j is semisimple. If I is an Abelian ideal of L , then the projection of I onto each summand L_j is Abelian and hence 0. This implies $I = 0$.

Conversely, if L is semisimple and any L_j has a nonzero Abelian ideal, that ideal is a nonzero Abelian ideal of L .

Alternative proof of (1), still using $L = I \oplus I^\perp$. The Killing form κ is the direct sum of the Killing forms κ_I and $\kappa_{I^\perp}$. Thus if κ is nondegenerate, κ_I must be nondegenerate.

Alternative proof of (3). The Killing form κ of L is the direct sum of the Killing forms κ_j of the summands, so κ is nondegenerate if and only if each κ_j is nondegenerate. ■

7.3. The structure of semisimple and reductive Lie algebras.

Proposition 2.7.11 (Semisimple is equivalent to sum of simples).

Let L be a finite-dimensional Lie algebra over a field of characteristic 0. If L is semisimple, then there are nonzero ideals $I_1, \dots, I_k$ of L , which are simple as Lie algebras and minimal as ideals of L , such that $L = I_1 \oplus \dots \oplus I_k$. Furthermore, any ideal of L is uniquely a sum of some of $I_1, \dots, I_k$.

REFERENCES AND REMARKS. See [Humphreys, §5.2, Theorem, p. 23; Jacobson, §III.5, Theorem 3, p. 71; Bourbaki 1-3, §I.6.2, Proposition 2 & Corollary 3, p. 53; Erdmann-Wildon, Theorem 9.11, p. 83; Wikipedia].

PROOF. By Proposition 2.7.6 (Cartan's criterion for semisimplicity, p. 74), the Killing form is a nondegenerate invariant form on L , and so Proposition 2.7.8 (Bracket and

orthogonal complement, p. 75) implies that $L = I \oplus I^\perp$ for any ideal I . All the claims follow from this fact. For example, pick any nonzero ideal I of minimal dimension. The direct sum decomposition $L = I \oplus I^\perp$ implies any ideal of I is actually an ideal of L , so I is simple as a Lie algebra. Since $I^\perp$ is semisimple by Proposition 2.7.10 (Semisimple for ideals and quotients, p. 76), induction implies $I^\perp$ is a direct sum of simple ideals. Thus we get the first claim.

Suppose now $L = I_1 \oplus \cdots \oplus I_k$ with each I_j minimal and let I be an ideal of L . We will show for every j , either $I_j \subseteq I$ or $[I_j, I] = I_j \cap I = 0$. Since $I \cap I_j$ is an ideal of L contained in I_j it must be 0 or I_j . The latter case is equivalent to $I_j \subseteq I$. Moreover, $[I_j, I] \subseteq I_j \cap I$, proving the final claim. Note this also shows $[I_j, I_\ell] = 0$ whenever $j \neq \ell$.

We claim that I is the sum of all I_j with $I_j \subseteq I$. To see this, suppose $i = \sum_{j=1}^k i_j$, where each $i_j \in I_j$ and choose an ℓ with $[I, I_\ell] = 0$ and $i_\ell \neq 0$. Since I_ℓ is minimal and non-Abelian, there must be an $x \in I_\ell$ with $[i_\ell, x] \neq 0$. (If not, then $i_\ell \in Z(I_\ell)$, which is an ideal. This forces I_ℓ to be Abelian.) For any $j \neq \ell$, we have $[I_j, I_\ell] = 0$, and so $0 = [i, x] = \sum_{j=1}^k [i_j, x] = [i_\ell, x] \neq 0$. This contradiction shows $i_\ell = 0$ whenever $I_\ell \not\subseteq I$, and hence I is the sum of the I_j with $I_j \subseteq I$. ■

An interesting corollary of this result is that $L = [L, L]$ for every semisimple Lie algebra. The converse is true if $\dim L \leq 4$ (Exercise!) but not in general. For example, if $L = \mathfrak{sl}(2)$ and V is the 2-dimensional irreducible representation of L , then the extension of L by V described in Proposition 3.1.11 (Extension of a Lie algebra by a module, p. 95) is a 5-dimensional perfect Lie algebra that is not semisimple, since V is a 2-dimensional Abelian ideal. (This Lie algebra is $L \oplus V$ with Lie bracket $[(x, v), (y, w)] = ([x, y], x.w - y.v)$.)

Proposition 2.7.12 (Semisimple Lie algebra equals its derived subalgebra).

Let L be a finite-dimensional Lie algebra over a field of characteristic 0. If L is semisimple, then $L = [L, L]$.

REFERENCES AND REMARKS. See [Humphreys, §5.2, Corollary, p. 23; Jacobson, §III.5, Corollary 2, p. 72; Bourbaki 1-3, §I.6.1, Theorem 1, p. 50; Wikipedia].

PROOF. We know $L = I_1 \oplus \cdots \oplus I_k$ for simple non-Abelian ideals. Thus each $[I_j, I_j]$ is a nonzero ideal of I_j and hence $[I_j, I_j] = I_j$. It now follows that $[L, L] = L$. ■

All derivations on a semisimple Lie algebra are inner.

Proposition 2.7.13 (All derivations are inner).

Let L be a finite-dimensional semisimple Lie algebra over a field of characteristic 0. Then all derivations on L are inner.

REFERENCES AND REMARKS. See [Humphreys, §5.3, Theorem, p. 23; Jacobson, §III.6, Theorem 6, p. 74; Bourbaki 1-3, §I.6.1, Corollary 3, p. 51; Erdmann-Wildon, Proposition 9.13, p. 84].

PROOF. We give two proofs.

The first is from [Jacobson]. Let κ be the Killing form on L and let δ be a derivation on L . Define $\phi \in L^*$ by $\phi(x) = \text{tr}(\delta \circ \text{ad } x)$. Since κ is nondegenerate, there exists $z \in L$ with $\kappa(z, x) = \phi(x)$ for all $x \in L$. Set $\epsilon = \delta - \text{ad } z$; this is a derivation on L . Moreover, for any $x \in L$, we have $\text{tr}(\epsilon \circ \text{ad } x) = \text{tr}(\delta \circ \text{ad } x) - \text{tr } \text{ad } z \circ \text{ad } x = \phi(x) - \kappa(z, x) = 0$.

Now take any $x, y \in L$. We have

$$\kappa(\epsilon(x), y) = \text{tr } \text{ad } \epsilon(x) \circ \text{ad } y = \text{tr}[\epsilon, \text{ad } x] \circ \text{ad } y = \text{tr } \epsilon \circ [\text{ad } x, \text{ad } y] = \text{tr } \epsilon \circ \text{ad}[x, y] = 0.$$

Thus $\epsilon(x)$ is in the radical of κ and hence is 0 for all x . This means $\delta = \text{ad } z$.

Our second proof follows our other sources. Let $D = \text{Der } L$ be the Lie algebra of derivations on L . Since $Z(L) = 0$, the map $\text{ad} : L \rightarrow D$ is injective, and so $I = \text{im } \text{ad}$ is isomorphic to L . In particular, I is semisimple. The comments after Proposition 2.2.1 (Operations on inner derivations, p. 49) show that I is an ideal of D . By Proposition 2.7.9 (Semisimple ideal is summand, p. 76), there is an ideal K of D with $D = I \oplus K$. In particular, $[K, I] \subseteq K \cap I = 0$.

Let $\delta \in K$, $x \in L$. Then $[\delta, \text{ad } x] = \text{ad } \delta(x)$; we've just seen that this is 0. Thus $\delta(x) \in Z(L) = 0$ for all $x \in L$, and so $\delta = 0$. This proves $I = D$, as desired. ■

Here is a completely different result, in the Jacobson style. If L is any subspace of $\mathfrak{gl}(V)$, then it generates an associative algebra, possibly without 1, and an associative algebra with $1 = I$. Recall that a finite-dimensional associative algebra is semisimple if and only if its Jacobson radical is 0, if and only if all one-sided ideals are direct sums of simple one-sided ideals. Semisimple associative algebras always have a multiplicative identity. Jacobson uses this connection to bring the power of associative algebra theory to bear on the study of Lie algebras. Bourbaki occasionally does so as well.

Proposition 2.7.14 (When is a linear Lie algebra reductive?).

Let V be a finite-dimensional vector space over a field of characteristic 0 and let $L \subseteq \mathfrak{gl}(V)$ be a linear Lie algebra, Let A be the associative subalgebra of $\text{End}(V)$ generated by L . Then A is semisimple if and only if L is reductive. In this case, $L = Z(L) \oplus I$ for a semisimple ideal I of L .

REFERENCES AND REMARKS. See [Jacobson, §II.5, Theorem 8, p. 44; §II.5, Corollary, p. 47; Bourbaki 1-3, §I.6.4 & 5, pp. 56-58].

PROOF. Omitted. ■

8. Extension of Scalars

In this section, we state results related to extension of scalars. The most desirable are of the form: if $\mathbb{E}$ is a field extension of $\mathbb{F}$, then L has property $\mathcal{P}$ if and only if $L_{\mathbb{E}}$ has property $\mathcal{P}$.

Proposition 2.8.1 (Isomorphism and extension of scalars).

Let L be a Lie algebra over the field $\mathbb{F}$ and let $\mathbb{E}$ be a field extension of $\mathbb{F}$. Let V, W be finite-dimensional L -modules.

- (1) *V, W are isomorphic as L -modules if and only if $V_{\mathbb{E}}, W_{\mathbb{E}}$ are isomorphic as $L_{\mathbb{E}}$ -modules.*
- (2) *If $V_{\mathbb{E}}$ is completely reducible as an $L_{\mathbb{E}}$ -module, then V is completely reducible as an L -module. If $\mathbb{F}$ has characteristic 0, then the converse is true.*

REFERENCES AND REMARKS. See [Jacobson, §I.8, pp. 26-28; §III.7, Lemma 4, p. 80; Bourbaki 1-3, §I.3.8, Proposition 13, p. 37].

The corresponding result for Lie algebras is not true, that is, if L, M are Lie algebras over $\mathbb{F}$ and $L_{\mathbb{E}} \cong M_{\mathbb{E}}$ as Lie algebras over $\mathbb{E}$, we cannot conclude that $L \cong M$ as Lie algebras over $\mathbb{F}$. For example, if $\mathbb{F} = \mathbb{R}$ and $\mathbb{E} = \mathbb{C}$, we have $\mathfrak{sl}(2, \mathbb{R}) \not\cong \mathfrak{o}(3, \mathbb{R})$ but $\mathfrak{sl}(2, \mathbb{C}) \cong \mathfrak{o}(3, \mathbb{C})$.

PROOF. **Maybe later.** ■

Proposition 2.8.2 (Derived/central series and extension of scalars).

Let L be a Lie algebra over the field $\mathbb{F}$ and let $\mathbb{E}$ be a field extension of $\mathbb{F}$.

- (1) *$(I_{\mathbb{E}})^k = (I^k)_{\mathbb{E}}$ and $(I_{\mathbb{E}})^{(k)} = (I^{(k)})_{\mathbb{E}}$ for all k .*
- (2) *$(Z_{\mathbb{E}})_k = (Z_k)_{\mathbb{E}}$ for all k .*
- (3) *Suppose L is finite-dimensional. The Killing form $\kappa_{\mathbb{E}}$ on $L_{\mathbb{E}}$ is the extension $(\kappa)_{\mathbb{E}}$ of the Killing form on L , and they have the same matrix. Thus if K is the radical of κ , then $K_{\mathbb{E}}$ is the radical of $\kappa_{\mathbb{E}}$.*

REFERENCES AND REMARKS. See [Jacobson, §III.4, Corollary, p. 70; Bourbaki 1-3, §I.4.5, p. 42; §I.5.6, p. 49; §I.5.6, p. 50; §I.6.10, p. 68].

PROOF. **Maybe later.** ■

Proposition 2.8.3 (Types of Lie algebras and extension of scalars).

Let L be a finite-dimensional Lie algebra over the field $\mathbb{F}$ of characteristic 0 and let $\mathbb{E}$ be a field extension of $\mathbb{F}$. Then L is any of the following if and only if $L_{\mathbb{E}}$ is: Abelian, nilpotent, solvable, semisimple, reductive.

REFERENCES AND REMARKS. See [**Jacobson**, §III.4, Corollary, p. 70; **Bourbaki 1-3**, §I.4.5, p. 42; §I.5.6, pp. 49-50; §I.5.6, p. 50].

Do we need the finite-dimensional or the characteristic 0 hypotheses? They're used in the proof when the Killing form is invoked.

It is possible for L to be simple but $L_{\mathbb{E}}$ to fail to be simple. **Example.** However, if $L_{\mathbb{E}}$ is simple, L must be simple.

[**Bourbaki 1-3**, §I.5.6, p. 50] also says that if M is a finite-dimensional semisimple Lie algebra over $\mathbb{E}$ and $\mathbb{E}$ is a finite extension of $\mathbb{F}$, then M is semisimple when regarded as a Lie algebra over $\mathbb{F}$. (This is called restriction.)

PROOF. **Maybe later.** ■

CHAPTER 3

Representations

In this chapter, we study representations of a Lie algebra. Representations occur throughout mathematics and are one of the main motivations for studying Lie algebras.

A (finite-dimensional) representation of a Lie algebra L amounts to assigning matrices of a given size or operators on a particular vector space in a way that respects the commutator bracket. In Section 1 (Definitions and Basics, p. 84), we make the basic definitions and introduce the key concept of an irreducible representation, which is a representation that can't be made any smaller. We give complete descriptions of 1-dimensional representations and representations of Abelian Lie algebras. We then describe some ways to construct new modules from given ones. We also define the notion of semidirect product, which enables us to create new Lie algebras.

In Section 2 (Lie's Theorem, p. 97), we give several versions of Lie's theorem on solvable Lie algebras. We prove all irreducible representations of a solvable Lie algebra over an algebraically closed field are 1-dimensional. This leads to the following characterization of solvable (over an algebraically closed field): a linear Lie algebra is solvable if and only if it is triangularizable. A general Lie algebra (over an algebraically closed field), is solvable if and only if it has a full flag of ideals (a chain with each ideal of codimension 1 in the one above it). When the field is not algebraically closed, these results apply to split Lie algebras. Finally, we show that if L is solvable, then the derived subalgebra $[L, L]$ is nilpotent. For a general L with solvable radical R , the image $\delta(R)$ is contained in the nil radical for every derivation δ on L .

In Section 3 (Representations of $\mathfrak{sl}(2)$, p. 103), we give a detailed description of all (finite-dimensional) irreducible representations of $\mathfrak{sl}(2)$. In Section 5 (Weyl's Theorem on Complete Reducibility, p. 114), we will show all representations of $\mathfrak{sl}(2)$ are completely reducible, so this gives a complete description of representations of $\mathfrak{sl}(2)$. We show that for each positive integer n , there is exactly one irreducible representation of $\mathfrak{sl}(2)$ of dimension n . Recall that $\mathfrak{sl}(2)$ is usually described by a basis e, h, f with $[e, f] = h$, $[h, e] = 2e$, $[h, f] = -2f$. The action of h on each irreducible module is diagonal, with integer eigenvalues, and the largest eigenvalue that occurs is one less than the dimension of the representation.

In Section 4 (Representations of $\mathfrak{sl}(3)$, p. 107), we will describe the representations of $\mathfrak{sl}(3)$. Here the eigenvalues and eigenvectors of h are replaced by the simultaneous eigenvalues and eigenvectors of several h 's; we call these eigenvalues and eigenvectors weights and weight vectors. The relation of the weights to each other puts us in a geometric setting. We also lose some of the uniformity of the $\mathfrak{sl}(2)$ case: irreducible representations do not occur in every dimension and some dimensions have multiple non-isomorphic irreducible representations. The results of this section give us a hint as to the nature of representation theory over general semisimple Lie algebras.

In Section 5 (Weyl's Theorem on Complete Reducibility, p. 114), we prove that all (finite-dimensional) representations of a semisimple Lie algebra are completely reducible. Hermann Weyl proved this theorem (with $\mathbb{F} = \mathbb{C}$) in 1925 using analysis via an inner product and the "unitarian trick" for the corresponding Lie group in a way that is similar to (but more complicated than) the proof of the same fact for finite groups. The proof we give starts with the Casimir element, a special element/operator that commutes with all elements of L . We then use the Casimir element to show a certain exact sequence splits, which gives us Weyl's theorem.

Weyl's theorem has many applications. For example, we've seen before that every operator has a Jordan-Chevalley decomposition and that the semisimple and nilpotent components of a derivation are derivations. In particular, $\text{ad } x$ has such a decomposition into derivations. But are the pieces inner derivations by elements of L , i.e., can we write $\text{ad } x = \text{ad } y + \text{ad } z$ with $\text{ad } y$ semisimple and $\text{ad } z$ nilpotent? We prove that for a semisimple Lie algebra L , there is such a decomposition with $y, z \in L$. This is sometimes called the abstract Jordan-Chevalley decomposition. If L is a semisimple linear Lie algebra, then x has its own Jordan-Chevalley decomposition $x = y' + z'$. We show that $y = y'$ and $z = z'$, so there is no ambiguity in the term "Jordan-Chevalley decomposition".

Finally, we show that a reductive Lie algebra is a Lie algebra that is the direct sum of an Abelian Lie algebra and a semisimple Lie algebra. We also describe for a linear Lie algebra $L \subseteq \mathfrak{gl}(V)$ when V is an irreducible or completely reducible L -module. In either case, L must be reductive and the elements of $Z(L)$ must be semisimple. In the irreducible case, the center of L must be small.

1. Definitions and Basics

1.1. Basics of modules and representations. A **module** over a Lie algebra L is a vector space V equipped with an **action** of L on V , that is, a map $L \times V \rightarrow V$ satisfying certain conditions. (As usual, we assume V is finite-dimensional unless otherwise stated.) We will denote the image of (x, v) for $x \in L, v \in V$ by $x.v$ or simply xv . The conditions that the action must satisfy are that it is bilinear (i.e., $(\lambda x).v = x.(\lambda v) = \lambda(x.v)$)

for a scalar λ and $(x_1 + x_2).v = x_1.v + x_2.v$ and $x.(v_1 + v_2) = x.v_1 + x.v_2$ and that $[x, y].v = x.(y.v) - y.(x.v)$. The second condition may appear a bit mysterious, but we will explain it shortly. An L -module is also called a **representation of L** , although technically, “representation” has a different meaning. A **representation** of L is a Lie algebra homomorphism $\phi : L \rightarrow \mathfrak{gl}(V)$ for a vector space V , or equivalently, a Lie algebra map $\phi : L \rightarrow \mathfrak{gl}(n)$ for some n . If we define $x.v = \phi(x)(v)$, then a representation defines a module action. The bilinearity is easy to see (since ϕ is linear) and $[x, y].v = \phi([x, y])(v) = \phi(x)(\phi(y)(v)) - \phi(y)(\phi(x)(v)) = x.(y.v) - y.(x.v)$, as required of a module action. This calculation should explain the definition of a module action. Of course, we can do this in reverse: given an action, define $\phi : L \rightarrow \mathfrak{gl}(V)$ by $\phi(x)(v) = x.v$. It’s easy to see that ϕ is a map of Lie algebras. We will use the words “module” and “representation” interchangeably.

If V is an L -module and S is a subalgebra of L , then V is easily seen to be an S -module, and the representation is given by $\phi|_S$. If S is a subspace of L , we use $S.V$ or just SV to denote the subspace of V spanned by all elements of the form $s.v$ for $s \in S, v \in V$. (Not all elements of $S.V$ will necessarily have the form $s.v$.) We leave the reader to verify that when S is a subalgebra of V , the space SV is an S -submodule of V and when S is an ideal of V , then SV is an L -submodule of V . It is not always the case that $LV = V$. (This always holds for modules over rings with identity).

A primary example of a module/representation occurs when L is a linear Lie algebra, say L is a subalgebra of $\mathfrak{gl}(V)$. We then define $A.v = Av$ and the representation map is the inclusion map: this makes V an L -module. Another example is the **regular** or **adjoint** representation $\text{ad} : L \rightarrow \mathfrak{gl}(L)$, which makes L a module over itself. For more examples, see Subsection 1.2 (Two examples, p. 87).

The **kernel** of a representation $\phi : L \rightarrow \mathfrak{gl}(V)$ is the kernel of the linear map ϕ . If V is an L -module, the kernel of the corresponding representation is $\{x \in L \mid x.v = 0 \text{ for all } v \in V\}$. The kernel K is an ideal of L , since if $x \in L, y \in K$, then $[x, y].v = x.y.v - y.x.v = 0 - 0 = 0$, so $[x, y] \in K$. If the kernel is 0, we say the representation or the module is **faithful**. It is always the case that V is faithful as a module over the quotient Lie algebra L/K . If the kernel of a representation is L , we say the representation or module is **trivial**. There is one trivial representation/module of every dimension. The trivial representation of dimension n is isomorphic to the direct sum of n copies of the 1-dimensional trivial representation. **Warning:** Some sources use the term trivial only for the 1-dimensional trivial representation. Some distinguish between *the* trivial representation, which is the 1-dimensional trivial representation, and *a* trivial representation, which can have any dimension.

The following definition is **not standard**. We say an L -module V is **split** if the map $v \mapsto x.v$ from V to V is split (triangularizable) for every $x \in L$. That is, if $\phi : L \rightarrow \mathfrak{gl}(V)$ is the corresponding representation, we ask that $\phi(x)$ is split for all $x \in V$. If $\mathbb{F}$ is algebraically closed, every module is split. If V is a split L -module, so is W/U for any submodules $U \subseteq W$ of V . In addition, V is a split K/J -module for any subalgebra K of L and ideal J of K with $K.V = 0$. Sometimes there are several Lie algebras at play. For example, when L is semisimple and H is a toral/Cartan subalgebra, we may be more interested in the action of H than that of all of L . In this case, we refer to modules as L -split or H -split.

For some examples, see Subsection 1.2 (Two examples, p. 87).

Just as for modules over rings or groups, the modules over a given Lie algebra form a category with homomorphisms, isomorphisms, submodules, and quotient modules. These are defined exactly as they are for vector spaces, with the extra requirement that the module action be respected. We recall some alternative terminology. Suppose V, W are modules over L corresponding to representations $\phi_V : L \rightarrow \mathfrak{gl}(V)$, $\phi_W : L \rightarrow \mathfrak{gl}(W)$. A homomorphism $\theta : V \rightarrow W$ is a linear map satisfying $\theta(x.v) = x\theta(v)$ for $x \in L, v \in V$. This can be rephrased as $\theta \circ \phi_V(x) = \phi_W(x) \circ \theta$ for all $x \in L$. Because of this, we also call θ an **intertwining map** or **equivariant map**.

We also have external and internal direct sums. For example, we say V is the internal direct sum of submodules $V_1, \dots, V_k$ if $V = \bigoplus_{i=1}^k V_i$ as a vector space. In this case, we say the V_i are **direct summands** (or just **summands** of V). More generally, a submodule W is a **direct summand** or **summand** of V if $V = W \oplus W'$ for some submodule W' . We say a module V is **indecomposable** if it cannot be written $V = V_1 \oplus V_2$ with V_1, V_2 nonzero submodules of V .

A module is called **simple** or an **irreducible** representation if it is not zero and has no nonzero proper submodules. A module or representation is called **completely reducible**, or sometimes **semisimple**, if it is a direct sum of simple modules.

Vista. We will see later that given any Lie algebra L over $\mathbb{F}$, there is an associative algebra $U = U(L)$ (infinite-dimensional) such that L is a subspace of U and the commutator in U when restricted to L is the Lie bracket. Moreover, L generates U as an algebra and elements of U look (as a vector space, not a ring) like polynomials in elements of L . This associative algebra is called the **universal enveloping algebra** of U and it has one more property. Modules over U (as a ring) are exactly the same as modules over L (as a Lie algebra). Thus if we are willing to grant the existence of this universal enveloping algebra, we do not need to introduce a separate module theory for Lie algebras. We will hold off on the use of the universal enveloping algebra for now.

1.2. Two examples. We now detail representation theory in the case of 1-dimensional representations and in the case of Abelian Lie algebras. Our first example shows that 1-dimensional representations are in bijective correspondence with the dual space $(L/[L, L])^*$.

Proposition 3.1.1 (One-dimensional representations).

Let L be a Lie algebra and let V be a 1-dimensional L -module.

- (1) *V is a split irreducible module.*
- (2) *There is an element $\lambda \in L^*$ with $x.v = \lambda(x)v$ for all $x \in L, v \in V$. Moreover, λ determines V up to isomorphism.*
- (3) *The λ in (1) vanishes on $[L, L]$. In fact, $\lambda \in L^*$ occurs for some 1-dimensional module V if and only if $\lambda([L, L]) = 0$.*
- (4) *If $L = [L, L]$ (for example, L is semisimple), then V is trivial.*

PROOF. (1) This is obvious.

(2) If v_0 is a basis for V , then $x.v_0 = \lambda v_0$ for some scalar λ . If we define $\lambda(x) = \lambda$, it's easy to see $x.v = \lambda(x)v$ for all $v \in V$ and it's easy to see $\lambda(x)$ defined in this way is linear.

Suppose W is another 1-dimensional module with $x.w = \mu(x)w$ for all $w \in W$ and suppose $\phi : V \rightarrow W$ is an L -module isomorphism. Then $\lambda(x)\phi(v) = \phi(x.v) = x.\phi(v) = \mu(x)\phi(v)$ for all $v \in V$. Since some $\phi(v) \neq 0$, this implies $\lambda(x) = \mu(x)$ for all $x \in L$.

(3) Let $x, y \in L$ and $v \in V$ be nonzero. Then $[x, y].v = x.y.v - y.x.v = \lambda(y)\lambda(x)v - \lambda(x)\lambda(y)v = 0$. This shows $\lambda([x, y]) = 0$.

Suppose $\lambda \in L^*$ satisfies $\lambda([L, L]) = 0$ and let V be a 1-dimensional vector space with basis v_0 . Then we claim the action $x.v = \lambda(x)v$ makes V into an L -module. To see this, we need to show $[x, y].v = x.(y.v) - y.(x.v) = x.(\lambda(y)v) - y.(\lambda(x)v) = \lambda(y)\lambda(x)v - \lambda(x)\lambda(y)v = 0$. Since $\lambda([x, y]) = 0$, this equality holds.

(4) This follows immediately from (3). ■

The next result summarizes representation theory for Abelian Lie algebras.

Proposition 3.1.2 (Representations of Abelian Lie algebras).

Let L be an Abelian Lie algebra with basis $x_1, \dots, x_n$ over a field $\mathbb{F}$ of characteristic 0 and let V be a finite-dimensional vector space over $\mathbb{F}$.

- (1) *If $A_1, \dots, A_n \in \mathfrak{gl}(V)$ are commuting operators, then there is a unique representation of L with $x_i.v = A_i.v$ for all i and all $v \in V$. Moreover, every representation of L on V arises in this way.*
- (2) *The representation in (1) is split if and only if the operators $A_1, \dots, A_n$ are split. In this case, the representation of L on V is triangularizable. (That is, there is a basis of V with respect to which all of the operators $v \mapsto x.v$ for $x \in L$ are triangular.)*
- (3) *If the representation in (1) is split, then it is irreducible if and only if $\dim V = 1$.*

- (4) *The representation in (1) is completely reducible if and only if the operators $A_1, \dots, A_n$ are semisimple.*
- (5) *Every finite-dimensional irreducible representation of L is 1-dimensional if and only if $\mathbb{F}$ is algebraically closed.*

REFERENCES AND REMARKS. **Any references?**

For a concrete example, let $L = \mathbb{F}x$ be a 1-dimensional Lie algebra and let $V = \mathbb{F}^2$. The representation of L on V defined by the matrix $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ is split but not completely reducible. If $\mathbb{F} = \mathbb{R}$, the representation of L on V defined by the matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ is not split and is irreducible of dimension greater than 1.

We may be able to eliminate the hypothesis that V is finite-dimensional in parts of the proposition.

PROOF. (1) This is easy to see.

(2) This is also easy to see. The triangularization result follows from Proposition 1.3.5 (Frobenius's theorem on triangularizability, p. 29).

(3) This follows from (2).

(4) Assume $\mathbb{F}$ is algebraically closed and $A_1, \dots, A_n$ are semisimple. Then $A_1, \dots, A_n$ are simultaneously diagonalizable, say with respect to a basis $v_1, \dots, v_r$ of V . Then each $\mathbb{F}v_i$ is a 1-dimensional subspace of V invariant under each A_i . This means that each $\mathbb{F}v_i$ is an irreducible L -module. Since $V = \bigoplus_{i=1}^r \mathbb{F}v_i$, this shows V is completely reducible.

Conversely, suppose V is completely reducible. Then by (3), every irreducible module is 1-dimensional and so has a simultaneous eigenvector for $A_1, \dots, A_n$. Since V is the sum of such modules, we see that V has a basis that makes each of $A_1, \dots, A_n$ diagonal. Thus these operators are all semisimple.

If $\mathbb{F}$ is not algebraically closed, we can use the fact that V is completely reducible if and only if $V_{\bar{\mathbb{F}}}$ is completely reducible and apply the algebraically closed case. As an alternative, one can develop a theory of decomposition into subspaces corresponding to the (irreducible) factors of the minimal polynomial, paralleling the theory of diagonalization.

(5) If $\mathbb{F}$ is algebraically closed, then all finite-dimensional modules are split and we can apply (3).

Suppose $\mathbb{F}$ is not algebraically closed. Then there is a finite extension $\mathbb{E}$ of $\mathbb{F}$ that contains a root of an irreducible $p \in \mathbb{F}[x]$ of degree greater than 1. There exists a vector space V over $\mathbb{F}$ and an operator A with both minimal and characteristic polynomial p . (For example, we can take $V = \mathbb{E} = \mathbb{F}[\alpha]$ for some root α of p in $\mathbb{E}$ and take A to be multiplication by α .) If W is a nonzero proper submodule of V , then $A|_W$ has

characteristic polynomial properly dividing p . Since p is irreducible, this is impossible. Thus V is an irreducible L -module of dimension greater than 1. ■

1.3. The Jordan-Hölder and Krull-Schmidt theorems. There are two ways to reduce modules into simpler ones: direct sums and composition series. The most desirable reduction is as a direct sum of irreducible modules. This is not always possible; we state below two results about decompositions. These results hold in general settings, such as group representations and modules over rings, when the representations are small enough.

The first result gives us a “unique factorization” into indecomposable modules.

Proposition 3.1.3 (Krull-Schmidt Theorem).

Let V be a finite-dimensional module over a Lie algebra L . Then there exist indecomposable submodules $V_1, \dots, V_k$ of V such that $V = \bigoplus_{i=1}^k V_i$. Moreover, if $V \cong \bigoplus_{i=1}^\ell W_i$ for some indecomposable L -modules $W_1, \dots, W_\ell$, then $k = \ell$ and there is a permutation σ of $\{1, \dots, k\}$ such that $W_i \cong V_{\sigma(i)}$ for each $i = 1, \dots, k$.

REFERENCES AND REMARKS. See a modern algebra textbook or **Wikipedia**.

We are most interested in this result for completely reducible modules V , in which case the uniqueness proof can be simplified somewhat.

PROOF. This is a standard, but not trivial, result. We invite the reader to prove it or look it up. ■

Not every module is completely reducible if our Lie algebra is not semisimple. We can still analyze modules in terms of irreducible modules. A **composition series** of a module V is a sequence $0 = V_0 \subseteq V_1 \subseteq \dots \subseteq V_k = V$ such that V_i/V_{i-1} is irreducible for each i . [Bourbaki 1-3] calls this a **Jordan-Hölder series** and uses a different indexing convention ($V_0 = V$). The factors V_i/V_{i-1} are called the **composition factors** of the series and k is called the **length** of the series. The next result shows they are the same (perhaps in a different order) for any composition series.

Proposition 3.1.4 (Jordan-Hölder Theorem).

The following statements hold for any finite-dimensional module V over a Lie algebra L .

- (1) V has a composition series.
- (2) Any two composition series of V have the same length and the same composition factors, up to order and isomorphism.
- (3) Any chain of submodules of V can be refined to a composition series by adding submodules.

REFERENCES AND REMARKS. See a modern algebra textbook or **Wikipedia**.

PROOF. This is a standard, but not trivial, result. We prove (1) and invite the reader to prove the rest or look it up.

(1) Let V_1 be a nonzero submodule of V of least dimension. Clearly V_1 is irreducible. Next, let V_2/V_1 be a nonzero submodule of V/V_1 of least dimension, where $V_1 \subseteq V_2$. Then V_2/V_1 is irreducible. If we continue in this fashion (or use induction), we obtain a chain $0 = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_k = V$ such that V_i/V_{i-1} is irreducible for $i = 1, \dots, k$. This is a composition series for V . ■

1.4. Completely reducible representations. Again, as with modules over rings, there is a theory of semisimple objects that says quotients, submodules, and direct sums of completely reducible modules are completely reducible and that the direct sum decomposition of a completely reducible module is unique up to order and isomorphism (an easier version of the Krull-Schmidt theorem).

Here are two more related results that hold in a general setting (i.e., for modules over rings). The following result relates complete reducibility and direct summands.

Proposition 3.1.5 (Complete reducibility is equivalent to all submodules being summands).

Let V be a finite-dimensional module over the Lie algebra L . The following statements are equivalent.

- (1) V is completely reducible.
- (2) The sum of all irreducible submodules of V is V .
- (3) The intersection of all maximal submodules of V is 0.
- (4) Every submodule of V is a direct summand.
- (5) Every irreducible submodule of V is a direct summand.
- (6) Every maximal submodule of V is a direct summand.

REFERENCES AND REMARKS. Conditions (1), (2), (4) are equivalent for arbitrary modules, but (3), (5), (6) require some finiteness condition on V .

PROOF. We may assume $V \neq 0$.

(1) $\Rightarrow$ (2). This is clear.

(2) $\Rightarrow$ (4). Let W be a submodule of V and let W' be a submodule of V maximal with respect to the property $W \cap W' = 0$. We will show $V = W \oplus W'$. Let S be an irreducible submodule of V . If $S \not\subseteq W \oplus W'$, then $S \cap (W \oplus W') = 0$ since S is irreducible. This means $W \cap (W' + S) = 0$ and so (by maximality), we must have $W' + S = W'$, i.e., $S \subseteq W'$. This contradicts $S \not\subseteq W \oplus W'$, so all irreducible submodules of V are contained in $W \oplus W'$. By hypothesis (2), this implies $W \oplus W' = V$.

(4) $\Rightarrow$ (5)&(6). These are obvious.

(5) $\Rightarrow$ (1)&(3) and (6) $\Rightarrow$ (1)&(3). In either case, we will prove that for each positive integer k up to some r , there are irreducible submodules $S_1, \dots, S_k$, and maximal submodules $M_1, \dots, M_k$ with $V = S_1 \oplus \dots \oplus S_k \oplus (M_1 \cap \dots \cap M_k)$, and when $k = r$, we have $M_1 \cap \dots \cap M_r = 0$.

To get started with $k = 1$, in case (5), we note V has an irreducible submodule S_1 (e.g., any nonzero submodule of minimal dimension), and by hypothesis there is a submodule M_1 with $S_1 \oplus M_1 = V$. Since $S_1 \cong V/M_1$, we see that M_1 is maximal. In case (6), we note that V has a maximal submodule M_1 (e.g., any proper submodule of maximal dimension), and by hypothesis, there is a submodule S_1 with $S_1 \oplus M_1 = V$. Since $S_1 \cong V/M_1$, we see that S_1 is irreducible. If $M_1 = 0$, then we take $r = 1$ and we are done.

Suppose now that $V = S_1 \oplus \dots \oplus S_k \oplus (M_1 \cap \dots \cap M_k)$ and $M_1 \cap \dots \cap M_k \neq 0$ for $S_1, \dots, S_k$ irreducible and $M_1, \dots, M_k$ maximal submodules.

In case (5), we know that since $M_1 \cap \dots \cap M_k \neq 0$, it has an irreducible submodule S_{k+1} . By (5), there is a maximal submodule M_{k+1} with $S_{k+1} \oplus M_{k+1} = V$. We claim $V = S_1 \oplus \dots \oplus S_k \oplus S_{k+1} \oplus (M_1 \cap \dots \cap M_k \cap M_{k+1})$. To see that the sum is direct, suppose $\sum_{i=1}^{k+1} s_i = w \in M_1 \cap \dots \cap M_k \cap M_{k+1}$, where $s_i \in S_i$. Then $\sum_{i=1}^k s_i = w - s_{k+1} \in M_1 \cap \dots \cap M_k$, whence $s_i = 0$ for $i \leq k$ and $w = s_{k+1} \in M_{k+1} \cap S_{k+1} = 0$. This proves directness. Now let $v \in V$. Then $v = s_1 + \dots + s_k + w$ for some $s_i \in S_i$, $w \in M_1 \cap \dots \cap M_k$. Moreover, $w = s_{k+1} + w'$ for $s_{k+1} \in S_{k+1}$ and $w' \in M_{k+1}$. Since $w, s_{k+1} \in M_1 \cap \dots \cap M_k$, it follows that $w' \in M_1 \cap \dots \cap M_k \cap M_{k+1}$. As $v = s_1 + \dots + s_k + s_{k+1} + w'$, this shows $V = S_1 \oplus \dots \oplus S_k \oplus S_{k+1} \oplus (M_1 \cap \dots \cap M_k \cap M_{k+1})$.

In case (6), we know that since $S_1 \oplus \dots \oplus S_k$ is a proper submodule of V , it is contained in a maximal submodule M_{k+1} . By (6) there is an irreducible submodule S_{k+1} with $V = S_{k+1} \oplus M_{k+1}$. We claim $V = S_1 \oplus \dots \oplus S_k \oplus S_{k+1} \oplus (M_1 \cap \dots \cap M_k \cap M_{k+1})$. To see that the sum is direct, suppose $\sum_{i=1}^{k+1} s_i = w \in M_1 \cap \dots \cap M_k \cap M_{k+1}$, where $s_i \in S_i$. Then $s_{k+1} = \sum_{i=1}^k s_i - w \in M_{k+1}$, whence $s_{k+1} = 0$. Thus $\sum_{i=1}^k s_i = w \in M_1 \cap \dots \cap M_k$, and this implies each $s_i = 0$ and $w = 0$. This proves directness. Now let $v \in V$. We can write $v = s_{k+1} + w$ where $s_{k+1} \in S_{k+1}$, $w \in M_{k+1}$. We can then write $w = \sum_{i=1}^k s_i + w'$ where each $s_i \in S_i$ and $w' \in M_1 \cap \dots \cap M_k$. Since each $s_i \in M_{k+1}$, this equality implies $w' \in M_1 \cap \dots \cap M_k \cap M_{k+1}$. As $v = s_1 + \dots + s_k + s_{k+1} + w'$, this proves $V = S_1 \oplus \dots \oplus S_k \oplus S_{k+1} \oplus (M_1 \cap \dots \cap M_k \cap M_{k+1})$.

In either case, if $M_1 \cap \dots \cap M_k \cap M_{k+1} = 0$, we quit with $r = k + 1$. Otherwise we continue to $k + 2$. We cannot continue forever, since $\dim S_1 \oplus \dots \oplus S_k \geq k$ and V is finite-dimensional.

(3) $\Rightarrow$ (1). Since V is finite-dimensional, there must be a finite set $M_1, \dots, M_k$ of maximal ideals with $M_1 \cap \dots \cap M_k = 0$. There is a natural map from M to $\bigoplus_{i=1}^k M/M_i$ given by $m \mapsto (m + M_1, \dots, m + M_k)$ with kernel $M_1 \cap \dots \cap M_k$. Thus this map embeds

M into $\bigoplus_{i=1}^k M/M_i$. Each M/M_i is irreducible, so M is isomorphic to a submodule of a completely reducible module. This implies M is completely reducible. ■

Proposition 3.1.6 (Direct summands are equivalent to projections).

Let V be a finite-dimensional module over the Lie algebra L and let W be a submodule. The following statements are equivalent.

- (1) W is a direct summand of V .
- (2) There is an L -module homomorphism $\pi : V \rightarrow W$ with $\pi|_W = \text{id}_W$.
- (3) There is an L -module homomorphism $\pi : V \rightarrow V$ with $\pi^2 = \pi$ and $\text{im } \pi = W$.

PROOF. Exercise. ■

One of our key results later is that a Lie algebra is semisimple if and only if all of its (finite-dimensional) modules are completely reducible. **reference**

Here is an example of a standard theorem from module theory.

Proposition 3.1.7 (Schur's Lemma).

Let V, W be irreducible modules over the Lie algebra L and let $\phi : V \rightarrow W$ be an L -module homomorphism.

- (1) Either $\phi = 0$ or ϕ is an isomorphism.
- (2) If $V = W$ and $\mathbb{F}$ is algebraically closed, then there is a scalar $\lambda \in \mathbb{F}$ with $\phi(v) = \lambda v$ for all $v \in V$.

REFERENCES AND REMARKS. See [Humphreys, §6.1, p. 26; Erdmann-Wildon, Lemma 7.13, p. 62; Wikipedia].

Part (2) can fail when $\mathbb{F}$ is not algebraically closed. For example, let $\mathbb{F} = \mathbb{R}$, let L be a 1-dimensional Abelian Lie algebra with basis x , and let V be the 2-dimensional vector space with basis v, w . We let x act by the matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, that is, $x.v = w$, $x.w = -v$. Since this matrix has no eigenvectors over $\mathbb{R}$, the L -module V is irreducible. Now define $\phi : V \rightarrow V$ by $\phi(u) = xu$. Since L is Abelian, ϕ is an L -module homomorphism, and obviously it is an isomorphism. However, ϕ is not multiplication by a scalar.

PROOF. (1) Suppose $\phi \neq 0$. Then $\ker \phi$ is a proper submodule of V and hence must be 0. Also, $\text{im } \phi$ is a nonzero submodule of W and hence must be W . Thus ϕ is injective and surjective.

(2) Assume V is finite-dimensional; then ϕ has an eigenvalue λ . Thus the λ -eigenspace U of V is nonzero. We claim U is a submodule; if that is true, then $U = V$ and we are done. To see this, let $u \in U, x \in L$. Then $\phi(x.u) = x.\phi(u) = x.(\lambda u) = \lambda x.u$, so $x.u \in U$.

The fact that (2) holds even if V is infinite dimensional (for $\mathbb{F}$ algebraically closed) is known as Quillen's Lemma. ■

1.5. New modules from old. We now introduce some ways of combining modules. When $x \in L$ and ϕ is a function with domain V , we will use $(x.\phi)(v)$ to denote $x.\phi$ (as defined below) applied to v . However, $x.\phi(v) = x.(\phi(v))$.

Proposition 3.1.8 (Combining modules).

Let V, W be modules over the Lie algebra L . Then the following are modules when given the listed actions with $x \in L$.

- (1) V^* with $(x.\phi)(v) = -\phi(x.v)$ for $\phi \in V^*, v \in V$.
- (2) $V \oplus W$ with $x.(v, w) = (x.v, x.w)$ for $v \in V, w \in W$.
- (3) $V \otimes W$ with $x.(v \otimes w) = x.v \otimes w + v \otimes x.w$ for $v \in V, w \in W$.
- (4) $\text{Hom}_{\mathbb{F}}(V, W)$ with $(x.\phi)(v) = x.\phi(v) - \phi(x.v)$ for $\phi : V \rightarrow W, v \in V$.

REFERENCES AND REMARKS. See [Humphreys, §6.1, pp. 26-27; Jacobson, §I.6, pp. 19-23; Bourbaki 1-3, §I.3, pp. 25-31 and §III.Appendix, pp. 368-369; Erdmann-Wildon, Exercise 7.12, p. 65].

The dual module V^* is also called the **contragredient module** of V .

There is also a vector space $\text{Hom}_L(V, W)$ of L -module homomorphisms from V to W , but this is not an L -module in a natural way. **Isn't it?? Often trivial.**

PROOF. (1) follows from (4) when we take $W = \mathbb{F}$ with trivial action $x.w = 0$.

$$(2) \quad x.y.(v, w) - y.x.(v, w) = (x.y.v - y.x.v, x.y.w - y.x.w) = ([x, y].v, [x, y].w) = [x, y].(v, w).$$

(3) To give a complete proof, we should note that for fixed x , the action $(x, (v, w)) \mapsto x.v \otimes w + v \otimes x.w$ is bilinear and induces a linear map from $V \otimes W \rightarrow V \otimes W$ and then show that action is linear in x . We leave this to the reader. We will verify the module rule on simple tensors (why is this enough?).

$$\begin{aligned} x.y.(v \otimes w) - y.x.(v \otimes w) &= x.(y.v \otimes w + v \otimes y.w) - y.(x.v \otimes w + v \otimes x.w) = \\ &= x.y.v \otimes w + y.v \otimes x.w + x.v \otimes y.w + v \otimes x.y.w - y.x.v \otimes w - x.v \otimes y.w - y.v \otimes x.w - v \otimes y.x.w \\ &= (x.y.v - y.x.v) \otimes w + v \otimes (x.y.w - y.x.w) = [x, y].v \otimes w + v \otimes [x, y].w \end{aligned}$$

(4)

$$\begin{aligned} (x.(y.\phi))(v) - (y.(x.\phi))(v) &= x.(y.\phi)(v) - (y.\phi)(x.v) - y.(x.\phi)(v) + (x.\phi)(y.v) = \\ &= x.y.\phi(v) - x.\phi(y.v) - y.\phi(x.v) + \phi(y.x.v) - y.x.\phi(v) + y.\phi(x.v) + x.\phi(y.v) - \phi(x.y.v) = \\ &= x.y.\phi(v) - y.x.\phi(v) - \phi(x.y.v - y.x.v) = [x, y].\phi(v) - \phi([x, y].v) \end{aligned}$$

■

Exercise: Use (1) and (3) along with the natural isomorphism of $\text{Hom}(V, W)$ and $V^* \otimes W$ (or $V \otimes W^*$) to prove (4). (Of course you then need an independent proof of (1).)

Examples of some of these constructions for $\mathfrak{sl}(2)$ can be seen in Subsection 13.1 (Small dimension $\mathfrak{sl}(2)$ representations, p. 191) in Chapter 8 (Miscellaneous Calculations and Observations, p. 173)

We can create many representations, besides the dual space, from a single representation V . Let $v_1, \dots, v_n$ be a basis for V . We start with the k^{th} **tensor power** $T^k(V) = V^{\otimes k}$ of V . This has basis the set of simple tensors $v_{i_1} \otimes v_{i_2} \otimes \dots \otimes v_{i_k}$ for arbitrary $i_1, \dots, i_k$ between 1 and n . The k^{th} **symmetric power** $S^k(V) = \text{Sym}^k(V)$ can be defined as either a quotient of $T^k(V)$ or a subspace (when $\mathbb{F}$ has characteristic 0). The latter definition says that $S^k(V)$ consists of tensors that are symmetric with respect to the natural action of the symmetric group Sym_k on $T^k(V)$ (permuting tensor factors). In either case, $S^k(V)$ has basis elements denoted $v_{i_1} v_{i_2} \dots v_{i_k}$ where $i_1 \leq i_2 \leq \dots \leq i_k$ and multiplication mimicking that of polynomials. We can also write the basis elements as $v_1^{j_1} v_2^{j_2} \dots v_n^{j_n}$ where $j_1 + \dots + j_n = k$. The k^{th} **exterior power** (or **alternating power**) $\Lambda^k(V) = \text{Alt}^k(V)$ can be defined as either a quotient of $T^k(V)$ or a subspace (when $\mathbb{F}$ has characteristic 0). The latter definition says that $\Lambda^k(V)$ consists of tensors that are skew-symmetric with respect to the natural action of the symmetric group Sym_k on $T^k(V)$. In either case, $\Lambda^k(V)$ has basis elements denoted $v_{i_1} \wedge v_{i_2} \wedge \dots \wedge v_{i_k}$ where $i_1 < i_2 < \dots < i_k$ and multiplication determined (from concatenation and bilinearity) by the rules $v \wedge v = 0$ and $v \wedge w = -w \wedge v$.

Note that when $\dim V = n$, we have $\dim T^k(V) = n^k$, $\dim S^k(V) = \binom{n+k-1}{k}$, and $\dim \Lambda^k(V) = \binom{n}{k}$. The next result shows how to put a module structure on these powers. All the formulas derive from the basic rule of $x.(v \otimes w) = x.v \otimes w + v \otimes x.w$ for the tensor product of two representations.

Proposition 3.1.9 (Module powers).

Let V be a module over the Lie algebra L . Then the following are modules when given the listed actions with $x \in L$.

- (1) *The tensor power $T^k(V) = V^{\otimes k}$ with $x.(v_1 \otimes \dots \otimes v_k) = \sum_{i=1}^k v_1 \otimes \dots \otimes x.v_i \otimes \dots \otimes v_k$ for $v_1, \dots, v_k \in V$.*
- (2) *The symmetric power $S^k(V) = \text{Sym}^k(V)$ with $x.(v_1 \dots v_k) = \sum_{i=1}^k v_1 \dots x.v_i \dots v_k$ for $v_1, \dots, v_k \in V$.*
- (3) *The exterior power $\Lambda^k(V) = \text{Alt}^k(V)$ with $x.(v_1 \wedge \dots \wedge v_k) = \sum_{i=1}^k v_1 \wedge \dots \wedge x.v_i \wedge \dots \wedge v_k$ for $v_1, \dots, v_k \in V$.*

PROOF. These follow from Proposition 3.1.8 (Combining modules, p. 93)(3) and induction. ■

Some of the above constructions behave well with respect to irreducibility and complete reducibility.

Proposition 3.1.10 (Properties of combined modules).

Let V, W be finite-dimensional modules over the Lie algebra L .

- (1) $(V^*)^* \cong V$.
- (2) V is irreducible if and only if V^* is irreducible.
- (3) If V is completely reducible, so are V^* , $V^{\otimes k}$, $S^k(V)$, and $\Lambda^k(V)$.
- (4) If V, W are completely reducible, so are $V \oplus W$, $V \otimes W$, and $\text{Hom}_{\mathbb{F}}(V, W)$.

REFERENCES AND REMARKS. See [Humphreys, §27.3, statement on p. 161; Jacobson, §III.7, Theorem 11, p. 83; Bourbaki 1-3, §I.3, statement on p. 30; §I.5, Corollaries 1 & 3, pp. 58-59].

PROOF. **Later.** ■

1.6. Semidirect products. Another construction we will run across is the extension of a Lie algebra by a module.

Proposition 3.1.11 (Extension of a Lie algebra by a module).

Let L be a Lie algebra and V be an L -module. Then $L \oplus V$ is a Lie algebra with bracket $[(x, v), (y, w)] = ([x, y], x.w - y.v)$ for $x, y \in L, v, w \in V$. Moreover, $L \oplus 0$ is a subalgebra isomorphic to L and $0 \oplus V$ is an Abelian ideal in this new Lie algebra.

PROOF. The only thing that is not obvious is the Jacobi identity. Let $x, y, z \in L$, $v, w, u \in V$.

$$\begin{aligned} [[(x, v), (y, w)], (z, u)] &= [[([x, y], x.w - y.v), (z, u)] = ([x, y], z], [x, y].u - z.x.w + z.y.v) \\ [[(y, w), (z, u)], (x, v)] &= [[([y, z], y.u - z.w), (x, v)] = ([y, z], x], [y, z].v - x.y.u + x.z.w) \\ [[(z, u), (x, v)], (y, w)] &= [[([z, x], z.v - x.u), (y, w)] = ([z, x], y], [z, x].w - y.z.v + y.x.u) \end{aligned}$$

If we add these three pairs, the first component is 0 by the Jacobi identity. Getting rid of the brackets, the second component becomes

$$x.y.u - y.x.u - z.x.w + z.y.v + y.z.v - z.y.v - x.y.u + x.z.w + z.x.w - x.z.w - y.z.v + y.x.u.$$

All terms cancel and we get 0. Thus the Jacobi identity holds. ■

We can extend the above construction to obtain the notion of a **semidirect product** (or **semidirect sum**) of Lie algebras. Suppose L, M are Lie algebras and M is an L -module such that the action of L is given by a derivation. That is, for $x \in L$, $m, n \in M$ we require that $x.[m, n] = [x.m, n] + [m, x.n]$. This is equivalent to being given a map

$\rho: L \rightarrow \text{Der } M$ and defining the L -action by $x.m = \rho(x)(m)$, or said differently, that M is a representation of L via derivations.

A common way for this to occur is when a Lie algebra $L = S \oplus I$, where S is a subalgebra and I is an ideal. Then S takes the place of L and I takes the place of M in our construction. The action of S on I is via $s.v = \text{ad } s(v) = [s, v]$, so the action is defined by derivations.

Proposition 3.1.12 (Semidirect product of Lie algebras).

Let L, M be Lie algebras and let L act on M via derivations, as defined before the proposition. Then $L \oplus M$ is a Lie algebra with bracket

$$[(x, m), (y, n)] = ([x, y], x.n - y.m + [m, n]) \quad \text{for } x, y \in L, m, n \in M.$$

Moreover, $L \oplus 0$ is a subalgebra isomorphic to L and $0 \oplus M$ is an ideal, isomorphic to M , in this new Lie algebra

REFERENCES AND REMARKS. See [Bourbaki 1-3, §I.1.8, pp. 8-9; Erdmann-Wildon, Exercise 3.9, p. 26; Wikipedia].

PROOF. Imitate the proof of Proposition 3.1.11 (Extension of a Lie algebra by a module, p. 95). **More, maybe?** ■

We may or may not get to the following result, but we state it for completeness. It says every Lie algebra is the semidirect product of a solvable Lie algebra and a semisimple Lie algebra. The result is sometimes known as the Levi-Malcev Theorem, although [Jacobson] refers to it as the Malcev–Harish-Chandra Theorem. It appears (1) is due to Levi and (2), (3) to Malcev and Harish-Chandra.

Proposition 3.1.13 (Levi decomposition).

Let L be a finite-dimensional Lie algebra over a field of characteristic 0 with solvable radical $R = \text{Rad } L$.

- (1) *There is a semisimple subalgebra K of L such that $L = R \oplus K$.*
- (2) *If J, K are semisimple subalgebras of L with $L = R \oplus J = R \oplus K$, then K and J are conjugate by an elementary automorphism of L (a product of automorphisms of the form $e^{\text{ad } x}$, where $x \in \text{Nil } L$).*
- (3) *If J is any semisimple subalgebra of L , then there is a semisimple subalgebra $K \supseteq J$ with $L = R \oplus K$.*

*A subalgebra K as in (1) is called a **Levi factor** or **Levi subalgebra** of L .*

REFERENCES AND REMARKS. See [Jacobson, §III.9, pp. 86-93; Bourbaki 1-3, §I.6.8, Theorem 5, p. 63; Wikipedia].

PROOF. **Later, maybe. Or maybe not.** ■

2. Lie's Theorem

2.1. Irreducible modules over solvable Lie algebras. We begin with a result that will be used more than once.

Proposition 3.2.1 (Invariance lemma).

Let L be a Lie algebra over a field of characteristic 0, let V be a finite-dimensional L -module, and let I be an ideal of L . If $\lambda \in I^*$, define $V_\lambda = \{v \in V \mid x.v = \lambda(x)v \text{ for all } x \in I\}$. Then V_λ is an L -submodule of V . Moreover if $V_\lambda \neq 0$, then $[L, I] \subseteq \ker \lambda$.

REFERENCES AND REMARKS. See [Humphreys, §4.1, step 3 in proof of Theorem, p. 16 and §19, proof of Proposition, p. 102; Erdmann-Wildon, Lemma 5.5, p. 40; Wikipedia, step 4 of proof].

The subspace $V_{[\lambda]} = \{v \in V \mid (x - \lambda(x)I)^n v = 0 \text{ for all } x \in I\}$, where $n = \dim V$, should also be an L -submodule. **Prove.**

PROOF. Let $w \in V_\lambda$ be nonzero and let $y \in L$. We will prove the second claim first.

By definition, $x.w = \lambda(x)w$ for $x \in I$; we use the notation $y^i w$ for $y.(y.(\dots y.w) \dots)$ for i occurrences of $y \in L$ (and $y^0 w = w$). We claim that for each nonnegative integer k , there are $\alpha_{0,k}, \dots, \alpha_{k,k} \in I^*$ with $\alpha_{0,k} = \lambda$ and $x.y^k w = \sum_{i=0}^k \alpha_{i,k}(x)y^{k-i}w$ for any $x \in I$. This is clear for $k = 0$, and for $k = 1$, we have $x.yw = y.x.w + [x, y].w = \lambda(x)yw + \lambda([x, y])w$. If we take $\alpha_{0,1} = \lambda$ and $\alpha_{1,1}(x) = \lambda([x, y])$, then we have the desired formula for $k = 1$. Now suppose the formula holds for k . Then

$$\begin{aligned} x.y^{k+1}w &= x.y.y^k w = y.x.y^k w + [x, y].y^k w = \\ &= y. \sum_{i=0}^k \alpha_{i,k}(x)y^{k-i}w + \sum_{i=0}^k \alpha_{i,k}([x, y])y^{k-i}w = \quad (\text{set } j = i + 1 \text{ in second sum}) \\ &= \alpha_{0,k}(x)y^{k+1}w + \sum_{j=1}^k \alpha_{j,k}(x)y^{k+1-j}w + \sum_{j=1}^k \alpha_{j-1,k}([x, y])y^{k+1-j}w + \alpha_{k,k}([x, y])w \end{aligned}$$

Thus we may take $\alpha_{0,k+1} = \alpha_{0,k} = \lambda$ (as required), $\alpha_{k+1,k+1}(x) = \alpha_{k,k}([x, y])$, and $\alpha_{j,k+1}(x) = \alpha_{j,k}(x) + \alpha_{j-1,k}([x, y])$ for $1 \leq j \leq k$.

Now let U be the span of $w, y.w, y^2 w, \dots$. Since V is finite-dimensional, U is spanned by a finite subset of the $y^i w$. If we let m be the smallest integer such that $y^m w$ is a linear combination of $w, yw, \dots, y^{m-1}w$, a standard argument shows that $w, yw, \dots, y^{m-1}w$ is a basis for U . (Try it.)

Let X and Y denote the operators giving the action of $x \in I$ and $y \in L$ respectively on V , i.e., $X(v) = x.v$, $Y(v) = y.v$. The formulas above for x, y remain valid for X, Y .

These formulas show that U is invariant under the action of any $x \in I$ and that the matrix of $X|_U$, with respect to the given basis, is an upper triangular matrix with every diagonal entry equal to $\lambda(x)$. Thus the trace of $X|_U$ is $m\lambda(x)$. In addition, the action of $[x, y]$ on U is given by $X|_U Y|_U - Y|_U X|_U$, so this action has trace 0. It also has trace $m\lambda([x, y])$. Since $m \neq 0$, this implies $\lambda([x, y]) = 0$. Since $x \in I, y \in L$ were arbitrary, this proves $\lambda([I, L]) = 0$.

Keeping the notation $w \in V_\lambda, x \in I, y \in L$, we now see that

$$x.(y.w) = y.x.w + [x, y].w = \lambda(x)y.w + \lambda([x, y])w = \lambda(x)y.w.$$

This proves $y.w \in V_\lambda$, which shows that V_λ is an L -submodule. ■

Recall that we say a module V is split if every element of L acts on V in such a way that all eigenvalues lie in $\mathbb{F}$. The following results improve our triangularizability results from Chapter 1 (Desirable Forms for Matrices and Operators, p. 17). The next result can be regarded as one version of Lie's theorem. In fact, let us be explicit. Any nonzero element of a 1-dimensional L -module is a common eigenvector for the elements of L . Conversely, if $v \in V$ is a common eigenvector for the elements of L , then the subspace $\mathbb{F}v$ is a submodule. Thus an irreducible module is 1-dimensional if and only if it is spanned by a common eigenvector for the elements of L .

Proposition 3.2.2 (Irreducibles are 1-dimensional for solvable Lie algebras).

Let L be a finite-dimensional solvable Lie algebra over a field of characteristic 0 and let V be a finite-dimensional split irreducible L -module. Then $\dim V = 1$.

REFERENCES AND REMARKS. See [Humphreys, §4.1, Theorem, p. 15; §6, Exercise 3, p. 30; Jacobson, §II.6, Lemma 5, p. 49; Bourbaki 1-3, §I.5.3, Corollary 2, p. 46; Erdmann-Wildon, Proposition 6.6, p. 50; Wikipedia].

PROOF. Since $L/[L, L]$ is a nonzero Abelian Lie algebra, any subspace is an ideal. Thus there is an ideal K of L containing $[L, L]$ and of codimension 1 in L . By induction, V has a 1-dimensional K -submodule W . This means that there is a $\lambda \in K^*$ with $x.w = \lambda(x)w$ for all $x \in K$. Now set $V_\lambda = \{v \in V \mid x.v = \lambda(x)w \text{ for all } x \in K\}$. By Proposition 3.2.1 (Invariance lemma, p. 97), this is an L -submodule of V and so by irreducibility $V_\lambda = V$.

Since $\dim L/K = 1$, there is a $z \in L$ with $L = K \oplus \mathbb{F}z$. Since V is split, z has an eigenvector $w \in V$, say $z.w = \zeta w$. The submodule $\mathbb{F}w \subseteq V$ is invariant under K and under z , so it is invariant under L , i.e., it is an L -submodule. Thus $V = \mathbb{F}w$ and so $\dim V = 1$. ■

For any Lie algebra L , we are interested in classifying the (finite-dimensional) irreducible L -modules. We will spend a great deal of time on this quest for semisimple

Lie algebras in later chapters. When L is solvable and $\mathbb{F}$ is algebraically closed, the previous result tells us all irreducible modules are 1-dimensional. This means (why?) that any irreducible module is annihilated by $[L, L]$. Since $L/[L, L]$ is Abelian, this tells us that irreducible L -modules are in bijective correspondence with the vector space $(L/[L, L])^* \cong L/[L, L]$.

The previous proposition applies to any solvable Lie algebra when $\mathbb{F}$ is algebraically closed. We state without proof a result when $\mathbb{F} = \mathbb{R}$.

Proposition 3.2.3 (Irreducibles for solvable Lie algebras over $\mathbb{R}$).

If $\mathbb{F} = \mathbb{R}$ and L is a finite-dimensional solvable Lie algebra over $\mathbb{F}$, then every finite-dimensional irreducible L -module has dimension 1 or 2.

REFERENCES AND REMARKS. See [Bourbaki 1-3, §I.5.3, Corollary 4, p. 46].

If $\mathbb{F} = \mathbb{Q}$ and L is the Abelian 1-dimensional Lie algebra, then L has irreducible modules of arbitrarily large finite dimension (since $\mathbb{Q}[x]$ has irreducible polynomials of arbitrary degree).

PROOF. Omitted. ■

There is a converse to the above results.

Proposition 3.2.4 (1-dimensional irreducibles implies solvable).

Let L be a finite-dimensional Lie algebra over a field of characteristic 0.

- (1) *Every finite-dimensional irreducible split L -module is 1-dimensional if and only if L is solvable and $\text{ad } x$ is split for every $x \in L$.*
- (2) *Every finite-dimensional irreducible L -module is 1-dimensional if and only if L is solvable and $\mathbb{F}$ is algebraically closed.*

PROOF. (1) First assume all finite-dimensional irreducible split L -modules are 1-dimensional. We prove $\text{ad } L$ is solvable. Consider L as an $\text{ad } L$ -module and let $0 = L_0 \subseteq L_1 \subseteq \cdots \subseteq L_n = L$ be a composition series for L . Then each L_i is an ideal of L . By hypothesis, each L_i/L_{i-1} is 1-dimensional: let $x_i + L_{i-1}$ span L_i/L_{i-1} . Then $x_1, \dots, x_n$ is a basis for L , and with respect to that basis, every element of $\text{ad } L$ in $\mathfrak{gl}(L)$ is upper triangular. (This uses the fact that L_{i-1} is an ideal.) Since L is a faithful $\text{ad } L$ -module, it follows that $\text{ad } L$ embeds in the Lie algebra $\mathfrak{t}(n)$ of upper triangular matrices, which is solvable by Proposition 2.3.5 (Derived and central series of triangular matrices, p. 56). By Proposition 2.4.1 (Properties of L from $\text{ad } L$, p. 61), this implies L is solvable. Moreover, each $\text{ad } x$ is upper triangular with respect to this basis and so is split.

The converse follows from Proposition 3.2.2 (Irreducibles are 1-dimensional for solvable Lie algebras, p. 98).

(2) Suppose all finite-dimensional irreducible modules are 1-dimensional. Then (1) shows L is solvable. Suppose $\mathbb{F}$ is not algebraically closed. Since L is solvable, $[L, L] \neq L$, whence $L/[L, L]$ is a nonzero Abelian Lie algebra. By Proposition 3.1.2 (Representations of Abelian Lie algebras, p. 87)(5), $L/[L, L]$ has a finite-dimensional irreducible representation of dimension greater than 1. This is also an irreducible representation of L .

The converse follows from Proposition 3.2.2 (Irreducibles are 1-dimensional for solvable Lie algebras, p. 98), since all finite-dimensional modules are split over an algebraically closed field. ■

I believe that if there is a bound on the dimensions of the (finite-dimensional) irreducible L -modules, then L is solvable. This is certainly true if $\mathbb{F}$ is algebraically closed or $\mathbb{F} = \mathbb{R}$. This is because every semisimple Lie algebra has irreducible representations of arbitrarily large (finite) dimension.

2.2. Solvability and triangularizability. In this subsection, we prove several results that have some claim to be called Lie's theorem.

Proposition 3.2.5 (Lie's theorem for representations).

Let L be a finite-dimensional solvable Lie algebra over a field of characteristic 0, let V be a finite-dimensional vector space, and let $\phi : L \rightarrow \mathfrak{gl}(V)$ be a split representation of L . Then there is a basis of V with respect to which $\phi(x)$ is upper triangular for all $x \in L$.

REFERENCES AND REMARKS. See [Jacobson, §II.6, Lie's theorem, p. 50].

PROOF. Let $0 = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_n = V$ be a composition series for V . Since V is split, each V_i/V_{i-1} is split, and so by Proposition 3.2.4 (1-dimensional irreducibles implies solvable, p. 99), each V_i/V_{i-1} is 1-dimensional. Thus $0 = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_n = V$ is a full flag of L -submodules. If we choose a basis for V consistent with this flag (i.e., choose $v_i \in V_i$ so that $v_i + V_{i-1}$ spans V_i/V_{i-1} and thus $v_1, \dots, v_i$ is a basis for V_i for $i = 1, \dots, n$), then the matrix of $\phi(x)$ with respect to this basis is upper triangular for each $x \in L$. ■

Proposition 3.2.6 (Concrete Lie's theorem).

Let V be a finite-dimensional vector space over a field of characteristic 0 and let $L \subseteq \mathfrak{gl}(V)$ be a split linear Lie algebra. Then L is triangularizable if and only if L is solvable. If L is solvable, then $[L, L]$ is nilpotent as a Lie algebra and consists of nilpotent operators.

REFERENCES AND REMARKS. See [Humphreys, §4.1, pp. 15-17; Jacobson, §II.6, p. 50; Bourbaki 1-3, §I.5.3, Corollary 2, p. 46; Erdmann-Wildon, §6.4, Theorem 6.5, pp. 49-51; Wikipedia].

PROOF. If L is triangularizable, then L is a subalgebra of the Lie algebra $\mathfrak{t}(n)$ of upper triangular matrices, with respect to the triangularization basis. Proposition 2.3.5 (Derived

and central series of triangular matrices, p. 56) tells us that $\mathfrak{t}(n)$ is solvable, whence L is solvable. Moreover, $[L, L]$ is a subalgebra of $\mathfrak{n}(n)$, whence $[L, L]$ consists of nilpotent linear operators by Proposition 8.2.1 (Multiplication of triangular matrices, p. 174)(4) and is nilpotent as a Lie algebra by Proposition 2.3.5 (Derived and central series of triangular matrices, p. 56)(3).

Assume now that L is solvable. Applying Proposition 3.2.5 (Lie's theorem for representations, p. 100) to the natural representation V , we see that L is triangularizable.

For reference, we include the proof from [Humphreys, §4.1, pp. 15-16] that solvable implies triangularizable, without using composition series and irreducible representations. Suppose that L is solvable and every linear, split solvable Lie algebra of dimension less than $\dim L$ is solvable. Suppose also that any solvable subalgebra of $\mathfrak{gl}(W)$ for $\dim W < \dim V$ is triangularizable. (We can think of this as induction on $\dim L + \dim V$.) Since $[L, L]$ is properly contained in L and $L/[L, L]$ is Abelian, there must be an ideal of $L/[L, L]$ of codimension 1. This is equivalent to the existence of an ideal K of L with $\dim L/K = 1$. Since K is solvable, the first induction hypothesis implies K is triangularizable. In particular, there is a common eigenvector $w_0 \in V$ for the elements of K . This means there is a linear functional $\lambda : K \rightarrow \mathbb{F}$ such that $x(w_0) = \lambda(x)w_0$ for all $x \in K$. Set $V_\lambda = \{w \in V \mid x(w)\lambda(x)w \text{ for all } x \in K\}$. By Proposition 3.2.1 (Invariance lemma, p. 97), V_λ is an L -submodule of V and $\lambda([L, K]) = 0$.

Since $\dim L/K = 1$, there is a $z \in L$ with $L = K \oplus \mathbb{F}z$. As z is triangularizable, so is $z|_{V_\lambda}$. If w is any eigenvector for z in V_λ , then w is also an eigenvector for K and hence for all of L . Thus $\mathbb{F}w$ is a submodule of V , and so we can form the L -module $V/\mathbb{F}w$. Let $I = \{x \in L \mid x(L) \subseteq \mathbb{F}w\}$. Then I is an ideal of L and L/I embeds in $\mathfrak{gl}(V/\mathbb{F}w)$. Moreover, L/I is identical with the set $\bar{L}$ of operators on $V/\mathbb{F}w$ induced by the operators in L . By the second induction hypothesis, L/I is triangularizable on the vector space $V/\mathbb{F}w$. By Proposition 1.3.2 (Desirable forms on subspaces and quotient spaces, p. 26), we see that L is triangularizable on the vector space V . ■

One way to state an abstract version of Lie's theorem is the following.

Proposition 3.2.7 (Lie's theorem on complete solvability).

Let L be a finite-dimensional Lie algebra over a field of characteristic 0. If L is solvable and $\text{ad } x$ is split for all $x \in L$, then L is completely solvable. That is, there is a chain $0 = L_0 \subseteq L_1 \subseteq \cdots \subseteq L_n = L$ of ideals of L such that $\dim L_i = i$.

REFERENCES AND REMARKS. See [Humphreys, §4.1, Corollary B, p. 16; Jacobson, §II.7, Theorem 14, p. 52; Bourbaki 1-3, §I.5.3, Corollary 3, p. 46].

Recall that another name for a chain of ideals as described above is a full flag of ideals or a full invariant flag.

PROOF. Apply Proposition 3.2.5 (Lie's theorem for representations, p. 100) to the regular L -module, i.e., the representation $\text{ad} : L \rightarrow \mathfrak{gl}(L)$, and note that L -submodules of L are ideals. ■

2.3. Solvability and nilpotence. Another important consequence of Lie's theorem is the following.

Proposition 3.2.8 (Solvable is close to nilpotent).

Let L be a finite-dimensional solvable Lie algebra over a field of characteristic 0.

- (1) *The derived subalgebra $[L, L]$ is nilpotent.*
- (2) *If N is the nil radical of L and δ is a derivation on L , then $\text{im } \delta \subseteq N$.*

REFERENCES AND REMARKS. See [Humphreys, §4.2, Corollary C, p. 16; Jacobson, §II.7, Theorem 13 and Corollaries 1 & 2, pp. 51-52; Bourbaki 1-3, §I.5.3, Theorem 1, p. 45 and §1.5.5, Proposition 6, p. 49].

PROOF. (1) is a consequence of (2), but we prove (1) first and then (2). We assume $\mathbb{F}$ is algebraically closed for the proof of (1).

Note that by Proposition 3.2.7 (Lie's theorem on complete solvability, p. 101), there is a sequence of ideals $0 = L_0 \subseteq L_1 \subseteq \cdots \subseteq L_n = L$ of ideals of L such that $\dim L_i = i$. If we let $\text{ad } L \subseteq \mathfrak{gl}(L)$ act on this series, we see that $\text{ad } L$ is triangularizable. By Proposition 3.2.6 (Concrete Lie's theorem, p. 100), we know that $[\text{ad } L, \text{ad } L]$ is nilpotent. Since ad is a Lie algebra homomorphism, $[\text{ad } L, \text{ad } L] = \text{ad}[L, L]$, and so $\text{ad}[L, L]$ is nilpotent. By Proposition 2.4.1 (Properties of L from $\text{ad } L$, p. 61), this implies $[L, L]$ is nilpotent.

For an alternative finish to the proof, triangularizability of $\text{ad } L$ implies it embeds in $\mathfrak{t}(n)$ ($n = \dim L$). Thus $[\text{ad } L, \text{ad } L]$ embeds in $[\mathfrak{t}(n), \mathfrak{t}(n)] = \mathfrak{n}(n)$, and so each element of $[\text{ad } L, \text{ad } L]$ is a nilpotent operator on L . Since ad is a Lie algebra homomorphism, $[\text{ad } L, \text{ad } L] = \text{ad}[L, L]$, and so $\text{ad}_L x$ is nilpotent for each $x \in [L, L]$. Thus $\text{ad}_{[L, L]} x$ is nilpotent for each $x \in [L, L]$. Engel's theorem, Proposition 2.4.4 (Engel's theorem for ad , p. 64), implies that $[L, L]$ is nilpotent.

If $\mathbb{F}$ is not algebraically closed, we use extension of scalars. Set $\mathbb{E} = \overline{\mathbb{F}}$. By Proposition 2.8.2 (Derived/central series and extension of scalars, p. 80), we have that $[L, L]_{\mathbb{E}} = [L_{\mathbb{E}}, L_{\mathbb{E}}]$. The right hand side is nilpotent by the algebraically closed case. Thus by Proposition 2.8.3 (Types of Lie algebras and extension of scalars, p. 80), $[L, L]$ is nilpotent.

(2) We can regard $\mathbb{F}\delta$ as a 1-dimensional Lie algebra with Lie bracket zero. We can thus define a Lie algebra structure on $M = \mathbb{F}\delta \oplus L$ as in Proposition 3.1.12 (Semidirect product of Lie algebras, p. 96), with $[(\lambda\delta, x), (\mu\delta, y)] = (0, \delta(\lambda y - \mu x) + [x, y])$ for scalars λ, μ . Let us identify L with $0 \times L$: then it's easy to see that $[M, M] = [L, L] + \text{im } \delta$.

Since $\mathbb{F}\delta$ and $M/\mathbb{F}\delta \cong L$ are solvable, we see that M is solvable. Thus by (1), $[M, M]$ is nilpotent. Thus $\text{im } \delta \subseteq [L, L] + \text{im } \delta \subseteq N$. ■

A stronger version of this result is true. We omit the proof of the hardest result (2).

Proposition 3.2.9 (Derivations take solvable to nilpotent).

Let L be a finite-dimensional Lie algebra over a field of characteristic 0. Let R be the solvable radical and N the nil radical of L .

- (1) $[L, R] \subseteq N$.
- (2) *If δ is a derivation on L , then $\delta(R) \subseteq N$.*
- (3) R and N are derivation and automorphism characteristic ideals of L .

REFERENCES AND REMARKS. See [Jacobson, §III.6, Theorem 7, pp. 74-75; Bourbaki 1-3, §1.5.5, Proposition 6, p. 49].

PROOF. (1) This follows from (2), but we are not providing a proof of (2).

If $x \in L$, then $[x, R] \subseteq R$. Thus $\text{ad } x$ is a derivation on R . By Proposition 3.2.8 (Solvable is close to nilpotent, p. 102), we see that $\text{ad } x(R) \subseteq N$, i.e., $[x, R] \subseteq N$.

(2) Omitted. ???

(3) Clearly R and N are automorphism characteristic. Since $\delta(R) \subseteq N \subseteq R$ and $\delta(N) \subseteq \delta(R) \subseteq N$ for any derivation δ , we see that R, N are derivation characteristic. ■

3. Representations of $\mathfrak{sl}(2)$

The representation theory of $L = \mathfrak{sl}(2)$ turns out to be fundamental to the theory of semisimple Lie algebras. Let e, h, f be the usual basis of L , with $[e, f] = h$, $[h, e] = 2e$, $[h, f] = -2f$. The reader who wishes to see some small dimensional examples before diving into the general theory may consult Subsection 13.1 (Small dimension $\mathfrak{sl}(2)$ representations, p. 191) in Chapter 8 (Miscellaneous Calculations and Observations, p. 173).

3.1. Concrete construction of irreducible representations. We begin with an example that involves a slight dip into the infinite-dimensional case. We consider linear operators on the polynomial ring $\mathbb{F}[x, y]$. Three operators of particular interest to us are $E = x \frac{\partial}{\partial y}$, $H = x \frac{\partial}{\partial x} - y \frac{\partial}{\partial y}$, $F = y \frac{\partial}{\partial x}$. Thus, for example, $E(x^3 y^2) = 2x^4 y$, $H(x^3 y^2) = 3x^3 y^2 - 2x^3 y^2 = x^3 y^2$, $F(x^3 y^2) = 3x^2 y^3$. Explicitly, we have

$$E(x^i y^j) = j x^{i+1} y^{j-1} \quad H(x^i y^j) = (i - j) x^i y^j \quad F(x^i y^j) = i x^{i-1} y^{j+1}.$$

Note that E, F, H all preserve total degree, and so they leave invariant the space V_n of homogeneous polynomials of degree n (plus 0), spanned by $x^n, x^{n-1}y, \dots, y^n$. One can show (see Subsection 13.2 (E, F, H relations, p. 192) in Chapter 8 (Miscellaneous Calculations and Observations, p. 173)) that $EF - FE = H$, $HE - EH = 2E$, $HF - FH = -2F$, so

the map $e \mapsto E, h \mapsto H, f \mapsto F$ gives us a(n infinite-dimensional) representation of L on $\mathbb{F}[x, y]$. If we restrict ourselves to V_n , we get an $n + 1$ -dimensional representation of L .

Before proceeding, let's look at the action of E, F, H on V_n for $n = 3, 4$.

basis	V_3				V_4				
	x^3	x^2y	xy^2	y^3	x^4	x^3y	x^2y^2	xy^3	y^4
$E = x \frac{\partial}{\partial y}$	0	x^3	$2x^2y$	$3xy^2$	0	x^4	$2x^3y$	$3x^2y^2$	$4xy^3$
$H = x \frac{\partial}{\partial x} - y \frac{\partial}{\partial y}$	$3x^3$	x^2y	$-xy^2$	$-3y^3$	$4x^4$	$2x^3y$	0	$-2xy^3$	$-4y^4$
$F = y \frac{\partial}{\partial x}$	$3x^2y$	$2xy^2$	y^3	0	$4x^3y$	$3x^2y^2$	$2xy^3$	y^3	0
H -eigenvalue	3	1	-1	-3	4	2	0	-2	-4

Each of the representations V_n is irreducible. To see this, note that (for $i + j = n$), $x^i y^j$ is an eigenvector for H with eigenvalue $i - j$ and these eigenvalues are distinct for pairs (i, j) with $i + j = n$. A standard procedure now enables one to reduce a nonzero sum of monomials to a single monomial by repeatedly applying H and taking linear combinations — see Proposition 1.1.5 (Consequences of distinct eigenvalues, p. 22). From a single monomial, we can apply E and F repeatedly to get any other monomial of degree n . Thus from any nonzero element of V_n we can get any other nonzero element by repeatedly applying E, F, H and taking linear combinations. (See Subsection 13.3 (Irreducibility of V_n , p. 193) in Chapter 8 (Miscellaneous Calculations and Observations, p. 173) for more detail.) It follows that V_n is an irreducible L -module. We will see below that every (finite-dimensional) L -module is a direct sum of copies of the V_n .

3.2. Abstract construction of irreducible representations. We now construct the modules V_n from a different point of view, starting from within a given representation V of L . Suppose $v' \in V$ is an eigenvector for h , say $h.v' = \lambda'v'$. Then $h.(e.v') = e.(h.v') + [h, e].v' = e.(\lambda'v') + 2e.v' = (\lambda' + 2)e.v'$. A similar calculation shows $h.(f.v') = (\lambda' - 2)f.v'$. Thus the action of e, f turn h -eigenvectors into either 0 or h -eigenvectors with different eigenvalues. For convenience, let us write $e^k v' = e.(e \dots (e.v') \dots)$, with k instances of e , with similar notation for f . Eigenvectors with different eigenvalues are linearly independent and V is finite-dimensional, so $v', e.v', e^2 v', \dots$ cannot all be nonzero. Choose a $k \geq 1$ so that $e^{k-1} v' \neq 0 = e^k v'$. We now replace v' by $v = e^{k-1} v'$ and λ' by $\lambda = \lambda' + 2(k - 1)$.

Thus we can assume v is an eigenvector for h with eigenvalue λ and $e.v = 0$. We will call such a v a **highest weight vector**. ([Humphreys, §7.1, p. 32] uses the term **maximal vector**.) As before, $f^i v$ is 0 or an h -eigenvector with eigenvalue $\lambda - 2i$. Let us set $v_i = f^i v$, so $v_0 = v, v_1 = f.v$, etc., and v_i is an h -eigenvector with eigenvalue $\lambda - 2i$. As before, $v_0, v_1, v_2, \dots$ cannot all be nonzero, so there is an $n \geq 0$ with $v_n \neq 0 = v_{n+1}$.

Let W denote the span of $v_0, v_1, v_2, \dots, v_n$. We claim W a submodule of V ; clearly it is closed under the action of f and h . Now $e.v_0 = 0$, $e.v_1 = e.f.v = f.e.v + [e, f]v = \lambda v_0$. In general, $e.v_i = i(\lambda + 1 - i)v_{i-1}$. Suppose this is true for a particular i and note that $e.v_{i+1} = e.f.v_i = f.e.v_i + [e, f].v_i = f.i(\lambda + 1 - i)v_{i-1} + h.v_i = i(\lambda + 1 - i)v_i + (\lambda - 2i)v_i = ((i + 1)\lambda + i - i^2 - 2i)v_i = (i + 1)(\lambda + 1 - (i + 1))v_i$, as desired. Thus W is indeed a submodule.

The observation $e.v_i = i(\lambda + 1 - i)v_{i-1}$ remains true for $i = n + 1$, in which case it says $0 = (n + 1)(\lambda + 1 - n - 1)v_n$. Since $n \geq 0$ and $v_n \neq 0$, we conclude that $\lambda = n = \dim W - 1$. Thus λ is always a nonnegative integer.

Let's look at the action of e, f, h on W for $m = \dim W = 4, 5$.

basis	dim $W = 4$				dim $W = 5$				
	v_0	v_1	v_2	v_3	v_0	v_1	v_2	v_3	v_4
e	0	$3v_0$	$4v_1$	$3v_2$	0	$4v_0$	$6v_1$	$6v_2$	$4v_3$
h	$3v_0$	v_1	$-v_2$	$-3v_3$	$4v_0$	$2v_1$	0	$-2v_3$	$-4v_4$
f	v_1	v_2	v_3	0	v_1	v_2	v_3	v_4	0
h -eigenvalue	3	1	-1	-3	4	2	0	-2	-4

In fact, W is isomorphic to the module V_n above. To see this, let us define $\alpha_0 = 1, \alpha_i = n(n-1) \cdots (n-(i-2))n - (i-1) \cdots n$ for $1 \leq i \leq n$. Thus $\alpha_{i+1} = (n-i)\alpha_i$. Now define $\theta : W \rightarrow V_{m-1}$ by $\theta(v_i) = \alpha_i x^{n-i} y^i$. This is a linear isomorphism. The following calculations prove θ is an isomorphism of modules.

$$\theta(h.v_i) = \theta((n-2i)v_i) = (n-2i)\alpha_i x^{n-i} y^i = H(\theta(v_i))$$

$$\theta(f.v_i) = \theta(v_{i+1}) = \alpha_{i+1} x^{n-(i+1)} y^{i+1} = (n-1)\alpha_i x^{n-i-1} y^{i+1} = F(\theta(v_i))$$

$$\theta(e.v_i) = \theta(i(n+1-i)v_{i-1}) = i(n+1-i)\alpha_{i-1} x^{n-(i-1)} y^{i-1} = \alpha_i i x^{n+1-i} y^{i-1} = E(\theta(v_i)).$$

Thus any (finite-dimensional) irreducible L -module such that the action of h has an eigenvector is one of the V_n 's, and in fact, is uniquely determined by its dimension. Moreover, all the eigenvalues of h are integers. If $\mathbb{F}$ is algebraically closed, then eigenvectors always exist, so in this case we have a complete classification of irreducible modules.

In fact, the above is a complete classification for any field $\mathbb{F}$ (of characteristic 0). To see this, let V be an irreducible L -module and let $\overline{\mathbb{F}}$ be the algebraic closure of $\mathbb{F}$. The action of h on the $L_{\overline{\mathbb{F}}}$ -module $V_{\overline{\mathbb{F}}}$ obtained by extension of scalars has the same matrix as the action of h on the L -module V , and so the extended action has the same eigenvalues as the original action. We've seen above that eigenvalues of the extended action are all integers, so they lie in $\mathbb{F}$. This means the h -action on V has eigenvectors and so we may repeat the arguments above to show V is uniquely one of the V_n .

To sum up an important part of the above: if n is even, the eigenvalues of h acting on V_n are $-n, -(n-2), \dots, -2, 0, 2, \dots, n-2, n$, while if n is odd, the eigenvalues are $-n, -(n-2), \dots, -1, 1, \dots, n-2, n$. We call these eigenvalues **weights** and the corresponding eigenvectors are called **weight vectors**. Thus for any irreducible $\mathfrak{sl}(2)$ -module, a weight of either 0 or 1 must occur, and the number of irreducible modules in the decomposition of any completely reducible $\mathfrak{sl}(2)$ -module equals the number of linearly independent weight vectors of weight 0 plus the number of linearly independent weight vectors of weight 1.

Warning: For semisimple Lie algebras bigger than $\mathfrak{sl}(2)$, the weights are not numbers but rather linear functionals on a subalgebra called the Cartan subalgebra.

Let us summarize our results in a formal statement.

Proposition 3.3.1 (Irreducible representations of $\mathfrak{sl}(2)$).

Let $\mathbb{F}$ have characteristic 0 and let $L = \mathfrak{sl}(2, \mathbb{F})$.

- (1) L has, up to isomorphism, exactly one irreducible module of each positive dimension.
- (2) Any finite-dimensional irreducible L -module has a basis of h -eigenvectors and each h -eigenspace has dimension 1.
- (3) The h -eigenvalues of a finite-dimensional irreducible L -module form an arithmetic progression with difference 2, and the dimension of the module is 1 plus the largest h -eigenvalue.

REFERENCES AND REMARKS. See [Humphreys, §7, pp. 31-34; Jacobson, §III.8, pp. 83-86; Bourbaki 7-9, §VIII.1, pp. 70-74; Erdmann-Wildon, §8.1 & 8.2, pp. 67-73; Fulton-Harris, §11.1, pp. 146-150; Wikipedia].

PROOF. It's up above. ■

3.3. Vista: infinite dimensional irreducible representations. We briefly consider infinite-dimensional representations. Let λ be an element of $\mathbb{F}$ that is not a nonnegative integer. We can define an $\mathfrak{sl}(2)$ -module with basis $v_0, v_1, \dots$ and action as follows for $i \geq 0$:

$$h.v_i = (\lambda - 2i)v_i \quad f.v_i = v_{i+1} \quad e.v_i = i(\lambda + 1 - i)v_{i-1}$$

To check that this does define an $\mathfrak{sl}(2)$ -module, we must check the actions of e, f, h and see that the relations of $\mathfrak{sl}(2)$ are obeyed.

Since each v_i is an h -eigenvector with a unique eigenvalue, any nonzero submodule of V must contain some v_j , just as in the proof of Proposition 1.1.5 (Consequences of distinct eigenvalues, p. 22). The action of e on v_i for $i > 0$ is never 0, since $\lambda \neq i + 1$ for any i . This means that if a submodule contains v_j for $j > 0$, then it contains v_0 . If a submodule contains v_0 , then the action of f shows that it contains every v_i , i.e., it equals

V . This shows V is an infinite-dimensional irreducible $\mathfrak{sl}(2)$ -module. The module V is called a Verma module. It can also be defined for λ a nonnegative integer, in which case it is not irreducible but has the irreducible module with highest weight λ as a factor.

There are many infinite-dimensional irreducible $\mathfrak{sl}(2)$ -modules that do not arise from this construction.

3.4. Generating irreducible representations from the natural representation.

Let V be the natural representation of $\mathfrak{sl}(2)$, that is, the representation as $\mathbb{F}^2$ with $\mathfrak{sl}(2)$ acting as matrix multiplication. This is often called the standard representation instead. Let v, w be a basis of V with v a highest weight vector (so $h.v = v$, $e.v = 0$) and $w = f.v$. We will show how to derive irreducible representations from powers of V , as in Subsection 1.5 (New modules from old, p.93). Recall that $V^{\otimes n}$ denotes the tensor product of n copies of V . Then it is easy to see that $v' = v \otimes v \otimes \cdots \otimes v$ is a highest weight vector in $V^{\otimes n}$: we have $h.v' = nv'$ and $e.v' = 0$. (For example, $h.(v \otimes v \otimes v) = h.v \otimes v \otimes v + v \otimes h.v \otimes v + v \otimes v \otimes h.v = 3(v \otimes v \otimes v)$ and $e.(v \otimes v \otimes v) = e.v \otimes v \otimes v + v \otimes e.v \otimes v + v \otimes v \otimes e.v = 0$.) Thus v' generates an irreducible submodule of $V^{\otimes n}$ of dimension $n + 1$. This shows all (finite-dimensional) irreducible $\mathfrak{sl}(2)$ -modules occur as submodules of some $V^{\otimes n}$.

The representation $V^{\otimes n}$ is generally a direct sum of more than just one irreducible representation. For example, $V \otimes V$ contains $v' = v \otimes v$, which generates a 3-dimensional irreducible representation V' (with basis $v \otimes v$, $\frac{1}{2}(v \otimes w + w \otimes v)$, $w \otimes w$) and $v'' = v \otimes w - w \otimes v$, which generates a 1-dimensional irreducible representation V'' . Moreover, $V \otimes V = V' \oplus V''$.

As above, let V_n denote the $(n + 1)$ -dimensional irreducible representation of $\mathfrak{sl}(2)$ and define $V^{\otimes 0} = V_0$ and $V^{\otimes 1} = V_1 = V$. We have the following decompositions.

$$V^{\otimes 0} = V_0 \quad V^{\otimes 1} = V_1 \quad V^{\otimes 2} \cong V_2 \oplus V_0 \quad V^{\otimes 3} \cong V_3 \oplus V_1^{\oplus 2} \quad V^{\otimes 4} \cong V_4 \oplus V_2^{\oplus 3} \oplus V_0^{\oplus 2}$$

Recall that the symmetric power $S^n(V) = \text{Sym}^n(V)$ has a basis of the form $v^i w^j$ where $i + j = n$ and the action of $x \in \mathfrak{sl}(2)$ is given by $x.v^i w^j = ix.v v^{i-1} w^j + jx.w v^i w^{j-1}$. Plainly v^n is a highest weight vector with $h.v^n = nv^n$, so it generates an irreducible submodule of dimension $n + 1$. Since $\dim S^n(V) = n + 1$, we see that $S^n(V)$ is an irreducible representation of dimension $n + 1$. It follows that the symmetric powers of V give us all (finite-dimensional) irreducible representations of $\mathfrak{sl}(2)$.

4. Representations of $\mathfrak{sl}(3)$

This material is covered in [Fulton-Harris, Lectures 12 & 13, pp.161-193].

We described the irreducible representations of $\mathfrak{sl}(2)$ in great detail in the previous section. In this section, we tackle $\mathfrak{sl}(3)$ and learn something about the similarities and

differences. We will prove in Proposition 3.5.3 (Weyl's theorem on complete reducibility, p. 118) that all representations of $\mathfrak{sl}(2)$ and $\mathfrak{sl}(3)$ are completely reducible; we freely make use of that fact in the current section.

The techniques developed in this section carry over to the theory of representations of general semisimple Lie algebras.

4.1. Roots, weights, and representations for $\mathfrak{sl}(3)$. Let $L = \mathfrak{sl}(3)$ and label the elements as follows.

$$h_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad h_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad h_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

$$e_1 = e_{12} \quad e_2 = e_{23} \quad e_3 = e_{13} \quad f_1 = e_1^T = e_{21} \quad f_2 = e_2^T = e_{32} \quad e_3 = e_3^T = e_{31}$$

We can condense most of this into $L = \begin{pmatrix} H & e_1 & e_3 \\ f_1 & H & e_2 \\ f_3 & f_2 & H \end{pmatrix}$. The elements e_i, h_i, f_i span L .

We don't really need h_3 since it equals $h_1 + h_2$, but we include it for symmetry. Clearly $e_1, e_2, e_3, h_1, h_2, f_1, f_2, f_3$ is a basis for L . Let H be the set of diagonal matrices in L , so H has basis h_1, h_2 .

Here is a multiplication/bracket table for $\mathfrak{sl}(3)$, including the redundant h_3 .

$[\cdot, \cdot]$	e_1	e_2	e_3	h_1	h_2	h_3	f_1	f_2	f_3
e_1	0	e_3	0	$-2e_1$	e_1	$-e_1$	h_1	0	$-f_2$
e_2	$-e_3$	0	0	e_2	$-2e_2$	$-e_2$	0	h_2	f_1
e_3	0	0	0	$-e_3$	$-e_3$	$-2e_3$	$-e_2$	e_1	h_3
h_1	$2e_1$	$-e_2$	e_3	0	0	0	$-2f_1$	f_2	$-f_3$
h_2	$-e_1$	$2e_2$	e_3	0	0	0	f_1	$-2f_2$	$-f_3$
h_3	e_1	e_2	$2e_3$	0	0	0	$-f_1$	$-f_2$	$-2f_3$
f_1	$-h_1$	0	e_2	$2f_1$	$-f_1$	f_1	0	$-f_3$	0
f_2	0	$-h_2$	$-e_1$	$-f_2$	$2f_2$	f_2	f_3	0	0
f_3	f_2	$-f_1$	$-h_3$	f_3	f_3	$2f_3$	0	0	0

For each $i = 1, 2, 3$, let L_i be the span of e_i, h_i, f_i . The above bracket table shows that L_i is a subalgebra of L isomorphic to $\mathfrak{sl}(2)$. If V is a representation of L , then it is automatically a representation of each L_i . By Proposition 3.5.3 (Weyl's theorem on complete reducibility, p. 118), V is a direct sum of irreducible L_i -representations for each i . In Section 3 (Representations of $\mathfrak{sl}(2)$, p. 103), we saw that each such irreducible representation of L_i is a direct sum of h_i -eigenspaces, with integral eigenvalues. Thus V is a direct sum of h_i -eigenspaces with integral eigenvalues for each i .

Since H is Abelian, we know that the images of H in the representation V are simultaneously diagonalizable. Thus there is a basis $v_1, \dots, v_n$ of V and linear functionals $\lambda_1, \dots, \lambda_n \in H^*$ such that $h.v_j = \lambda_j(h)v_j$ for all $j = 1, \dots, n$. We saw above that $\lambda_j(h_i)$ is an integer for all i, j . We call these functionals λ_j **eigenvalues** or **weights** and we call the v_i **eigenvectors** or **weight vectors**. The subspace of V consisting of 0 plus all weight vectors of a given weight is called an **eigenspace** or **weight space**.

This is the big difference between $\mathfrak{sl}(2)$ -modules and modules over other semisimple Lie algebras: eigenvalues go from being numbers to being linear functionals. In some ways, this is an easy generalization, but in other ways it is much more complicated. For example, integers are ordered, but how do we order linear functionals?

Let's first consider some special representations of $L = \mathfrak{sl}(3)$. We know the 1-dimensional representation is trivial, so all weights are 0. Next, let V be the natural 3-dimensional representation (with action by matrix multiplication). Consider the unit vectors $\mathbf{e}_j$. (E.g., $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$; we use bold to distinguish these from the vectors $e_i \in \mathfrak{sl}(3)$ defined above.) It's easy to see that these are weight vectors with weights $\beta_1, \beta_2, \beta_3$, where

$$\begin{array}{lll} \beta_1(h_1) = 1 & \beta_1(h_2) = 0 & \beta_1(h_3) = 1 \\ \beta_2(h_1) = -1 & \beta_2(h_2) = 1 & \beta_2(h_3) = 0 \\ \beta_3(h_1) = 0 & \beta_3(h_2) = -1 & \beta_3(h_3) = -1 \end{array}$$

The weights β_i can be described by the formula $\beta_i\left(\begin{pmatrix} d_1 & & \\ & d_2 & \\ & & d_3 \end{pmatrix}\right) = d_i$. Note that $\beta_1 + \beta_2 + \beta_3 = 0$ and β_1, β_2 is a basis for H^* . [Fulton-Harris, Lecture 12, p. 163] uses the notation L_i in place of β_i .

It's not hard to see directly that V is irreducible. Since the weights β_j are distinct, we can also show that V is irreducible as in the $\mathfrak{sl}(2)$ -case, using Proposition 1.3.3 (Consequences of distinct eigenvalues for sets, p. 28).

Note that $\mathbf{e}_1$ is annihilated by all the upper triangular matrix units e_i , while $f_1\mathbf{e}_1 = \mathbf{e}_2$, $f_2\mathbf{e}_2 = \mathbf{e}_3$. Thus V is generated by $\mathbf{e}_1$, which is annihilated by all the e_i . Moreover, it is a weight vector with weight $\beta_1(h_1)$ being larger than $\beta_i(h_1)$ for $i = 2, 3$. Thus $\mathbf{e}_1$ might be called a highest weight vector in V . We will be more explicit about the meaning of "highest weight vector" below.

Now consider L as an L -module via the adjoint representation (so it is an 8-dimensional representation). Since submodules under the adjoint representation are ideals, the adjoint representation is irreducible for any simple Lie algebra. When dealing with the adjoint representation, it is traditional to call the weights **roots** and to call the weight vectors

root vectors and the weight spaces **root spaces**, with the exception that we do not call 0 a root and we do not call the corresponding 0-eigenspace (which is H) a root space. We will use α_j in place of λ_j as well. This means the root vectors are $e_1, e_2, e_3, f_1, f_2, f_3$ and the roots are defined by $[h_i, e_j] = \alpha_j(h_i)e_j$.

The multiplication table for $\mathfrak{sl}(2)$ gives us the following.

$$\begin{array}{lll} \alpha_1(h_1) = 2 & \alpha_1(h_2) = -1 & \alpha_1(h_3) = 1 \\ \alpha_2(h_1) = -1 & \alpha_2(h_2) = 2 & \alpha_2(h_3) = 1 \\ \alpha_3(h_1) = 1 & \alpha_3(h_2) = 1 & \alpha_3(h_3) = 2 \end{array}$$

We also have α_j for $j = 4, 5, 6$, defined by $[h, f_{j-3}] = \alpha_j(h)f_{j-3}$. The multiplication table shows that $\alpha_j = -\alpha_{j-3}$ for $j = 4, 5, 6$. It also is the case that $\alpha_3 = \alpha_1 + \alpha_2$. Thus the roots for $\mathfrak{sl}(3)$ are $\pm\alpha_1, \pm\alpha_2, \pm(\alpha_1 + \alpha_2)$.

We can write all the roots in terms of α_1, α_2 with integral coefficients, all nonnegative or all nonpositive. For example, $\alpha_3 = \alpha_1 + \alpha_2$ and $-\alpha_3 = -\alpha_1 - \alpha_2$.

If we regard the α_i as $(\alpha_i(h_1), \alpha_i(h_2), \alpha_i(h_3))$, elements of the plane $x + y + z = 0$, we can compute the angle between the various roots. For example, if θ is the angle between α_1 and α_2 , we have

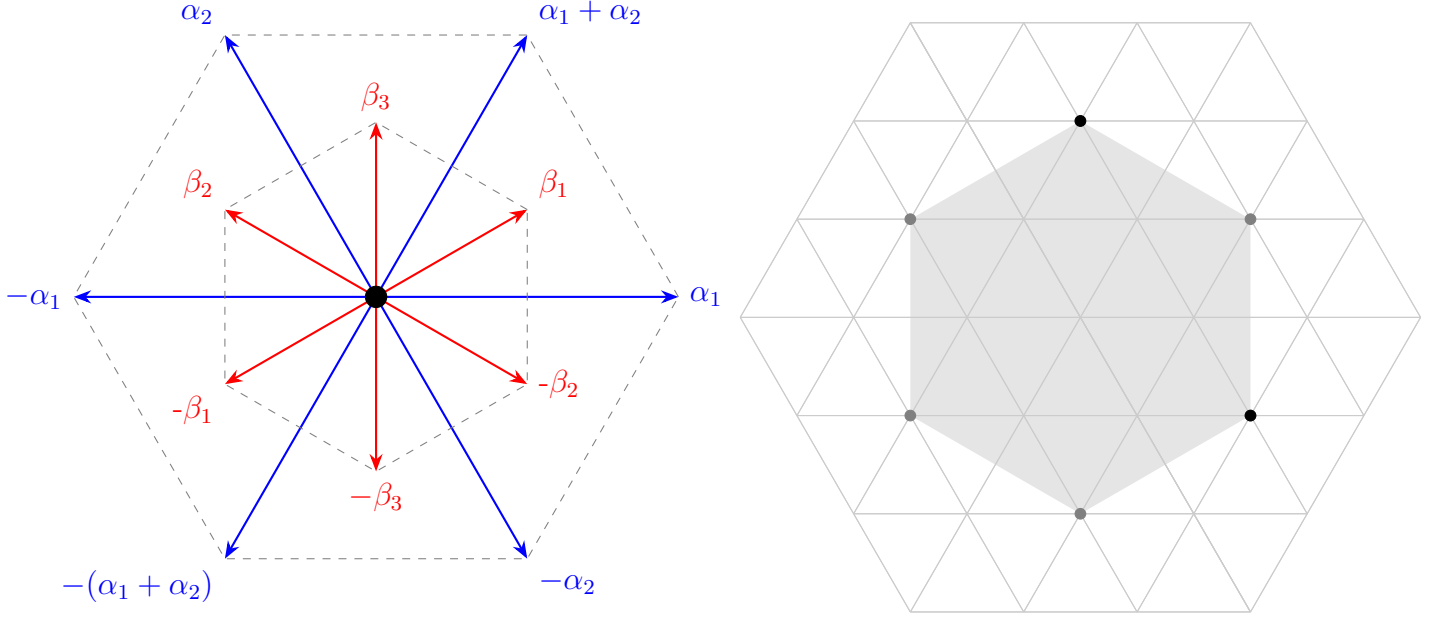
$$\cos \theta = \frac{\alpha_1 \cdot \alpha_2}{\|\alpha_1\| \|\alpha_2\|} = \frac{-2 + -2 + 1}{\sqrt{6}\sqrt{6}} = -\frac{1}{2}.$$

Thus $\theta = 120^\circ$. Likewise, the angles between α_1 and $-\alpha_3$ and α_2 and $-\alpha_3$ are both 120° . The angles between α_1 and α_3 and α_2 and α_3 are both 60° . The six roots for L are the vertices of a regular hexagon. See below for a picture of the roots.

We also can write $\alpha_1 = \beta_1 - \beta_2$, $\alpha_2 = \beta_2 - \beta_3$, $\alpha_3 = \beta_1 - \beta_3$. Regarding the α_i as vectors in the plane $x + y + z = 0$, with the β_j as coordinate axes, the angle between α_1 and α_2 is 120° , and the angle between α_i and α_3 is 60° for $i = 1, 2$, just as we computed above.

We also have $\beta_1 = \frac{1}{3}(\alpha_1 + \alpha_3) = \frac{2}{3}\alpha_1 + \frac{1}{3}\alpha_2$, $\beta_2 = -\frac{1}{3}\alpha_1 + \frac{1}{3}\alpha_2$, $\beta_3 = -\frac{1}{3}(\alpha_2 + \alpha_3) = -\frac{1}{3}\alpha_1 - \frac{2}{3}\alpha_2$.

Below are two picture of the root system, both drawn with `tikzpicture`; they are rotated differently. The first shows the α 's in blue and the β 's in red. The second is drawn with the `rank-2-roots` and `rootSystem` package and shows the roots α in a lattice. You can see pictures for other low-dimensional Lie algebras in Section 15 (Root System Pictures, p. 206) of Chapter 8 (Miscellaneous Calculations and Observations, p. 173).



We give one more example; we leave the details to the reader. Let V_n be the space of homogeneous polynomials in x, y, z of degree n (plus 0). We can define a representation of L on V_n by letting e_i, h_i, f_i act by E_i, H_i, F_i as below.

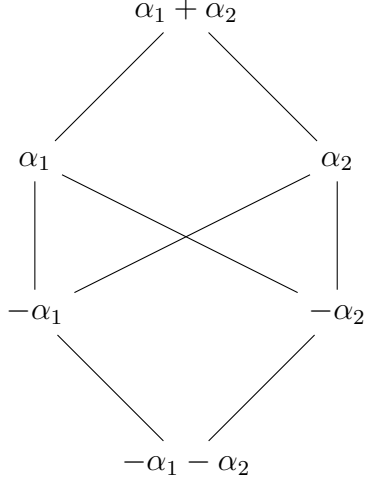
$$\begin{aligned} E_1 &= x \frac{\partial}{\partial y} & H_1 &= x \frac{\partial}{\partial x} - y \frac{\partial}{\partial y} & F_1 &= y \frac{\partial}{\partial x} \\ E_2 &= y \frac{\partial}{\partial z} & H_2 &= y \frac{\partial}{\partial y} - z \frac{\partial}{\partial z} & F_2 &= z \frac{\partial}{\partial y} \\ E_3 &= x \frac{\partial}{\partial z} & H_3 &= x \frac{\partial}{\partial x} - z \frac{\partial}{\partial z} & F_3 &= z \frac{\partial}{\partial x} \end{aligned}$$

Then $x^a y^b z^c$, with $a + b + c = n$, is a weight vector with weight $(a - b, b - c, a - c) = a\beta_1 + b\beta_2 + c\beta_3$, and these are all of the weight vectors (up to scalar multiples).

4.2. Highest weight vectors for $\mathfrak{sl}(3)$. Let us define a partial order on the set of weights of the representation V by the rule $\lambda \leq \mu$ if $\mu = \lambda + a\alpha_1 + b\alpha_2 + c\alpha_3$ for nonnegative integers a, b, c . Since $\alpha_3 = \alpha_1 + \alpha_2$, this is equivalent to $\lambda \leq \mu$ if $\mu = \lambda + a\alpha_1 + b\alpha_2$ for nonnegative integers a, b . This is easily seen to be a partial order, but it is not a total order for most representations.

For example, for the natural representation, we have $\beta_1 = \beta_2 + \alpha_1 = \beta_3 + \alpha_3$ and $\beta_2 = \beta_3 + \alpha_2$. Thus the weights are ordered $\beta_1 > \beta_2 > \beta_3$.

For the adjoint representation, α_3 is the sum of α_1 and α_2 , so it is greater than either of these. In addition, $\alpha_i = -\alpha_j + \alpha_i + \alpha_j$, so each α_i is greater than any $-\alpha_j$. However, neither α_1 nor α_2 is greater than the other. Below is a Hasse diagram of the partial order of the α 's.



For the representation V_n of homogeneous polynomials in x, y, z of degree n , we noted the weights are $a\beta_1 + b\beta_2 + c\beta_3$ with $a + b + c = n$. Using the formulas for the β_i given above, we see $a\beta_1 + b\beta_2 + c\beta_3 = a\alpha_1 - c\alpha_2 + \frac{n}{3}(-\alpha_1 + \alpha_2)$. In particular, if we subtract this weight from weight $n\beta_1$ (with weight vector x^n), we get $(n - a)\alpha_1 + c\alpha_2 = (b + c)\alpha_1 + c\alpha_2$. Thus $n\beta_1$ is greater than or equal to any other weight.

The weight vector in the natural representation with the largest weight, β_1 , is $\mathbf{e}_1$, which is annihilated by all of the e_i . The weight vector in the adjoint representation with largest weight, α_3 , is e_3 , which is annihilated by all of the e_i . The weight vector in the representation V_n with largest weight, $n\beta_1$, is x^n , which is annihilated by all of the e_i . This is not a coincidence as the following discussion shows.

Suppose v is a weight vector for the representation V with weight λ . Then $h_i.e_jv = e_j.h_iv + [h_i, e_j]v = \lambda(h_i)e_jv + \alpha_j(h_i)e_jv = (\lambda + \alpha_j)(h_i)e_jv$. This shows that $e_j.v$ is a weight vector with weight $\lambda + \alpha_j$ or else $e_j.v = 0$. In particular, either $e_j.v$ has greater weight than the weight of v or $e_j.v = 0$. Suppose we pick a weight vector v whose weight λ is maximal with respect to our order. Then $e_j.v = 0$ for each j by the this last comment.

A similar calculation shows that $f_j.v$ is either 0 or a weight vector with weight $\lambda - \alpha_j$. We can iterate this: if $1 \leq j_1, \dots, j_k \leq 3$, then $f_{j_1}f_{j_2} \dots f_{j_k}v$ is either 0 or a weight vector with weight $\lambda - \sum_{j=1}^k \alpha_{j_k} < \lambda$.

We define a **highest weight vector** to be a weight vector v with the property that $e_i.v = 0$ for all i . ([Humphreys, §20.2, p. 108] uses the term **maximal vector**.) The discussion above shows that if v is a weight vector whose weight λ is maximal with respect to our ordering of weights, then v is a highest weight vector. We now discuss a converse to this statement.

Suppose v is a highest weight vector of weight λ and let W_m denote the span of all $f_{j_1}f_{j_2} \dots f_{j_k}v$ with $k \leq m$. We will show that W_m is closed under the action of each h_i

and e_i . It follows that $W = \sum_m W_m$ is a submodule of V , and our remarks above show W has highest weight vector v and that all the weights of W are less than or equal to λ .

It is clear that each W_m is closed under the action of the h_i , since each $f_{j_1}f_{j_2}\dots f_{j_k}v$ is a weight vector or 0. By definition of highest weight vector, $W_0 = \mathbb{F}v$ is annihilated by each e_i . Now $e_i f_j v = f_j e_i v + [e_i, f_j]v = [e_i, f_j]v$. The multiplication table for L shows that each $[e_i, f_j]$ is an h , a $\pm f$, or a $\pm e$. $h_\ell v = \lambda(h_\ell)v \in W_0$ and $f_\ell v \in W_1$, while $e_\ell v = 0$. Thus W_1 is closed under the action of the e_i .

Suppose by induction that W_m is closed under the action of the e_i . Every element of W_{m+1} is a linear combination of elements of the form $w = f_j w'$ with $w' \in W_m$. As above, $e_i w = f_j e_i w' + [e_i, f_j]w'$. By induction, $e_i w' \in W_m$, so $f_j e_i w' \in W_{m+1}$. If $[e_i, f_j]$ is an h , then $[e_i, f_j]w' \in \mathbb{F}w' \subseteq W_m$. If $[e_i, f_j]$ is a $\pm f$, then $[e_i, f_j]w' \in W_{m+1}$ by definition. If $[e_i, f_j]$ is a $\pm e$, then $[e_i, f_j]w' \in W_m$ by our inductive hypothesis. Thus in all cases $e_i w \in W_{m+1}$; this completes the inductive step.

It follows that if V is irreducible and v is a highest weight vector in V of weight λ , then λ is the greatest weight for V . We summarize our results below and add some more information.

Proposition 3.4.1 (Irreducible representations of $\mathfrak{sl}(3)$).

Let $\mathbb{F}$ be a field of characteristic 0 and let $L = \mathfrak{sl}(3, \mathbb{F})$. Let $\beta_1, \beta_2, \beta_3$ be defined as above.

- (1) If V is a finite-dimensional irreducible L -module, then V contains a unique (up to scalar multiple) highest weight vector v . Moreover, V is a direct sum of H -weight spaces, and the weights that occur are less than or equal to the weight of v .
- (2) For every pair of nonnegative integers a, b , there is an irreducible L -module $\Gamma_{a,b}$, of dimension $\frac{1}{2}(a+1)(b+1)(a+b+2)$, with highest weight $a\beta_1 - b\beta_3 = (a+b)\beta_1 + b\beta_2$. Moreover, $\Gamma_{a,b} \cong \Gamma_{a',b'}$ if and only if $a = a'$ and $b = b'$.
- (3) Every finite-dimensional irreducible L -module is isomorphic to $\Gamma_{a,b}$ for some a, b .

REFERENCES AND REMARKS. See [Fulton-Harris, Lecture 12, Proposition 12.11, p. 169 and Lecture 13, Theorem 13.1, p. 176].

One can also say something about which weights occur in an irreducible $\mathfrak{sl}(3)$ -representation (analogous to the weights for an irreducible representation of $\mathfrak{sl}(2)$ forming an arithmetic progression): see [Fulton-Harris, Lecture 12, Propositions 12.15 & 12.18, pp. 172-174].

The natural representation has highest weight β_1 , so it is isomorphic to $\Gamma_{1,0}$. The adjoint representation has highest weight $\alpha_3 = \beta_1 - \beta_3$, so it is isomorphic to $\Gamma_{1,1}$. The representation V_n of homogeneous polynomials in x, y, z of degree n has highest weight is $n\beta_1$, so it is isomorphic to $\Gamma_{n,0}$.

PROOF. (1) We noted above that a weight vector v of maximal weight is a highest weight vector, so highest weight vectors exist in any finite-dimensional module. Since v generates V (V is irreducible), we showed the weights that occur are $\leq$ the weight of v . Moreover, if v, v' are linearly independent highest weight vectors, neither is the module generated by the other; this contradicts irreducibility.

(2), (3) We leave these to the reader to look up, perhaps in [Fulton-Harris] or another source. There are several ways to construct the irreducible modules $\Gamma_{a,b}$. It is a fact that the dual $\Gamma_{a,b}^*$ is isomorphic to $\Gamma_{b,a}$. Thus $\Gamma_{a,0} \otimes \Gamma_{b,0}^*$ has a highest weight vector of weight $a\beta_1 - b\beta_3$, which generates a submodule isomorphic to $\Gamma_{a,b}$. **Check!** ■

4.3. Dimensions of irreducible representations. We have seen that there is exactly one irreducible representation of $\mathfrak{sl}(2)$ in each (finite) dimension. This is no longer true for $\mathfrak{sl}(n)$ when $n \geq 3$. It follows from Proposition 3.4.1 (Irreducible representations of $\mathfrak{sl}(3)$, p. 113) that for every pair of positive integers x, y , there is an irreducible $\mathfrak{sl}(3)$ -module of dimension $\frac{1}{2}xy(x+y)$ and two such modules are isomorphic if and only if the ordered pair (x, y) is the same. (The module is $\Gamma_{x-1, y-1}$.) Moreover, this formula gives the dimensions of all irreducible $\mathfrak{sl}(3)$ -modules.

For example $\frac{1}{2}xy(x+y) = 120$ has 4 solutions: $(x, y) = (15, 1), (1, 15), (10, 2), (2, 10)$. Thus there are four irreducible $\mathfrak{sl}(3)$ -modules of dimension 120, up to isomorphism. (In the notation of the previous subsection, they are $\Gamma_{14,0}, \Gamma_{0,14}, \Gamma_{9,1}, \Gamma_{1,9}$.)

Here is a list of the dimensions of irreducible representations of $\mathfrak{sl}(3)$ up to 100:

1, 3, 6, 8, 10, 15, 21, 24, 27, 28, 35, 36, 42, 45, 48, 55, 60, 63, 64, 66, 78, 80, 81, 90, 91, 99

This list is the start of sequence A121741 from the OEIS (On-Line Encyclopedia of Integer Sequences).

Each dimension on the list that is not a perfect cube is the dimension of exactly two non-isomorphic irreducible representations; each cube is the dimension of exactly one irreducible (up to isomorphism). There are exactly four non-isomorphic irreducible representations in dimension 105 and in dimension 120.

Here is a list of the dimensions of irreducible representations of $\mathfrak{sl}(4)$ up to 100:

1, 4, 6, 10, 15, 20, 35, 36, 45, 50, 56, 60, 64, 70, 84

This list is the start of sequence A121214 from the OEIS.

5. Weyl's Theorem on Complete Reducibility

We aim to show that every (finite-dimensional) representation of a semisimple Lie algebra is completely reducible. To do this, we introduced a special element or operator.

5.1. The Casimir element and operator. We now define a very special element or operator that commutes with the action of our Lie algebra L .

Suppose $\langle, \rangle$ is an invariant nondegenerate symmetric bilinear form on the Lie algebra L . (This forces L to be semisimple.) Let $x_1, \dots, x_n$ be a basis of L and let $y_1, \dots, y_n$ be a dual basis relative to $\langle, \rangle$, i.e., $\langle x_i, y_j \rangle = \delta_{ij}$. Set $c = \sum_{i=1}^n x_i y_i$; we call this the **Casimir element** or **Casimir invariant** of L . If $L \subseteq \mathfrak{gl}(V)$ is linear, then we can regard c as an element of $\mathfrak{gl}(V)$ (but generally not of L). In the abstract, we can regard c as an element of the universal enveloping algebra $U(L)$ of L . Recall that $U(L)$ is an associative algebra generated by L , with multiplication determined by $xy - yx = [x, y]$ for $x, y \in L$, and recall that elements of $U(L)$ look like (commutative) polynomials in a basis for L but with a twisted multiplication.

As an example, let's compute the Casimir elements for $\mathfrak{sl}(2)$ and $\mathfrak{sl}(3)$. First take $L = \mathfrak{sl}(2)$ with standard basis $e = e_{12}, h = e_{11} - e_{22}, f = e_{21}$ and let our form be the concrete trace form τ . The dual basis is $f, \frac{1}{2}h, e$, so the Casimir element is $c = ef + fe + \frac{1}{2}h^2$. (If we use Killing form, we get $c = \frac{1}{4}(ef + fe) + \frac{1}{8}h^2$, which is $\frac{1}{4}$ times the result using the concrete trace form.) We can regard c as an element of $U(L)$, in which case it can also be written as $2ef + \frac{1}{2}h^2 - h$. If we regard it as an element of $\mathfrak{gl}(2)$, we have $c = e_{11} + e_{22} + \frac{1}{2}I = \frac{3}{2}I$. This is not surprising, since we are about to show c commutes with elements of L . In fact, we will confirm it below in Proposition 3.5.1 (Casimir properties, p. 117)(3).

Let us continue this example using the notation from Section 3 (Representations of $\mathfrak{sl}(2)$, p. 103). Let us compute $\frac{1}{2}H^2 + EF + FE$ (the operator version of c). We'll do this by operating on a polynomial $p = p(x, y)$; we use p_x, p_y , etc. for the partial derivatives of p . We have $H^2(p) = H(xp_x - yp_y) = xp_x + x^2p_{xx} - xyp_{xy} - xyp_{xy} + y^2p_{yy} + yp_y = x^2p_{xx} + y^2p_{yy} - 2xyp_{xy} + xp_x + yp_y$, $EF(p) = E(yp_x) = xp_x + xyp_{xy}$, $FE(p) = F(xp_y) = yp_y + xyp_{xy}$. Thus

$$\begin{aligned} \left(\frac{1}{2}H^2 + EF + FE\right)(p) &= \frac{1}{2}x^2p_{xx} + \frac{1}{2}y^2p_{yy} - xyp_{xy} + \frac{1}{2}xp_x + \frac{1}{2}yp_y + xp_x + xyp_{xy} + yp_y + xyp_{xy} \\ &= \frac{1}{2}(x^2p_{xx} + y^2p_{yy}) + xyp_{xy} + \frac{3}{2}(xp_x + yp_y) \end{aligned}$$

It follows that $c = \frac{1}{2}(x^2 \frac{\partial^2}{\partial x^2} + y^2 \frac{\partial^2}{\partial y^2}) + xy \frac{\partial^2}{\partial x \partial y} + \frac{3}{2}(x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y})$. Thus

$$\begin{aligned} c(x^i y^j) &= \left(\frac{1}{2}(i(i-1) + j(j-1)) + ij + \frac{3}{2}(i+j)\right) x^i y^j \\ &= \frac{1}{2}(i^2 - i + j^2 - j + 2ij + 3i + 3j) x^i y^j = \frac{1}{2}((i+j)^2 + 2(i+j)) x^i y^j. \end{aligned}$$

In particular, c acts on the irreducible module V_n by the scalar $\frac{n^2 + 2n}{2} = \frac{n(n+2)}{2}$. The dimension of V_n is $n+1$, so $\text{tr } c_{V_n} = \frac{n(n+1)(n+2)}{2}$. Said differently, c acts on the irreducible $\mathfrak{sl}(2)$ -module of dimension m by the scalar $\frac{1}{2}(m^2 - 1)$, with trace $\frac{1}{2}m(m-1)m(m+1)$.

Next let $L = \mathfrak{sl}(3)$ with standard basis $e_1, e_2, e_3, h_1, h_2, f_1, f_2, f_3$ (notation as in Section 4 (Representations of $\mathfrak{sl}(3)$, p. 107)) and again use the concrete trace form τ . The dual basis is $f_1, f_2, f_3, \frac{2}{3}h_1 + \frac{1}{3}h_2, \frac{1}{3}h_1 + \frac{2}{3}h_2, e_1, e_2, e_3$. Since $h_1(\frac{2}{3}h_1 + \frac{1}{3}h_2) + h_2(\frac{1}{3}h_1 + \frac{2}{3}h_2) = \frac{2}{3}(h_1^2 + h_2^2 + h_1h_2)$, we have $c = e_1f_1 + e_2f_2 + e_3f_3 + f_1e_1 + f_2e_2 + f_3e_3 + \frac{2}{3}(h_1^2 + h_2^2 + h_1h_2) = \frac{2}{3}(h_1^2 + h_2^2 + h_1h_2 - 3h_1 - 3h_2) + 2e_1f_1 + 2e_2f_2 + 2e_3f_3$ in $U(L)$. In $\mathfrak{gl}(3)$, the products e_jf_j and f_je_j consist of exactly two copies of e_{ii} for each $i = 1, 2, 3$, and so the sum $e_1f_1 + e_2f_2 + e_3f_3 + f_1e_1 + f_2e_2 + f_3e_3$ is $2I$. We can also compute that $h_1(\frac{2}{3}h_1 + \frac{1}{3}h_2) + h_2(\frac{1}{3}h_1 + \frac{2}{3}h_2) = \frac{2}{3}(h_1^2 + h_2^2 + h_1h_2) = \frac{2}{3}I$. Thus $c = \frac{8}{3}I$ in $\mathfrak{gl}(3)$. Compare this to Proposition 3.5.1 (Casimir properties, p. 117)(3) below.

Let us first show that c is independent of the choice of dual bases. Note that the Casimir element may still depend on $\langle, \rangle$. Using matrix calculations (treating $\mathbf{x} = (x_1, \dots, x_n)$, $\mathbf{y} = (y_1, \dots, y_n)$ as column vectors), note that $\mathbf{x}, \mathbf{y}$ being dual bases means $\mathbf{xy}^T = I$, and we have $c = \mathbf{x}^T \mathbf{y}$. A change of basis from $\mathbf{x}$ to $\mathbf{x}'$ is given by a matrix P with $P\mathbf{x} = \mathbf{x}'$. Likewise, when we change bases from $\mathbf{y}$ to $\mathbf{y}'$, there is a matrix Q with $Q\mathbf{y} = \mathbf{y}'$. For $\mathbf{x}', \mathbf{y}'$ to be dual bases means $I = \mathbf{x}'\mathbf{y}'^T = P\mathbf{xy}^TQ^T = PQ^T$. This implies $P^TQ = I$. Now $c' = \mathbf{x}'^T \mathbf{y}' = \mathbf{x}^T P^T Q \mathbf{y} = \mathbf{x}^T \mathbf{y} = c$.

We wish to show c commutes with elements of L , i.e., $[x, c] = 0$ for $x \in L$. The following calculation can be thought of either as taking place in $\mathfrak{gl}(V)$ or in $U(L)$. For $x \in L$, write $[x, x_i] = \sum_j a_{ij}x_j$ and $[x, y_i] = \sum_j b_{ij}y_j$ for scalars a_{ij}, b_{ij} . Then $a_{ij} = \sum_k a_{ik}\langle x_k, y_j \rangle = \langle \sum_k a_{ik}x_k, y_j \rangle = \langle [x, x_i], y_j \rangle$. Likewise $b_{ij} = \langle [x, y_i], x_j \rangle = \langle x, [y_i, x_j] \rangle = -\langle x, [x_j, y_i] \rangle = -\langle [x, x_j], y_i \rangle = -a_{ji}$. Thus

$$[x, c] = \sum_i [x, x_i]y_i + \sum_i x_i[x, y_i] = \sum_{i,j} a_{ij}x_jy_i + \sum_{i,j} x_i(-a_{ji})y_j = \sum_{r,s} a_{rs}x_sy_r - \sum_{r,s} a_{rs}x_sy_r = 0.$$

When $\phi : L \rightarrow \mathfrak{gl}(V)$ is a representation of L in a vector space V , we modify the above construction to get the **Casimir operator** $c_\phi = \sum_{i=1}^n \phi(x_i)\phi(y_i)$. The above computation for c carries over to this setting and proves that c_ϕ commutes with the action of each $x \in L$, that is, $c_\phi : V \rightarrow V$ is an L -module homomorphism.

REMARK. Some sources, such as [Humphreys], call the “abstract” Casimir element c we have defined the **universal Casimir element**. The Casimir element we have defined is more or less the unique quadratic element of the center of $U(L)$ for L semisimple. There are higher degree Casimir elements that are elements of the center of $U(L)$. When

these higher order elements are in play, we call our c the **quadratic Casimir element**. Casimir introduced the (quadratic) Casimir operator in 1931 and Casimir and van der Waerden used it to prove complete reducibility of representations as we will below in 1935. In the 1950s, Chevalley and Harish-Chandra described the center of the enveloping algebra $U(L)$ for L semisimple (or even reductive).

We sum up our discussion and add one piece below.

Proposition 3.5.1 (Casimir properties).

Let L be a finite-dimensional semisimple Lie algebra over a field of characteristic 0 and let $\langle, \rangle$ be an invariant nondegenerate symmetric bilinear form on L . Let $x_1, \dots, x_n$ and $y_1, \dots, y_n$ be dual bases of L relative to $\langle, \rangle$ and let $\phi : L \rightarrow \mathfrak{gl}(V)$ be a representation of L .

- (1) $c = \sum_{i=1}^n x_i y_i$ is a central element of the universal enveloping algebra $U(L)$.
- (2) $c_\phi = \sum_{i=1}^n \phi(x_i) \phi(y_i) \in \mathfrak{gl}(V)$ is an L -module endomorphism of V .
- (3) Suppose ϕ is faithful and define $\langle x, y \rangle = \text{tr}(\phi(x) \phi(y))$ for $x, y \in L$. Then $\langle, \rangle$ is an invariant nondegenerate symmetric bilinear form on L with $\text{tr } c_\phi = \dim L$. If V is also irreducible and $\mathbb{F}$ is algebraically closed, then $c_\phi(v) = \frac{\dim L}{\dim V} v$ for all $v \in V$.

REFERENCES AND REMARKS. See [Humphreys, §6.2, pp. 27-28; Jacobson, §III.7, proof of Lemma 3, pp. 77-78; Exercises V.8-9, p. 195; Bourbaki 1-3, §I.3.7, pp. 35-36; Erdmann-Wildon, Exercise 8.6, pp. 75-76; Wikipedia].

PROOF. (1) and (2) are proved in the discussion prior to the proposition.

(3) Since ϕ is faithful, it embeds L in $\mathfrak{gl}(V)$, and the form $\langle, \rangle$ defined is just the concrete trace form. Now $\text{tr } c_\phi = \sum_i \text{tr}(\phi(x_i) \phi(y_i)) = \sum_i \langle x_i, y_i \rangle = \sum_i 1 = n = \dim V$. If V is irreducible, Schur's Lemma tells us $c_\phi(v) = \lambda v$ for all v , for some scalar λ . Thus $\dim L = \text{tr } c_\phi = (\dim V) \lambda$. ■

5.2. Weyl's theorem. We single out one piece of the proof of Weyl's theorem.

Proposition 3.5.2 (Distinct central characters split sequences).

Let L be a Lie algebra and let $0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0$ be a short exact sequence of L -modules. Let $\theta : B \rightarrow B$ be an L -module homomorphism with $\theta(A) \subseteq A$. If either of the following conditions holds, then this sequence splits.

- (1) There are distinct scalars $\alpha, \gamma \in \mathbb{F}$ with $\theta(a) = \alpha a$ for all $a \in A$ and $\bar{\theta}(c) = \gamma c$ for all $c \in C$.
- (2) $\theta(B) \subseteq A$ and $\theta|_A$ is an automorphism of A .

PROOF. We have to define a map $\eta : B \rightarrow A$ such that $\eta|_A = \text{id}_A$.

(1) Define $\eta(b) = \frac{1}{\alpha - \gamma}(\theta(b) - \gamma b)$. Since $\bar{\theta}(\bar{b}) = \gamma \bar{b}$, we see $\eta(b) \in A$. Moreover, $\eta(a) = \frac{1}{\alpha - \gamma}(\alpha a - \gamma a) = a$.

(2) Define $\eta(b) = (\theta|_A)^{-1}(\theta(b))$. Clearly $\eta(a) = a$. Since $\theta(b) \in A$, η makes sense and maps B to A . ■

Proposition 3.5.3 (Weyl's theorem on complete reducibility).

If L is a finite-dimensional semisimple Lie algebra over a field of characteristic 0, then every finite-dimensional representation of L is completely reducible.

REFERENCES AND REMARKS. See [Humphreys, §6.3, Theorem, p. 28; Jacobson, §III.7, Theorem 8, p. 79; Bourbaki 1-3, §I.6.2, Theorem 2, p. 52; Erdmann-Wildon, Theorem 8.7, p. 74; Appendix B, Theorem 17.4, p. 212; Wikipedia].

PROOF. We wish to show every submodule W of an L -module V is a direct summand of V . The proof proceeds in three steps. We first prove that if W is irreducible and if $\dim V/W = 1$, then W is a direct summand of V . We then prove that if W is an arbitrary submodule of V and if $\dim V/W = 1$, then W is a direct summand of V . Finally, we prove any W is a direct summand of V .

Thus we first assume W is irreducible and $\dim V/W = 1$. If K is the kernel of the representation of V , we can also pass to L/K and assume that V is faithful. The action of $[L, L]$ on any 1-dimensional module is trivial (why?). Since $L = [L, L]$, the action of L on V/W is thus 0.

Consider the action of the Casimir element c , as discussed in Subsection 5.1 (The Casimir element and operator, p. 115). We've just said that c acts like 0 on V/W , hence $cV \subseteq W$. Since W is irreducible, Schur's Lemma tells us c acts as either zero or an isomorphism on W . But $\text{tr } c = \dim L \neq 0$, so c cannot act like 0. Thus c acts like an isomorphism. By Proposition 3.5.2 (Distinct central characters split sequences, p. 117)(2), this implies the sequence $0 \rightarrow W \rightarrow V \rightarrow V/W \rightarrow 0$ splits and so W is a summand of V . (When $\mathbb{F}$ is algebraically closed, we can also apply Proposition 3.5.2 (Distinct central characters split sequences, p. 117)(1), since c will act on W as the scalar $\dim L / \dim W \neq 0$ and on V/W as 0.)

Next we assume that a submodule W' of any L -module V' with $\dim V'/W' = 1$ is a summand whenever $\dim V' < \dim V$. We have to show W is a summand of V , and we may plainly assume $W \neq 0$. If W is irreducible, then we know W is a summand of V by the first step in the proof. Thus we may suppose W has a nonzero proper submodule W' . Then W/W' is a submodule of V/W' and $\dim V/W' < \dim V$, so W/W' is a summand of V/W' by induction. Let $V/W' = W/W' \oplus U/W'$ for some submodule U containing W' . This means that $V = W + U$ and $W \cap U = W'$. Now $\dim U < \dim V$,

so by induction, $U = W'' \oplus W'$ for some submodule W'' of U . We claim $V = W \oplus W''$. Since $W' \subseteq W$, we have $W + W'' = W + W' + W'' = W + U = V$. We also have $W \cap W'' \subseteq W \cap U \cap W'' \subseteq W' \cap W'' = 0$. This proves the claim.

For the third and final step, let W be any nonzero submodule of V . We have seen before that the space of linear maps $\text{Hom}_{\mathbb{F}}(V, W)$ is an L -module under the action $(x.\phi)(v) = x.\phi(v) - \phi(x.v)$. Let $\mathcal{V}$ be the subspace of $\text{Hom}(V, W)$ consisting of those ϕ such that $\phi|_W(w) = \alpha w$ for some scalar $\alpha \in \mathbb{F}$, all $w \in W$. Let $\mathcal{W}$ be the subspace of all $\phi \in \text{Hom}(V, W)$ such that $\phi|_W = 0$, so $\mathcal{W} \subseteq \mathcal{V}$. Suppose $\phi \in \mathcal{V}$, say $\phi(u) = \alpha u$ for all $u \in W$, and $x \in L, w \in W$. Then $(x.\phi)(w) = x.\phi(w) - \phi(x.w) = x.(\alpha w) - \alpha x.w = 0$, so $x.\phi \in \mathcal{W}$. Thus $x.\mathcal{V} \subseteq \mathcal{W}$ for all $x \in L$. This shows $\mathcal{V}, \mathcal{W}$ are submodules of $\text{Hom}(V, W)$.

Define $\Phi : \mathcal{V} \rightarrow \mathbb{F}$ (where $\mathbb{F}$ is the trivial L -module) by setting $\Phi(\phi)$ equal to the scalar α above. Since $W \neq 0$, α is unique and thus Φ is easily seen to be a well-defined linear map with kernel $\mathcal{W}$. In fact, Φ is a module homomorphism, since $\Phi(x.\phi) = 0$ and $x.\Phi(\phi) = 0$. Since linear maps from V to W can be extended from any linear map from W to W , we see that Φ is onto. Thus $\mathcal{V}/\mathcal{W}$ is one-dimensional, and thus $\mathcal{V} = \mathcal{W} \oplus \mathcal{U}$ for some L -submodule $\mathcal{U}$ of $\mathcal{V}$ by the second step. Note that $\mathcal{U}$ is the trivial L -module. Let ϕ be a nonzero element of $\mathcal{U}$ with scalar $\alpha \neq 0$. If we replace ϕ by $\alpha^{-1}\phi$, we can assume $\alpha = 1$, i.e., $\phi|_W = \text{id}_W$. Since $\mathcal{U}$ is the trivial L -module, we have $0 = (x.\phi)(v) = x.\phi(v) - \phi(x.v)$ for all $x \in L, v \in V$. Thus $\phi(x.v) = x.\phi(v)$, and so ϕ is an L -module homomorphism and a projection with image W . This implies W is a direct summand of V . ■

5.3. Abstract Jordan-Chevalley decomposition. If L is a Lie algebra and $x \in L$, then $\text{ad } x \in \mathfrak{gl}(L)$ is a derivation. If $\text{ad } x = \phi + \eta$ is the Jordan-Chevalley decomposition of $\text{ad } x$ in $\mathfrak{gl}(L)$, then Proposition 2.2.3 (Jordan-Chevalley decomposition of derivations, p. 51) tells us ϕ, η are derivations on L . But are they inner? If either one is inner, so is the other, say $\phi = \text{ad } y', \eta = \text{ad } z'$, then $\text{ad } x = \text{ad}(y' + z')$, so $x - (y' + z') \in Z(L)$. Adding that element of $Z(L)$ to either y' or z' , we can get y, z with $y + z = x$ and $\text{ad } x = \text{ad } y + \text{ad } z$ the Jordan-Chevalley decomposition of $\text{ad } x$ in $\mathfrak{gl}(L)$, giving us a kind of ad-Jordan-Chevalley decomposition. This is sometimes called the **abstract Jordan-Chevalley decomposition**. (We don't know that y is semisimple or z is nilpotent.) If L is semisimple, any derivation on L is inner by Proposition 2.7.13 (All derivations are inner, p. 78), and so the abstract Jordan-Chevalley decomposition exists.

Let us look at this from a different angle. If $L \subseteq \mathfrak{gl}(V)$ is a linear Lie algebra, then any $x \in L$ has a Jordan-Chevalley decomposition $x = s + n$ for some $s, n \in \mathfrak{gl}(V)$. Moreover, Proposition 2.4.2 (Semisimple/nilpotent and ad, p. 61) tells us $\text{ad } s$ is semisimple, $\text{ad } n$ is nilpotent, and $[\text{ad } s, \text{ad } n] = \text{ad}[s, n] = 0$. Thus if $s, n \in L$, we have $\text{ad } x = \text{ad } s + \text{ad } n$ the Jordan-Chevalley decomposition of $\text{ad } x$. But are $s, n \in L$? If L is semisimple, then as in

the previous paragraph, we have $y, z \in L, s, n \in \mathfrak{gl}(V)$ with $\text{ad } x = \text{ad } y + \text{ad } z = \text{ad } s + \text{ad } n$. By uniqueness of the Jordan-Chevalley decomposition, we have $\text{ad } y = \text{ad } s$ and $\text{ad } z = \text{ad } n$. Can we say $y = s$ or $z = n$?

We summarize our results.

Proposition 3.5.4 (Abstract Jordan-Chevalley decomposition).

Let L be a finite-dimensional semisimple Lie algebra over a field of characteristic 0 and let $x \in L$.

- (1) *There are unique $x_s, x_n \in L$ with $x = x_s + x_n$, $[x_s, x_n] = 0$, and $\text{ad } x_s, \text{ad } x_n$ the semisimple and nilpotent parts of the Jordan-Chevalley decomposition of $\text{ad } x$ in $\mathfrak{gl}(L)$. Moreover, if $[x, y] = 0$ for $y \in L$, then $[x_s, y] = [x_n, y] = 0$.*
- (2) *If $L \subseteq \mathfrak{gl}(V)$ is linear and $x = s + n$ is the Jordan-Chevalley decomposition of x in $\mathfrak{gl}(V)$, then $s, n \in L$ and in the notation of (1), $x_s = s, x_n = n$.*

REFERENCES AND REMARKS. See [Humphreys, §5.4, p. 24; Jacobson, §III.11, Theorem 17 & its proof, pp. 100-101; Bourbaki 1-3, §.3, Proposition 3, p. 54; Erdmann-Wildon, Theorem 9.15, p. 86].

[Jacobson, §III.11, Definition 3, p. 98] calls a linear Lie algebra L **almost algebraic** if whenever $x = s + n$ is the Jordan-Chevalley decomposition of x in $\mathfrak{gl}(V)$, we have $s, n \in L$. Our result shows semisimple linear Lie algebras are almost algebraic. [Bourbaki 7-9, §VII.5.1, Definition 1, p. 34] uses the term **decomposable** for such a linear Lie algebra.

PROOF. (1) The discussion prior to the proposition gives us the existence of x_s, x_n with $x = x_s + x_n$ and $\text{ad } x = \text{ad } x_s + \text{ad } x_n$ the Jordan-Chevalley decomposition of $\text{ad } x$ in $\mathfrak{gl}(L)$. In particular, $\text{ad}[x_s, x_n] = [\text{ad } x_s, \text{ad } x_n] = 0$, so $[x_s, x_n] \in Z(L) = 0$. If $[x, y] = 0$, then $[\text{ad } x, \text{ad } y] = 0$. By the properties of the Jordan-Chevalley decomposition, $0 = [\text{ad } x_s, \text{ad } y] = \text{ad}[x_s, y]$. As above, this implies $[x_s, y] = 0$, and likewise $[x_n, y] = 0$. If $x = y + z$ is another decomposition with $\text{ad } x = \text{ad } y + \text{ad } z$ the Jordan-Chevalley decomposition of x , then uniqueness implies $\text{ad } y = \text{ad } x_s$ and $\text{ad } z = \text{ad } x_n$. Since $Z(L) = 0$, this implies $x_s = y$ and $x_n = z$.

(2) What needs to be proved is that if $x = s + n$ is the Jordan-Chevalley decomposition of x in $\mathfrak{gl}(V)$, then $s, n \in L$. Set $G = \mathfrak{gl}(V)$. In this proof, we will use two actions of L : one is the ordinary action of L on V as operators, i.e., $y.v = y(v)$. The other is the adjoint action of L on G , i.e., $y.z = \text{ad}_G y(z) = [y, z]$ for $z \in G$. To (we hope) avoid confusion, we will refer to modules and submodules of the second action as $\text{ad } L$ -modules.

We first notice that since $\text{ad } s, \text{ad } n$ are polynomials in $\text{ad } x$, and $\text{ad } x(L) = [x, L] \subseteq L$, we see that L is invariant under $\text{ad } x_s, \text{ad } x_n$. If we set $N = N_G(L) = \{y \in G \mid [y, L] \subseteq L\}$, then $s, n \in N$. Of course, N is a subalgebra of G and L is an ideal of N .

Let W be any L -submodule of V and define $B_W = \{y \in G \mid y(W) \subseteq W\}$ and $A_W = \{y \in B_W \mid \operatorname{tr}(y|_W) = 0\}$. If $y, z \in B_W$, then $[y, z](W) \subseteq y(z(W)) + z(y(W)) \subseteq W$ and $[y, z]|_W = [y|_W, z|_W]$, so $\operatorname{tr}[y, z]|_W = 0$. Thus $[B_W, B_W] \subseteq A_W$ and so A_W, B_W are both subalgebras of G . Since W is an L -submodule of V , we have $L \subseteq B_W$. Thus $L = [L, L] \subseteq [B_W, B_W] \subseteq A_W$ as well. In particular, then, if $A = N \cap \bigcap_W A_W$ (where the intersection is taken over all L -submodules W of V), we have $L \subseteq A$. In fact, A is a subalgebra of G contained in N , so L is an ideal of A .

Moreover, s, n are polynomials in x , so $s(W) \subseteq W, n(W) \subseteq W$ for all W . Furthermore, $n|_W$ is nilpotent, so $\operatorname{tr} n|_W = 0$ and we saw above that $\operatorname{tr} x|_W = 0$, so $\operatorname{tr} s|_W = 0$. This shows that $s, n \in A_W$ and so $s, n \in A$. Our goal now is to show $A = L$, which will give us $s, n \in L$.

Since L is an ideal of A , we have $[L, A] \subseteq L \subseteq A$. Thus A is an $\operatorname{ad} L$ -submodule of G , and L is an $\operatorname{ad} L$ -submodule of A . Since $\operatorname{ad} L \cong L$ is semisimple, Proposition 3.5.3 (Weyl's theorem on complete reducibility, p. 118) tells us A is a completely reducible $\operatorname{ad} L$ -module, and so $A = L \oplus K$ for some $\operatorname{ad} L$ -submodule K of A . Since $\operatorname{ad} L(A) = [L, A] \subseteq L$, this shows that $\operatorname{ad} L(K) = 0$.

Assume now that $\mathbb{F}$ is algebraically closed, W is an irreducible L -submodule of V , and $y \in K$. Then $[L, y] = 0$ implies that $y|_W : W \rightarrow W$ is an L -module homomorphism, since for $z \in L, w \in W$, we have $y|_W(z.w) = y(z(w)) = z(y(w)) + [y, z](w) = z.y|_W(w)$. Since W is irreducible, Proposition 3.1.7 (Schur's lemma, p. 92) implies there is a scalar $\lambda \in \mathbb{F}$ with $y|_W(w) = \lambda w$ for all $w \in W$. Thus $0 = \operatorname{tr} y|_W = (\dim W)\lambda$, and so $\lambda = 0$. It follows that $y(W) = 0$ for all irreducible L -submodules of V . By Proposition 3.5.3 (Weyl's theorem on complete reducibility, p. 118), V is a direct sum of irreducible submodules, and so $y(V) = 0$. This proves $y = 0$ and hence $K = 0$, as desired.

The non-algebraically closed case can be handled by extension of scalars — see [Bourbaki 1-3, §.3, Proposition 3, p. 54]. **Details?** ■

5.4. Reductive algebras. We give concrete and abstract versions of results about complete reducibility of either the space V or L , and this leads to a characterization of reductive Lie algebras.

Proposition 3.5.5 (Reductive algebras as direct sums).

Let L be a finite-dimensional Lie algebra over a field of characteristic 0. Then the following statements are equivalent.

- (1) L is reductive.
- (2) $L = [L, L] \oplus Z(L)$ and $[L, L]$ is semisimple.
- (3) L is completely reducible as an $\operatorname{ad} L$ -module.

REFERENCES AND REMARKS. See [Humphreys, §7.1, Exercise 5, pp.30-31; §19.1, Proposition, p.102; Jacobson, Chapter III, Exercise 19, p.104; Bourbaki 1-3, I.6.4, p.56; Wikipedia].

PROOF. (3) $\Rightarrow$ (2). Write $L = Z(L) \oplus I$ as $\text{ad } L$ -modules. This means I is an ideal of L and $I \cong L/Z(L)$. Suppose I contains an Abelian ideal J , which is automatically an ideal of L . By complete reducibility, $I = J \oplus K$ for some $\text{ad } L$ -submodule (i.e., ideal of L) K . This implies $[J, I] = 0$ and of course $[J, Z(L)] = 0$, so $J \subseteq Z(L)$. This forces $J = 0$ and so I is semisimple. Finally, $[L, L] = [Z(L), Z(L)] + [Z(L), I] + [I, I] = [I, I] = I$, since I is semisimple.

(2) $\Rightarrow$ (1). This is immediate.

(1) $\Rightarrow$ (3). Since $\text{ad } L \cong L/Z(L)$ is semisimple, Proposition 3.5.3 (Weyl's theorem on complete reducibility, p.118) implies L is completely reducible as an $\text{ad } L$ -module. ■

Here is a concrete result on reducibility.

Proposition 3.5.6 (Irreducible and completely reducible linear Lie algebras).

Let V be a finite-dimensional vector space over a field of characteristic 0 and let $L \subseteq \mathfrak{gl}(V)$ be a linear Lie algebra.

(1) V is a completely reducible L -module if and only if the following conditions hold:

- $L = Z(L) \oplus [L, L]$.
- $[L, L]$ is semisimple.
- The elements of $Z(L)$ are all semisimple operators

(2) If V is an irreducible L -module, then $\dim Z(L) \leq 1$.

(3) If $L \subseteq \mathfrak{sl}(V)$ and V is an irreducible L -module, then $Z(L) = 0$ and so L is semisimple.

REFERENCES AND REMARKS. See [Humphreys, §19.1, Proposition, p.102; Jacobson, §II.5, Theorem 11, p.47; §III.7, Theorem 10, p.81; Erdmann-Wildon, Chapter 12, Exercise 12.4, pp.139 & 245-246].

PROOF. Assume $\mathbb{F}$ is algebraically closed, let R be the solvable radical of L , and let W be any L -submodule of V . By Proposition 3.2.6 (Concrete Lie's theorem, p.100), there is a common eigenvector w for the elements of R , i.e., $xw = \lambda(x)w$ for some $\lambda \in R^*$. By Proposition 3.2.1 (Invariance lemma, p.97), the λ -eigenspace W_λ is an L -submodule of W . In particular, if W is irreducible, then $W = W_\lambda$.

(1) Since V is completely reducible, it is a direct sum of modules of the form W_λ . This means that the elements of R act diagonally on V and so R is Abelian. In addition, all elements of R are semisimple. Moreover, elements of $[L, R]$ act like 0 on every W_λ by Proposition 3.2.1 (Invariance lemma, p.97), and so $[L, R] = 0$. This means that $R = Z(L)$,

and therefore $L/Z(L)$ is semisimple. By Proposition 3.5.5 (Reductive algebras as direct sums, p. 121), we have $L = [L, L] \oplus Z(L)$ and $[L, L]$ semisimple.

(2) If V is irreducible, then $V = V_\lambda$. This means the elements of $R \subseteq \mathfrak{gl}(V)$ are all scalar matrices, so they are certainly central. This shows $R = Z(L)$ and $\dim R \leq 1$ (since the scalar matrices are a subspace of dimension 1).

(3) If $L \subseteq \mathfrak{sl}(V)$, each element of R has trace 0. The only scalar matrix with trace 0 is 0, so $R = 0$. ■

CHAPTER 4

Semisimple Lie algebras and Root Space Decomposition

In this chapter, we start our study of the structure of semisimple Lie algebras. Such Lie algebras contain a special Abelian subalgebra, called either a maximal toral subalgebra or a Cartan subalgebra. We can decompose the entire Lie algebra as a direct sum of eigenspaces, called root spaces, relative to this Abelian Lie algebra. These root spaces and the corresponding roots possess a number of remarkable properties. This will ultimately lead us to a classification of semisimple Lie algebras when $\mathbb{F}$ is algebraically closed.

1. Toral, Cartan, and Borel subalgebras

We now define several special subalgebras of an arbitrary Lie algebra, although these subalgebras are mainly of interest for semisimple or reductive Lie algebras. We will not make use of all of these, but we have gathered them together in one place for reference. The most important of these is the Cartan subalgebra, which will turn out to be the same as a maximal toral subalgebra (see Proposition 4.2.3 (Maximal toral subalgebras self-normalize, p. 131)). Later applications to representation theory will make use of Borel subalgebras. In the definitions below, $\text{ad } x$ means $\text{ad}_L x$ operating on all of L unless otherwise stated.

We say a subalgebra S of L is a **split subalgebra** if $\text{ad}_L x$ is split for all $x \in S$, i.e., all eigenvalues of $\text{ad } x$ lie in $\mathbb{F}$. We say a subalgebra T of L is **toral** if T is Abelian and $\text{ad}_L x$ is semisimple for every element $x \in T$. For example, the set of diagonal matrices is a split toral subalgebra of $\mathfrak{gl}(n)$ or of $\mathfrak{sl}(n)$. The next result tells us the Abelian condition is automatic if $\mathbb{F}$ is algebraically closed, and the example in the remarks shows that this is not true in general.

Proposition 4.1.1 (ad-diagonalizable implies Abelian).

Let L be a Lie algebra such that $\text{ad } x$ is diagonalizable for each $x \in L$. Then L is Abelian.

REFERENCES AND REMARKS. See [Humphreys, §8.1, Lemma, p. 35; Erdmann-Wildon, Lemma 18.6, p. 221; Wikipedia].

In Subsection 4.5 (The Jordan-Chevalley decomposition, p. 180) of Chapter 8 (Miscellaneous Calculations and Observations, p. 173), we show that every element of the Lie algebra $L = \mathfrak{so}(3, \mathbb{R})$ is semisimple, even though L is not Abelian. This algebra has basis

x, y, z and brackets $[x, y] = z, [y, z] = x, [z, x] = y$. Over $\mathbb{C}$, it is isomorphic to $\mathfrak{sl}(2, \mathbb{C})$ and not all elements are semisimple.

PROOF. Let $x, y \in L$ be nonzero and suppose $[x, y] = \alpha y$ for some scalar α . Then $\text{ad } y(x) = -\alpha y$ and so $(\text{ad } y)^2(x) = 0$. Since $\text{ad } y$ is diagonalizable, this implies $\text{ad } y(x) = 0$ and this in turn implies $\alpha = 0$. [To see this, suppose $v_1, \dots, v_n$ is a basis of $\text{ad } y$ -eigenvectors with corresponding eigenvalues $\alpha_1, \dots, \alpha_n$. Write $x = \sum_i \lambda_i v_i$ and note $0 = (\text{ad } y)^2(x) = \sum_i \alpha_i^2 \lambda_i v_i$. This implies every i with $\lambda_i \neq 0$ must have $\alpha_i = 0$, and so $\text{ad } y(x) = 0$.] Thus 0 is the only eigenvalue of $\text{ad } x$. As $\text{ad } x$ is diagonalizable, this implies $\text{ad } x = 0$, i.e., $x \in Z(L)$. This is true for every $x \in L$, so L is Abelian. ■

We say a subalgebra H of L is a **Cartan subalgebra** or **CSA** if H is nilpotent and $H = N_L(H)$ (where $N_L(H) = \{x \in L \mid [x, H] \subseteq H\}$). The reader may wish to show that the set of diagonal matrices is a split Cartan subalgebra of $\mathfrak{sl}(n)$ or $\mathfrak{gl}(n)$.

For $x \in L$, we define $L_{[0]}(x)$ to be $\cup_{n=0}^{\infty} \ker(\text{ad } x)^n$. This is the **Fitting null component** of $\text{ad } x$; we abuse terminology by also calling it the Fitting null component of x . That is, $L_{[0]}(x)$ is the set of $y \in L$ such that $[x, [x, \dots, [x, y] \dots]] = 0$ for some number of x s. A subalgebra of L of the form $L_{[0]}(x)$ for some $x \in L$ is called an **Engel subalgebra** of L . [Humphreys, §15.1, p. 78] and [Jacobson] use $L_0(x)$ or $L_0(\text{ad } x)$ and [Bourbaki 7-9, §VIII.2.2, p. 78] uses $L^0(x)$ for what we have denoted $L_{[0]}(x)$. We reserve $L_0(x)$ for $\{y \in L \mid [x, y] = 0\} = C(x)$.

We say $x \in L$ is **regular** if its Fitting null component $L_{[0]}(x)$ has smallest possible dimension for Fitting null components of elements of L . This definition follows [Jacobson, §III.1, Definition 2, p. 58] and [Bourbaki 7-9, §VII.2.2, pp. 16-18] (though Bourbaki's official definition is different). However, **Wikipedia** says an element is regular if its centralizer $C(x) = \{y \in L \mid [x, y] = 0\}$ has smallest possible dimension for centralizers of elements of L . If x is semisimple, then $L_{[0]}(x) = L_0(x) = C(x)$, and so the definitions agree for semisimple elements but they do not agree in general. (For example, in $\mathfrak{sl}(2)$, every nonzero x has $\dim C(x) = 1$, but x is regular in our definition if and only if it is not nilpotent, equivalently, is nonzero and semisimple.) [Humphreys, §15.3, p. 80] calls a semisimple element x with $C(x)$ a maximal toral subalgebra **regular semisimple**. [Bourbaki 7-9, §VII.2.4, Corollary 2, p. 20] shows every regular element of a semisimple Lie algebra is semisimple. On the other hand, if L is nilpotent, all elements of L are regular: see [Bourbaki 7-9, §VII.2.2, Example 1, p. 16].

When $L = \mathfrak{gl}(n)$, the element x is regular if and only if its minimal polynomial equals its characteristic polynomial, i.e., if and only if there is only one Jordan block for each eigenvalue. It turns out $x \in \mathfrak{gl}(n)$ is regular if and only if $C(x)$ is equal to the space of polynomials in x , and in this case, $\dim C(x) = n$. If x is not regular, then

$\dim C(x) > n$. In $\mathfrak{sl}(n)$, an element is regular if and only if it has n distinct eigenvalues in $\overline{\mathbb{F}}$. **Justify/reference**

Proposition 4.1.2 (Cartan subalgebras and regular elements).

Let L be a finite-dimensional semisimple Lie algebra over a field of characteristic 0 and let $x \in L$.

- (1) *x is regular if and only if $L_{[0]}(x)$ is a Cartan subalgebra of L .*
- (2) *Every Cartan subalgebra of L has the form $L_{[0]}(x)$ for some regular element $x \in L$.*
- (3) *All Cartan subalgebras of L have the same dimension.*
- (4) *The Cartan subalgebras of L are precisely the minimal Engel subalgebras of L .*
- (5) *The Cartan subalgebras of L are precisely the maximal toral subalgebras of L .*

REFERENCES AND REMARKS. See [Humphreys, §8, Exercise 7, p. 41; §IV.15.3, p. 80; Jacobson, §III.1, Theorem 1, p. 59; §IX.2, Remarks on p. 274; Bourbaki 7-9, §VII.2.3, Theorem 1, p. 18 and §VII.3.3, Theorem 2, p. 24; Wikipedia].

PROOF. **Later or omitted.** One half of (5) is Proposition 4.2.3 (Maximal toral subalgebras self-normalize, p. 131). ■

We say a subalgebra B of L is a **Borel subalgebra** if B is a maximal solvable subalgebra. ([Bourbaki 7-9, §VIII.3.3, Definition 1, p. 90 and §VIII.10, Corollary to Theorem 3, p. 163] also requires that B contain a Cartan subalgebra and then proves that this second condition is automatic if $\mathbb{F}$ is algebraically closed.) For example, the set of upper triangular matrices is a Borel subalgebra of $\mathfrak{gl}(n)$ or $\mathfrak{sl}(n)$, as is the set of lower triangular matrices. The next proposition gives a classification of Borel subalgebras of $\mathfrak{gl}(n)$.

Proposition 4.1.3 (Borel subalgebras and flags).

Let V be an n -dimensional vector space over a field of characteristic 0 and let B be a subalgebra of $\mathfrak{gl}(V)$. Then B is a Borel subalgebra if and only if there is a full flag $0 = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_n = V$ such that $B = \{x \in \mathfrak{gl}(V) \mid x(V_i) \subseteq V_i \text{ for all } i \geq 1\}$.

REFERENCES AND REMARKS. See [Bourbaki 7-9, §VIII.13.1, p. 191; Wikipedia].

PROOF. **Later or omitted.** ■

Our next result requires some terminology, which does not seem to be universally agreed upon. An automorphism of L of the form $e^{\text{ad } x}$, for $x \in L$ with $\text{ad } x$ nilpotent is called **inner** or **elementary**. We will extend this terminology to finite compositions of such automorphisms. [Jacobson, Chapter IX, p. 265] calls these invariant automorphisms.

[**Humphreys**, §16, pp. 81-82] introduces a putatively stronger notion that in the end turns out to be the same.

Proposition 4.1.4 (Borel subalgebras are conjugate).

Let L be a finite-dimensional Lie algebra over an algebraically closed field $\mathbb{F}$ of characteristic 0.

- (1) *Any two Cartan subalgebras of L are conjugate under an elementary automorphism of L .*
- (2) *Any two Borel subalgebras of L are conjugate under an elementary automorphism of L .*

REFERENCES AND REMARKS. See [**Humphreys**, §16.4, Theorem & Corollary, p. 84; **Jacobson**, §IX.2, Theorem 3, p. 273; **Bourbaki 7-9**, §VII.3.2, Theorem 1, p. 23 and §VIII.3.3, Corollary, p. 92]; **Wikipedia**.

The proposition requires algebraic closure. However, in general, all split Cartan subalgebras are conjugate by elementary automorphisms.

PROOF. **Later or omitted.** ■

When $\mathbb{F}$ is algebraically closed, a **parabolic subalgebra** of L is a subalgebra that contains a Borel subalgebra. For general $\mathbb{F}$, we say P is parabolic if $P_{\overline{\mathbb{F}}}$ is a parabolic subalgebra of $L_{\overline{\mathbb{F}}}$. See [**Humphreys**, Chapter 4, Exercise 6, pp. 87-88; **Bourbaki 7-9**, §VIII.3.5, Definition 3, p. 95)]. These definitions are usually only made for reductive Lie algebras and may not be used in the general case.

Proposition 4.1.5 (Trace form characterization of parabolic subalgebras).

*Let V be a finite-dimensional vector space over an algebraically closed field of characteristic 0. Let $L \subseteq \mathfrak{gl}(V)$ be a reductive linear Lie algebra, let $\langle \cdot, \cdot \rangle = \tau$ be the concrete trace form on L , and let P be a subalgebra of L . If $P^\perp$ is a subalgebra of L consisting of nilpotent operators, then P is a parabolic subalgebra of L . **Does this require $\mathbb{F}$ algebraically closed?***

REFERENCES AND REMARKS. See [**Bourbaki 7-9**, §VIII.10, Theorem 1, p. 159]; **Wikipedia**.

PROOF. **Omitted.** ■

2. Roots and Root Spaces

In this section, we begin the study of the root space decomposition of a semisimple Lie algebra and state a number of its properties. The Killing form plays a prominent role, as does the representation theory of $\mathfrak{sl}(2)$.

2.1. Root space decomposition. If h is a diagonalizable operator on V , then there are a basis $v_1, \dots, v_n$ for V and scalars $\lambda_1, \dots, \lambda_n$ such that $hv_i = \lambda_i v_i$ for each i . Another way to say this is that $V = \mathbb{F}v_1 \oplus \dots \oplus \mathbb{F}v_n$ where h acts on $\mathbb{F}v_i$ by the scalar λ_i . If we combine those v_j s with the same eigenvalue λ , we could also think of this as $V = \bigoplus_{i=1}^k V_i$ where h acts on each eigenspace V_i by scalar multiplication.

If we replace h by a diagonalizable set H , we can write $V = \bigoplus_{i=1}^k V_i$ where each $h \in H$ acts on each space V_i by scalar multiplication. If $hv = \lambda v$, then $(5h)v = (5\lambda)v$ and if $h'v = \lambda'v$, then $(h + h')v = (\lambda + \lambda')v$. Thus the map ($h \mapsto$ the scalar with which h acts on V_i) is a linear map from H to $\mathbb{F}$, i.e., it is a member of the dual space H^* . If $\alpha \in H^*$, we can define the α -**eigenspace** of H on V to be $V_\alpha = \{v \in V \mid hv = \alpha(h)v \text{ for all } h \in H\}$. It's easy to see that V_α is indeed a subspace of V . If $V_\alpha \neq 0$, we say α is an **eigenvalue** of H . The terminology we've just introduced is natural, but it is not the terminology most sources use. Instead, we generally call an eigenvector a **weight** and an eigenspace a **weight space**. (Usually weight spaces must be nonzero.) The fact that H is (simultaneously) diagonalizable is equivalent to the statement that $V = \bigoplus_{\alpha \in H^*} V_\alpha$. Of course only finitely many V_α can be nonzero.

We now apply and improve the above observations in the case H is a split toral subalgebra of L and $V = L$. The best results will require that H be a maximal toral subalgebra.

Proposition 4.2.1 (Root space decomposition).

Let L be a finite-dimensional Lie algebra over a field of characteristic 0. Let H be a split toral subalgebra of L and let $\langle, \rangle$ be a symmetric invariant bilinear form on L .

- (1) $L = \bigoplus_{\alpha \in H^*} L_\alpha$ and $[L_\alpha, L_\beta] \subseteq L_{\alpha+\beta}$ for all $\alpha, \beta \in H^*$.
- (2) If $\alpha, \beta \in H^*$ and $\beta \neq -\alpha$, then $\langle L_\alpha, L_\beta \rangle = 0$.
- (3) Suppose $\langle, \rangle$ is nondegenerate. (This means L is semisimple.) If $\alpha \in H^*$, $x \in L_\alpha$, and $\langle x, L_{-\alpha} \rangle = 0$, then $x = 0$.

REFERENCES AND REMARKS. See [Humphreys, §8.2, Proposition, p. 36; Jacobson, §III.2, Corollary, p. 64; §IV.1, p. 108; Bourbaki 7-9, §VII.1.3, Proposition 10, pp. 9-10; §VIII.2.2, Proposition 1, p. 78; Erdmann-Wildon, Lemma 10.1, p. 93].

This is a good place to remark that not all semisimple Lie algebras have a split maximal toral subalgebra. For example, $\mathfrak{su}(2, \mathbb{R}) = L_{1,1,1}$, does not contain a split maximal toral subalgebra. Moreover, it is possible for a semisimple Lie algebra to have a split maximal toral subalgebra and a non-split maximal toral subalgebra. See [Bourbaki 7-9, §VIII.Exercises.2, Exercises 1 & 2, p. 226]; Wikipedia.

PROOF. Suppose throughout that $x \in L_\alpha$, $y \in L_\beta$, $h \in H$.

(1) The decomposition is a result of the general theory, discussed above. Moreover,

$$[h, [x, y]] = [x, [h, y]] + [[h, x], y] = [x, \beta(h)y] + [\alpha(h)x, y] = (\alpha(h) + \beta(h))[x, y].$$

Thus $[x, y] \in L_{\alpha+\beta}$.

(2) Suppose $\beta \neq -\alpha$. Then

$$\alpha(h)\langle x, y \rangle = \langle [h, x], y \rangle = -\langle [x, h], y \rangle = -\langle x, [h, y] \rangle = -\beta(h)\langle x, y \rangle.$$

Thus $(\alpha(h) + \beta(h))\langle x, y \rangle = 0$ for all $h \in H$. Since $\beta \neq -\alpha$, there is at least one h with $\alpha(h) + \beta(h) \neq 0$. This implies $\langle x, y \rangle = 0$.

(3) Suppose $\langle x, L_{-\alpha} \rangle = 0$. By (2), we have $\langle x, L_{\beta} \rangle = 0$ for all $\beta \neq -\alpha$. By (1), this implies $\langle x, L \rangle = 0$. Since $\langle, \rangle$ is nondegenerate, we have $x = 0$. ■

When L is semisimple and we are concerned with the adjoint action of H on L , we call weights **roots**, weight vectors **root vectors**, and weight spaces **root spaces**, with the exception that the weight 0 is not called a root and likewise for vectors and spaces. We denote the set of roots by Φ . This Φ is just a set, but it turns out to have a very interesting structure when L is semisimple and H is a split maximal toral subalgebra of L .

Recall that $C(H) = \{x \in L \mid [h, x] = 0 \text{ for all } h \in H\}$ is the centralizer of H and $N(H) = \{x \in L \mid [h, x] \in H \text{ for all } h \in H\}$ is the normalizer of H . These are subalgebras of L and $H \triangleleft N(H)$. In our case, $C(H)$ is the weight space L_0 and $H \subseteq C(H) \subseteq N(H)$. They are all equal when H is a maximal toral subalgebra. The following result requires L to be semisimple and H to be maximal toral.

Proposition 4.2.2 (Killing form and toral subalgebras).

Let L be a finite-dimensional semisimple Lie algebra over a field of characteristic 0 and let H be a toral subalgebra. Then the Killing form restricted to $C(H)$ is nondegenerate.

REFERENCES AND REMARKS. See [Humphreys, §8.2, Corollary, p. 36; Jacobson, §IV.1, II, p. 109; Bourbaki 7-9, §VII.3.1, Proposition 10(iii), p. 10; Erdmann-Wildon, Lemma 10.1, p. 93].

Does this result require H to be split?

PROOF. Suppose H is split and $x \in C = C(H) = L_0$ satisfies $\kappa(x, y) = 0$ for all $y \in C$. Let $v \in L$. Then we can write $v = y + z$ where $y \in L_0$, $z \in \bigoplus_{0 \neq \alpha \in H^*} L_{\alpha}$. Since $x \in L_0$, Proposition 4.2.1 (Root space decomposition, p. 129) tells us $\kappa(x, z) = 0$. By hypothesis, $\kappa(x, y) = 0$, so we have $\kappa(x, v) = 0$. Since κ is nondegenerate, we must have $x = 0$.

non-split? ■

We now proceed to prove H , $C(H)$, L_0 , and $N(H)$ are all the same for a maximal toral subalgebra H . This shows a maximal toral subalgebra is a Cartan subalgebra. We

will eventually (**will we?**) show every Cartan subalgebra is a maximal toral subalgebra; the main thing to show is that Cartan subalgebras are Abelian.

Proposition 4.2.3 (Maximal toral subalgebras self-normalize).

Let L be a finite-dimensional semisimple Lie algebra over a field of characteristic 0 and let H be a maximal toral subalgebra. Then $N(H) = H$, and so H is a Cartan subalgebra of L .

REFERENCES AND REMARKS. See [Humphreys, §8.2, pp. 36-37; §15.3, top of p. 80; Jacobson, §IV.1, IV on p. 110; Bourbaki 7-9, §§VII.2.3 & 4, Proposition 10 & Theorem 2, p. 19; Erdmann-Wildon, Definition 10.2, p. 94; Exercise 10.9, p. 106; Wikipedia].

When using these references, translation is often required. For example, most sources use Cartan subalgebras in place of maximal toral subalgebras. These turn out to be the same thing, but that requires proof. In some sense, then, one could regard the statement that $N(H) = H$ for a maximal toral subalgebra as the same as the statement that a Cartan subalgebra is Abelian.

PROOF. Set $C = C(H)$ and assume H is split. Throughout, if $x \in L$, we let $x = x_s + x_n$ be the abstract Jordan-Chevalley decomposition of x (Proposition 3.5.4 (Abstract Jordan-Chevalley decomposition, p. 120)). Thus $x_s, x_n \in L$, $\text{ad } x_s$ is semisimple, $\text{ad } x_n$ is nilpotent, and $\text{ad } x_s, \text{ad } x_n$ commute and are polynomials in $\text{ad } x$ with zero constant coefficient.

Our strategy is to first prove that $C = H$ and then use the root space decomposition to conclude $N(H) = H$. Using the Jordan-Chevalley decomposition, we show C is nilpotent and then show C is Abelian. We can then show the semisimple part of $x \in C$ is in H and the nilpotent part is 0, so $C = H$. The final conclusion $N(H) = H$ follows quickly.

We first prove some facts we will use. The first two don't require C to be maximal, but the third does.

- (a) If x and y commute and $\text{ad } x$ is nilpotent, then $\kappa(x, y) = 0$.
- (b) If $x \in Z(C)$ and $\text{ad } x$ is nilpotent, then $x = 0$.
- (c) If $x \in C$, then $x_s \in H$ and $x_n \in C$.

To prove (a), note that $\text{ad } x$ commutes with $\text{ad } y$, so $\text{ad } x \text{ad } y$ is nilpotent. Thus $\kappa(x, y) = \text{tr } \text{ad } x \text{ad } y = 0$.

(b) is a corollary of (a), since (a) implies $\kappa(x, y) = 0$ for all $y \in C$. Proposition 4.2.2 (Killing form and toral subalgebras, p. 130) now implies $x = 0$.

To prove (c), note that if $x \in C$, then $\text{ad } x(H) = 0$. Thus $\text{ad } x_s(H) = \text{ad } x_n(H) = 0$, since both operators are polynomials in $\text{ad } x$ with zero constant coefficient. This shows $x_s, x_n \in C$. Since sums of commuting semisimple operators are semisimple, all elements of $H + \mathbb{F}x_s$ commute and are semisimple. Thus $H + \mathbb{F}x_s$ is a toral subalgebra of L ; maximality forces $x_s \in H$.

Suppose we show C is Abelian. By (c), $x_s \in H$ and by (b), $x_n = 0$, whence $x \in H$. Thus our goal now is to show C is Abelian.

Suppose y is in the radical of $\kappa|_H$ and suppose $x \in C$. Then $x_s \in H, x_n \in C$, and so $\kappa(x_s, y) = 0$. Since $y \in H$, we have that x_n, y commute, and so (a) gives $\kappa(x_n, y) = 0$. Thus $\kappa(x, y) = 0$ for all $x \in C = L_0$. Proposition 4.2.2 (Killing form and toral subalgebras, p. 130) shows $y = 0$. This proves $\kappa|_H$ is nondegenerate.

Take an arbitrary $x \in C$. Since $x_s \in H$, we have $\text{ad}_C x_s = 0$, while $\text{ad}_C x_n$ is nilpotent. Thus $\text{ad}_C x$ is nilpotent. By Proposition 2.4.4 (Engel's theorem for ad , p. 64), this implies C is nilpotent.

If C is not Abelian, then $[C, C] \neq 0$. As C is nilpotent, this implies $Z(C) \cap [C, C] \neq 0$ by Proposition 2.3.3 (Ideals meet the center in nilpotent algebras, p. 55). Take a nonzero $x \in Z(C) \cap [C, C]$. Now $x_s \in H \subseteq Z(C), x_n \in C$, and so $x_n \in Z(C)$. Since ad_{x_n} is nilpotent, fact (c) gives us $x_n = 0$, and so $x = x_s \in H$. For any $y \in H$, we have $\kappa([C, C], y) = \kappa(C, [C, y]) = \kappa(C, 0) = 0$. Thus $\kappa(x, y) = 0$; since $\kappa|_H$ is nondegenerate, this implies $x = 0$. This contradiction shows C is indeed Abelian. As noted above, this implies $C = H$.

Now let $x \in N(H)$, say $x = h + \sum_{\alpha \neq 0} x_\alpha$, where $h \in H, x_\alpha \in L_\alpha$. If $y \in H$, we have $H \ni [x, y] = \sum_{\alpha \neq 0} -\alpha(y)x_\alpha$. This means each term $\alpha(y)x_\alpha$ is 0. If $x \notin H$, then some $x_\alpha \neq 0$, and so $\alpha(y) = 0$. This is true for all $y \in H$, and so $\alpha = 0$. Thus in fact, $x = h \in H$.

non-split case? ■

2.2. Root space decomposition for $\mathfrak{sl}(n)$ and $\mathfrak{gl}(n)$. Let's examine the root space decomposition of L when $L = \mathfrak{gl}(n)$, which is reductive but not semisimple, or $L = \mathfrak{sl}(n)$, which is semisimple. The reader is encouraged to compare our discussion to Subsection 2.3 (Root space decomposition for semisimple Lie algebras, p. 134). We sum up our results for $\mathfrak{sl}(n)$ in Proposition 4.2.4 (Properties of the root space decomposition for $\mathfrak{sl}(n)$, p. 134), which should be compared to Proposition 4.2.5 (Properties of the root space decomposition, p. 135) in the next subsection.

In both cases we take H to be the algebra of diagonal matrices in L . Throughout our discussion, we will take $h = \text{diag}(d_1, \dots, d_n) \in H$ and $x = (a_{ij}) \in L$. Note hx multiplies row i of x by d_i and xh multiplies column j of x by d_j , and so $[h, x] = hx - xh = ((d_i - d_j)a_{ij})$. It's clear that $[h, x] = \lambda x$ if and only if $d_i - d_j = \lambda$ whenever $a_{ij} \neq 0$. In particular, if any $a_{ii} \neq 0$, we must have $\lambda = 0$. If A has two nonzero entries, say a_{ij} and $a_{k\ell}$, then we must have $d_i - d_j = d_k - d_\ell$ for all choices of $d_i, d_j, d_k, d_\ell \in \mathbb{F}$. This is impossible unless $(i, j) = (k, \ell)$. (Exercise!) Thus the only eigenspaces for nonzero eigenvalues, a.k.a. root

spaces, are the spans $L_{ij} = \mathbb{F}e_{ij}$ of the matrix units for $i \neq j$. The roots are the linear functionals $\pm\alpha_{ij}$ for $1 \leq i < j \leq n$, where $\alpha_{ij}(h) = d_i - d_j$.

Recall that τ is the concrete trace form given by $\tau(a, b) = \text{tr}ab$ and that τ is nondegenerate on $\mathfrak{gl}(n)$ or $\mathfrak{sl}(n)$. We stated in Proposition 2.7.1 (Killing form for classical Lie algebras, p. 71) that $\kappa = 2n\tau$ for $\mathfrak{sl}(n)$; we will use this fact freely. Assume now that $i < j$ and set $h_{ij} = e_{ii} - e_{jj}$ and $f_{ij} = e_{ji}$. Then $\tau(h_{ij}, h) = d_i - d_j = \alpha_{ij}(h)$. In particular $\tau(h_{ij}, h_{ij}) = \alpha_{ij}(h_{ij}) = 2$. Moreover, $\tau(e_{ij}, e_{ji}) = 1$ and $[e_{ij}, e_{ji}] = h_{ij}$. Also $[h_{ij}, e_{ij}] = 2e_{ij}$ and $[h_{ij}, f_{ij}] = -2f_{ij}$. This shows that the span of e_{ij}, h_{ij}, f_{ij} is a copy of $\mathfrak{sl}(2)$.

When is a linear combination of roots a root? First note that no multiple of α_{ij} by a number other than ± 1 is a root. If $\alpha = \alpha_{ij}$, $\beta = \alpha_{k\ell}$ are roots, then $\beta + \alpha$ is a root precisely when $j = k$ or $i = \ell$, and $\beta - \alpha$ is a root precisely when $j = \ell$ or $i = k$. In either case, $\beta + 2\alpha$ and $\beta - 2\alpha$ are never roots. Thus we have at most a string of roots $\beta - \alpha, \beta, \beta + \alpha$, and no other $\beta + r\alpha$ is a root. However, other linear combinations are possible. Note if $i < j$, then $\alpha_{ij} = \alpha_{i,i+1} + \alpha_{i+1,i+2} + \cdots + \alpha_{j-1,j}$ and $\alpha_{ji} = -\alpha_{i,i+1} - \alpha_{i+1,i+2} - \cdots - \alpha_{j-1,j}$. It's easy to see that $\alpha_{12}, \alpha_{23}, \dots, \alpha_{n-1,n}$ are linearly independent. Not only is this set a basis of H^* in the $L = \mathfrak{sl}(n)$ case, but every root is either a (consecutive) sum of some of the $\alpha_{i,i+1}$ or a (consecutive) sum of the negatives of some of the $\alpha_{i,i+1}$. In the $L = \mathfrak{gl}(n)$ case, we need to add the linear functional tr to get a basis of H^* .

We are interested also in the form τ^* on H^* that is dual to the form τ on H (really $\tau|_H$); it is given by $\tau^*(\beta, \alpha) = \tau(h_\beta, h_\alpha) = \beta(h_\alpha)$. We will see that in general this should be an integer of small magnitude. In our situation, this is given by:

$$\tau^*(\alpha_{k\ell}, \alpha_{ij}) = \tau(h_{k\ell}, h_{ij}) = \alpha_{k\ell}(h_{ij}) = \begin{cases} 2, & \text{if } (i, j) = (k, \ell) \\ -2, & \text{if } (i, j) = (\ell, k) \\ 1, & \text{if } k = i \text{ or } \ell = j \text{ but } (i, j) \neq (k, \ell) \\ -1, & \text{if } k = j \text{ or } \ell = i \text{ but } (i, j) \neq (\ell, k) \\ 0, & \text{if } \{i, j\} \cap \{k, \ell\} = \emptyset. \end{cases}$$

The root space decomposition depends on the choice of the Cartan/toral subalgebra H . For example, if $L = \mathfrak{sl}(2)$, we can take a basis e, h, f with the usual relations by setting

$$e = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \quad h = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad f = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}.$$

The Cartan/maximal toral subalgebra is $H = \mathbb{F}h$ and the root vectors are now the new e, f , not e_{12}, e_{21} as before.

We now state Proposition 4.2.5 (Properties of the root space decomposition, p. 135) from the next subsection for the special case of $L = \mathfrak{sl}(n)$.

Proposition 4.2.4 (Properties of the root space decomposition for $\mathfrak{sl}(n)$).

Let $L = \mathfrak{sl}(n, \mathbb{F})$ over a field $\mathbb{F}$ of characteristic 0 with Killing form κ and let H be the split maximal toral subalgebra consisting of the diagonal matrices in L . Let $\alpha_{ij} \in H^*$ be as defined above, namely $\alpha_{ij}(h)$ is the $d_i - d_j$ for $h = \text{diag}(d_1, \dots, d_n)$. The set Φ of roots of L relative to H consists of the $\pm\alpha_{ij}$ for $1 \leq i < j \leq n$. Let e_{ij} be the i, j matrix unit, let $f_{ij} = e_{ji}$, and let $h_{ij} = [e_{ij}, f_{ij}] = e_{ii} - e_{jj}$. Set $L_{ij} = \mathbb{F}e_{ij}$; this is the weight space for α_{ij} . Let $t_{ij} = \frac{1}{2n}h_{ij}$, whence $\alpha_{ij}(h) = \kappa(t_{ij}, h)$ for all $h \in H$, and note $h_{ij} = \frac{2t_{ij}}{\kappa(t_{ij}, t_{ij})}$.

- (1) $\alpha_{ij}(t_{ij}) = \kappa(t_{ij}, t_{ij}) = \frac{1}{n} \neq 0$ and $\alpha_{ij}(h_{ij}) = 2$.
- (2) Φ spans H^* and $\{h_{ij} \mid 1 \leq i < j \leq n\}$ spans H .
- (3) If $x \in L_{ij}$, $y \in L_{ji}$, then $[x, y] = \kappa(x, y)t_{ij} = \frac{1}{2n}\kappa(x, y)h_{ij}$.
- (4) If $x \in L_{ij}$, $y \in L_{ji}$, then $[h_{ij}, x] = 2x$ and $[h_{ij}, y] = -2y$. Thus e_{ij}, h_{ij}, f_{ij} span a subalgebra of L isomorphic to $\mathfrak{sl}(2)$.
- (5) $[L_{ij}, L_{ji}] = \mathbb{F}h_{ij}$ and $S_{ij} = L_{ij} \oplus [L_{ij}, L_{ji}] \oplus L_{ji}$ is a subalgebra of L isomorphic to $\mathfrak{sl}(2)$.
- (6) $\alpha_{kl}(h_{ij}) \in \{0, \pm 1, \pm 2\}$, with ± 2 only occurring if $(i, j) = (k, \ell)$ or $(i, j) = (\ell, k)$. Moreover, $\alpha_{kl} - \alpha_{kl}(h_{ij})\alpha_{ij} \in \Phi$.
- (7) If $\alpha_{ij} + \alpha_{kl} \in \Phi$, then $[L_{ij}, L_{kl}] = L_{\alpha_{ij} + \alpha_{kl}}$.
- (8) L is the sum, as a vector space, of the subalgebras S_{ij} defined in (5). L is generated, as a Lie algebra, by the root spaces L_{ij} , $1 \leq i, j \leq n$.
- (9) If $\lambda \in \mathbb{F}$, then $\lambda\alpha_{ij} \in \Phi$ if and only if $\lambda = \pm 1$.
- (10) If $\alpha = \alpha_{ij}, \beta = \alpha_{kl}$ are roots with $\beta \neq \pm\alpha$, then $\beta + r\alpha$ cannot be a root unless $r = \pm 1$. Moreover, if $\{i, j\}, \{k, \ell\}$ are disjoint, then neither $\beta + \alpha$ nor $\beta - \alpha$ are roots.

If $\{i, j\}, \{k, \ell\}$ are neither disjoint nor equal, then exactly one of $\beta + \alpha, \beta - \alpha$ is a root. More precisely, we have the following. If $i = k, j \neq \ell$, then $\beta - \alpha$ is a root and $\beta + \alpha$ is not. If $i = \ell, j \neq k$, then $\beta + \alpha$ is a root and $\beta - \alpha$ is not. If $j = k, i \neq \ell$, then $\beta + \alpha$ is a root and $\beta - \alpha$ is not. If $j = \ell, i \neq k$, then $\beta - \alpha$ is a root and $\beta + \alpha$ is not.

PROOF. See the comments earlier in this section and Proposition 4.2.5 (Properties of the root space decomposition, p. 135). Other calculations are left to the reader. ■

2.3. Root space decomposition for semisimple Lie algebras. For general semisimple Lie algebras, we may not have a concrete trace form available, so we use the Killing form. Let's translate the results for $\mathfrak{sl}(n)$ to those using the Killing form. For example, we want $t_{ij} \in H$ with $\alpha_{ij}(h) = \kappa(t_{ij}, h)$. Using the facts that $\kappa = 2n\tau$ and $\alpha_{ij}(h) = \tau(h_{ij}, h)$, we see that we should take $t_{ij} = \frac{1}{2n}h_{ij}$. Then $\kappa(t_{ij}, t_{ij}) =$

$2n \frac{1}{4n^2} \tau(h_{ij}, h_{ij}) = \frac{1}{n}$. In the next proposition, we are particularly interested in $\frac{2}{\kappa(t_{ij}, t_{ij})} t_{ij} = 2n \frac{1}{2n} h_{ij}$, which is just the h_{ij} we have already defined.

Let L be a semisimple Lie algebra with maximal toral subalgebra H . The fact that the Killing form is nondegenerate when restricted to H means that for every $\alpha \in H^*$, there is a unique $t_\alpha \in H$ with $\alpha(h) = \kappa(t_\alpha, h)$ for all $h \in H$. (The fact that κ is nondegenerate is equivalent to the statement that the map $H \rightarrow H^*$ given by $t \mapsto t^*$, where t^* is defined by $t^*(h) = \kappa(t, h)$, is an isomorphism of vector spaces. Thus we must have $\alpha = (t_\alpha)^*$ for some unique t_α .) The map $\alpha \mapsto t_\alpha$ is linear, so $t_{\lambda\alpha} = \lambda t_\alpha$ and $t_{\alpha+\beta} = t_\alpha + t_\beta$ for a scalar λ and $\alpha, \beta \in H^*$. We can now define a nondegenerate symmetric bilinear form κ^* on H^* (really $\kappa|_H$) by $\kappa^*(\alpha, \beta) = \kappa(t_\alpha, t_\beta) = \alpha(t_\beta) = \beta(t_\alpha)$. We call this the **dual Killing form**; this may not be the only name used for this form and our notation κ^* is unusual. In Chapter 7 (More Representation Theory, p. 165), we will write (α, β) for $\kappa^*(\alpha, \beta)$.

We will show in the next proposition that $\kappa(t_\alpha, t_\alpha)$ is nonzero, and so we may define $h_\alpha = \frac{2t_\alpha}{\kappa(t_\alpha, t_\alpha)} \in H$. Since $t_{\lambda\alpha} = \lambda t_\alpha$, we have $h_{\lambda\alpha} = \lambda^{-1} h_\alpha$ for a nonzero scalar λ .

Warning: Our notation is probably the most common, but [Jacobson] and [Bourbaki 7-9] use h_α for what we have called t_α . [Jacobson] uses h'_α for what we have called h_α , while [Bourbaki 7-9] uses H_α .

The following result gives many facts about the root space decomposition. One of the things it shows is that a subset of Φ forms a basis of H^* and some of the t_α (or the h_α) with $\alpha \in \Phi$ form a basis of H . Another thing it makes clear is the ubiquity of both the Lie algebra $\mathfrak{sl}(2)$ and the Killing form. The proof is correspondingly long.

Proposition 4.2.5 (Properties of the root space decomposition).

Let L be a finite-dimensional semisimple Lie algebra over a field of characteristic 0 with Killing form κ and let H be a split maximal toral subalgebra. Let Φ be the set of roots of L relative to H . Let $\alpha \in \Phi$, let t_α be the unique element of H with $\alpha(h) = \kappa(t_\alpha, h)$ for all $h \in H$, and let $h_\alpha = \frac{2t_\alpha}{\kappa(t_\alpha, t_\alpha)}$. [We will show $\kappa(t_\alpha, t_\alpha) \neq 0$.]

- (1) $\alpha(t_\alpha) = \kappa(t_\alpha, t_\alpha) \neq 0$ (so h_α is defined) and $\alpha(h_\alpha) = 2$.
- (2) Φ spans H^* and $\{h_\alpha \mid \alpha \in \Phi\}$ spans H .
- (3) If $x \in L_\alpha$, $y \in L_{-\alpha}$, then $[x, y] = \kappa(x, y)t_\alpha = \frac{1}{2}\kappa(x, y)\kappa(t_\alpha, t_\alpha)h_\alpha$.
- (4) If $x \in L_\alpha$, $y \in L_{-\alpha}$, then $[h_\alpha, x] = 2x$ and $[h_\alpha, y] = -2y$. If $e_\alpha \in L_\alpha$ is nonzero, there exists a unique $f_\alpha \in L_{-\alpha}$ with $[e_\alpha, f_\alpha] = h_\alpha$, and the span of $e_\alpha, h_\alpha, f_\alpha$ is a subalgebra of L isomorphic to $\mathfrak{sl}(2)$.
- (5) L_α and $[L_\alpha, L_{-\alpha}]$ are 1-dimensional, the latter with basis h_α . Moreover, $S_\alpha = L_\alpha \oplus [L_\alpha, L_{-\alpha}] \oplus L_{-\alpha}$ is a subalgebra of L isomorphic to $\mathfrak{sl}(2)$.
- (6) If $\beta \in \Phi$, then $\beta(h_\alpha) \in \mathbb{Z}$ and $\beta - \beta(h_\alpha)\alpha \in \Phi$.

- (7) If $\beta, \alpha + \beta \in \Phi$, then $[L_\alpha, L_\beta] = L_{\alpha+\beta}$.
- (8) L is the sum, as a vector space, of the subalgebras S_α , $\alpha \in \Phi$, defined in (5). L is generated, as a Lie algebra, by the root spaces L_α , $\alpha \neq 0$.
- (9) If $\lambda \in \mathbb{F}$, then $\lambda\alpha \in \Phi$ if and only if $\lambda = \pm 1$.
- (10) Suppose $\beta \in \Phi$, $\beta \neq \pm\alpha$. Suppose r, q are nonnegative integers and $\beta - r\alpha$, $\beta + q\alpha$ are roots. Then $\beta + i\alpha$ is a root for all integers i with $-r \leq i \leq q$. If r, q are the largest integers with $\beta - r\alpha, \beta + q\alpha \in \Phi$, then $\beta(h_\alpha) = r - q$ and $r + q \leq 3$.

REFERENCES AND REMARKS. See [Humphreys, §§8.3-8.5, pp. 37-40; Jacobson, §§IV.1-2, pp. 108-119; Bourbaki 7-9, §VIII.2.2, pp. 78-80, 82; Erdmann-Wildon, §§10.3-10.6, pp. 97-105; Wikipedia].

The numbers $\beta(h_\alpha)$ in (6) are called the **Cartan integers** of L . One can show they do not depend on the choice of maximal toral/Cartan subalgebra. We can also write the number as $\beta(h_\alpha) = 2 \frac{\kappa(t_\alpha, t_\beta)}{\kappa(t_\alpha, t_\alpha)} = 2 \frac{\kappa^*(\alpha, \beta)}{\kappa^*(\alpha, \alpha)}$.

If r, q are the largest integers with $\beta - r\alpha, \beta + q\alpha \in \Phi$ as in (10), then the set of roots $\beta - r\alpha, \beta - (r-1)\alpha, \dots, \beta, \dots, \beta + (q-1)\alpha, \beta + q\alpha$ is called the **α -string through β** . Such sequences are called **root strings**. Such strings have at most 4 elements by (10).

PROOF. (2) We will show the t_α , $\alpha \in \Phi$, span H . Since h_α is a scalar multiple of t_α , this will prove the claim about the h_α . Let W be the subspace of H spanned by the t_α , $\alpha \in \Phi$. If $W \neq H$, then there is a nonzero $h \in W^\perp$. Then $0 = \kappa(t_\alpha, h) = \alpha(h)$ for all $\alpha \in \Phi$. This means $[h, L_\alpha] = 0$ for all $\alpha \in \Phi$, and we also have $[h, H] = 0$. Proposition 4.2.1 (Root space decomposition, p. 129) implies $[h, L] = 0$, i.e., $h \in Z(L) = 0$. This proves $W = H$.

Let $\phi \in H^*$ and let $t_\phi = \sum_{\alpha \in \Phi} \lambda_\alpha t_\alpha$. Then $\phi(h) = \kappa(t_\phi, h) = \sum_{\alpha \in \Phi} \lambda_\alpha \kappa(t_\alpha, h) = \sum_{\alpha \in \Phi} \lambda_\alpha \alpha(h)$ for any $h \in H$, so $\phi = \sum_{\alpha \in \Phi} \lambda_\alpha \alpha$. This proves the α span H^* .

(3) Let $x \in L_\alpha$, $y \in L_{-\alpha}$, $h \in H$. Then

$$\kappa(h, [x, y]) = \kappa([h, x], y) = \alpha(h)\kappa(x, y) = \kappa(h, t_\alpha)\kappa(x, y) = \kappa(h, \kappa(x, y)t_\alpha).$$

Thus $\kappa(h, [x, y]) - \kappa(x, y)\kappa(h, t_\alpha) = 0$ for all $h \in H$. Since $\kappa|_H$ is nondegenerate by Proposition 4.2.2 (Killing form and toral subalgebras, p. 130), this implies $[x, y] = \kappa(x, y)t_\alpha$.

(1) Suppose $0 = \kappa(t_\alpha, t_\alpha) = \alpha(t_\alpha)$ and choose a nonzero $x \in L_\alpha$. Proposition 4.2.1 (Root space decomposition, p. 129)(3) implies there is a $y \in L_{-\alpha}$ with $\kappa(x, y) \neq 0$. Let S be the span of x, y, t_α . Then $[t_\alpha, x] = \alpha(t_\alpha)x = 0$. Recalling that $t_{-\alpha} = -t_\alpha$, we likewise have $[t_\alpha, y] = 0$. As $[x, y] = \kappa(x, y)t_\alpha$ by (3) and $\kappa(x, y) \neq 0$, we see that S is a subalgebra isomorphic to the Heisenberg Lie algebra and so is nilpotent. Moreover, $t_\alpha \in [S, S]$.

Regard L as an S -module via the adjoint representation. By Proposition 3.2.6 (Concrete Lie's theorem, p. 100), every element of $[S, S]$ acts nilpotently on L . In particular,

$\text{ad}_L t_\alpha$ is nilpotent. However, $\text{ad}_L t_\alpha$ is also semisimple, since H is toral. Thus $\text{ad}_L t_\alpha = 0$, i.e., $t_\alpha \in Z(L)$. Since L is semisimple, this is impossible.

(4) We see $[h_\alpha, x] = \frac{2}{\kappa(t_\alpha, t_\alpha)}[t_\alpha, x] = \frac{2}{\kappa(t_\alpha, t_\alpha)}\alpha(t_\alpha)x = 2x$ and $[h_\alpha, y] = -[h_{-\alpha}, y] = -2y$. Given a nonzero $e_\alpha \in L_\alpha$, Proposition 4.2.1 (Root space decomposition, p. 129) tells us there is a $y \in L_{-\alpha}$ with $\kappa(e_\alpha, y) \neq 0$. If we define $f_\alpha = \frac{2}{\kappa(e_\alpha, y)\kappa(t_\alpha, t_\alpha)}y$, then $\kappa(e_\alpha, f_\alpha) = \frac{2}{\kappa(t_\alpha, t_\alpha)}$ and $[e_\alpha, f_\alpha] = \frac{2}{\kappa(t_\alpha, t_\alpha)}t_\alpha = h_\alpha$. The relations between $e_\alpha, h_\alpha, f_\alpha$ show they span a copy of $\mathfrak{sl}(2)$. The uniqueness of f_α follows from the fact in (5) that $\dim L_{-\alpha} = 1$.

For the rest of the proof, $\alpha \in \Phi$ will be fixed with $L_\alpha \neq 0$. We set $K = \ker \alpha \subseteq H$. By (4), $L_{-\alpha} \neq 0$, so $-\alpha \in \Phi$. As in (4), we let S be the subalgebra of L spanned by $e_\alpha, h_\alpha, f_\alpha$. We will consider some subspaces of L as S -modules via the adjoint representation, and when we speak of weights, we mean weights relative to S (i.e., eigenvalues of h_α) unless otherwise stated.

(9) Set $M = H \oplus \bigoplus_{\lambda \neq 0} L_{\lambda\alpha}$. Then M is a subalgebra of L and is an S -module under the adjoint action. Each $L_{\lambda\alpha}$ is a weight space with weight $\lambda\alpha(h_\alpha) = 2\lambda$, while H is a weight space with weight 0. No weights other than 0 and $\lambda\alpha$ can occur in M .

Since $K \cap S = 0$ and all $\mathfrak{sl}(2)$ -modules are completely reducible, we have $M = K \oplus S \oplus N$ for some S -submodule $N \subseteq \bigoplus_{\lambda \neq 0} L_{\lambda\alpha}$. Only integers can occur as weights of $\mathfrak{sl}(2)$ -modules, whence $L_{\lambda\alpha} = 0$ unless $\lambda \in \frac{1}{2}\mathbb{Z}$. The only occurrences of weight 0 in M are the ones in K and the one in S , and the submodule S has weights $-2, 0, 2$. Thus there are no other even weights in M and in particular, $\lambda\alpha$ is not a root if $\lambda \in \mathbb{Z}$, $\lambda \neq \pm 1$. If $\lambda\alpha$ is a root for any $\lambda = \frac{n}{2}$ with n odd, then M contains an element of weight $2\lambda = n$, and that element must be in N . This implies N contains an element of weight 1, which must be in $L_{\frac{1}{2}\alpha}$. But if $\frac{1}{2}\alpha$ is a root, then $\alpha = 2(\frac{1}{2}\alpha)$ cannot be a root by the above. This contradiction shows $\lambda\alpha$ is a root only if $\lambda = \pm 1$.

(5) Suppose now $\dim L_\alpha > 1$. Then M contains a second weight vector of weight 0 and so N contains a weight vector of weight 0. We saw above that this is impossible, so $\dim L_\alpha = 1$. It follows from (4) that $[L_\alpha, L_{-\alpha}]$ is nonzero of dimension at most 1, so its dimension is exactly 1. The claim that $S_\alpha = S$ now follows.

(6), (10) We continue with $S = S_\alpha$ and the convention that weight refers to S -weight. If $\beta = \pm\alpha$, then (9) implies (6) and (10) hold, so we assume $\beta \neq \pm\alpha$. Set $W = \sum_{i \in \mathbb{Z}} L_{\beta+i\alpha}$, which is an S -module. By (9), no $\beta + i\alpha = 0$, and by (5), each $L_{\beta+i\alpha}$ is 1-dimensional of integral weight $\beta(h_\alpha) + 2i$. Thus $\beta(h_\alpha) \in \mathbb{Z}$. Clearly the weights 0 and 1 cannot both occur among the $\beta(h_\alpha) + 2i$ and 0 or 1 can occur at most once. The remarks prior to Proposition 3.3.1 (Irreducible representations of $\mathfrak{sl}(2)$, p. 106) thus imply W is an irreducible S -module. Since $\beta(h_\alpha)$ is one of the weights that occur in W , the highest weight of W must have the form $\beta(h_\alpha) + 2q$ and lowest weight must have the form $\beta(h_\alpha) - 2r$ for some nonnegative integers q, r . Moreover, every $\beta(h_\alpha) + 2i$ with $-r \leq i \leq q$ must be

a weight of W , and so $\beta + i\alpha$ must be a root. Furthermore, $\beta(h_\alpha) - 2r = -(\beta(h_\alpha) + 2q)$, so $\beta(h_\alpha) = r - q$. Since $-r \leq -(r - q) \leq q$, we also see that $\beta - (r - q)\alpha = \beta - \beta(h_\alpha)\alpha$ is a weight and hence is a root.

This proves all of the claims in (6) and (10) except that $r + q \leq 3$. Suppose $r + q \geq 4$. By replacing β by $\beta + i\alpha$ for an appropriate i , we can assume $\beta - 2\alpha, \beta - \alpha, \beta, \beta + \alpha, \beta + 2\alpha$ are roots. We know $2\alpha = (\beta + 2\alpha) - \beta$ and $2(\beta + \alpha) = (\beta + 2\alpha) + \beta$ are not roots. Thus the β -string through $\beta + 2\alpha$ has just one term, namely $\beta + 2\alpha$. By (6), this implies $(\beta + 2\alpha)(h_\beta) = 0$. Similarly, $\beta - 2\alpha - \beta, \beta - 2\alpha + \beta$ are not roots and so $(\beta - 2\alpha)(h_\beta) = 0$. Adding these two equations gives $0 = 2\beta(h_\beta) = 4$. This contradiction shows we cannot have a root string of length 5, and so $r + q \leq 3$.

Note: we can also express the conclusion of the last paragraph in terms of the dual Killing form. The fact that the β -string through $\beta + 2\alpha$ has just one term, namely $\beta + 2\alpha$, implies $\kappa^*(\beta + 2\alpha, \beta) = 0$. Similarly, the β -string through $\beta - 2\alpha$ has just one term, namely $\beta - 2\alpha$, whence $\kappa^*(\beta - 2\alpha, \beta) = 0$. Adding these two equalities gives $0 = 2\kappa^*(\beta, \beta) = 2\kappa(t_\beta, t_\beta)$, contrary to (1).

(7) Suppose $\beta, \alpha + \beta$ are roots and continue the notation in the proof of (10). Let $e_\alpha \in L_\alpha, e_\beta \in L_\beta$ be as in (4) and suppose $[e_\alpha, e_\beta] = 0$. This means the action of $e_\alpha \in S$ on e_β is 0, so e_β is a highest weight vector of W , with weight $\beta(h_\alpha)$. On the other hand, since $\alpha + \beta$ is a root, there is a nonzero $x \in L_{\alpha+\beta}$. We have $[h_\alpha, x] = (\alpha + \beta)(h_\alpha)x = (\alpha(h_\alpha) + \beta(h_\alpha))x = (2 + \beta(h_\alpha))x$. Since $2 + \beta(h_\alpha)$ is a larger weight than $\beta(h_\alpha)$, this contradicts e_β being a highest weight vector. Thus $[e_\alpha, e_\beta] \neq 0$. Since $L_{\alpha+\beta}$ is 1-dimensional, this shows $[L_\alpha, L_\beta] = L_{\alpha+\beta}$.

(8) This follows from (2), (4), and (5). ■

2.4. The dual Killing form. Above we defined the dual Killing form κ^* (really $\kappa|_H^*$) by $\kappa^*(\alpha, \beta) = \kappa(t_\alpha, t_\beta)$. Note that $2\frac{\kappa^*(\beta, \alpha)}{\kappa^*(\alpha, \alpha)} = \beta(h_\alpha)$. We state some of the results of Subsection 2.3 (Root space decomposition for semisimple Lie algebras, p. 134) in terms of κ^* .

Proposition 4.2.6 (Properties of the dual Killing form).

Let L be a finite-dimensional semisimple Lie algebra over a field of characteristic 0 and let H be a split maximal toral subalgebra. Let Φ be the set of roots of L relative to H , let κ be the Killing form on L , and let κ^* be the dual Killing form on H^* .

(1) $\kappa^*(\alpha, \alpha) = \alpha(t_\alpha) \neq 0$.

(2) Φ contains a basis of H^* such that every element of Φ is a linear combination of elements of this basis with integer coefficients. integer or rational? Do we know this?

(3) If $\alpha, \beta \in \Phi$, then $2\frac{\kappa^*(\beta, \alpha)}{\kappa^*(\alpha, \alpha)} \in \mathbb{Z}$.

(4) If $\alpha, \beta \in \Phi$, then $\beta - 2\frac{\kappa^*(\beta, \alpha)}{\kappa^*(\alpha, \alpha)}\alpha \in \Phi$.

REFERENCES AND REMARKS. See [Humphreys, §8.5, pp. 39-40].

PROOF. **Later.**

■

CHAPTER 5

Root Systems

We now turn to a geometric side of linear algebra. We will work in a vector space V over $\mathbb{R}$ with an inner product $(,)$. This means that $(,)$ is a symmetric bilinear form that is positive definite: $(v, v) > 0$ for any nonzero $v \in V$. Such a form is automatically nondegenerate. Of course you can take $V = \mathbb{R}^n$ and take $(,)$ to be the usual dot product. Note that our use of $\mathbb{R}$ applies no matter what field $\mathbb{F}$ our (semisimple) Lie algebras are defined over.

Our motivation for this study is that the roots of each semisimple Lie algebra form a special kind of root system, and we can classify all of them. The terminology “root” and “root system” comes from Lie theory. However, root systems have a much wider application. They classify finite Coxeter groups, that is, finite subgroups of $\mathbf{O}(n, \mathbb{R})$ that are generated by reflections. (This is just one definition of Coxeter group.) That is, they classify finite symmetry groups in $\mathbb{R}^n$ that are generated by reflections.

In Section 1 (Coxeter Groups and Root Systems, p. 142), we introduce the study of finite reflection groups and define root systems. In Section 2 (Two-dimensional Root Systems, p. 145), we show what root systems look like in two dimensions and show this restricts the possible angles between roots in a root system. In Section 3 (Positive Roots and Bases, p. 147), we introduce the crucial notions of “positive root” and “base”. We show that bases always exist and that the angles between elements of a base are obtuse. Positive roots and bases are not unique, but we show they are conjugate under the action of the Coxeter group. In Section 4 (Coxeter Graphs and Matrices, p. 151), we define a graph associated to each root system and use those graphs to classify all root systems. In Section 5 (The Crystallographic Condition, p. 157), we introduce the crystallographic condition, a condition that is satisfied by all root systems coming from semisimple Lie algebras. We classify all crystallographic root systems and replace Coxeter graphs with Dynkin diagrams. This will lead to the classification of semisimple Lie algebras in Chapter 6 (Classification of Semisimple Lie Algebras, p. 159).

For the classification of semisimple Lie algebras, we restrict our attention to root systems that satisfy the crystallographic condition, which restricts the rotations that can occur.

1. Coxeter Groups and Root Systems

Let V be an n -dimensional vector space over $\mathbb{R}$ and let $(,)$ be an inner product on V . We also have norm and distance functions in this event: $\|v\| = \sqrt{(v,v)}$ and $d(v,w) = \|v-w\|$. The reader should feel free to take $V = \mathbb{R}^n$ and $(,)$ to be the dot product. We will write $\mathbf{O}(V)$ for $\mathbf{O}(V, (,))$ and $\mathbf{O}(n)$ for $\mathbf{O}(\mathbb{R}^n, \cdot)$. That is, $\mathbf{O}(V)$ is the group of operators ϕ with $(\phi(v), \phi(w)) = (v, w)$ for all $v, w \in V$. An operator ϕ on V is in $\mathbf{O}(V)$ if and only if $(\phi(v), \phi(v)) = (v, v)$ for all $v \in V$. (This is due to polarization, the identity $(v, w) = \frac{1}{2}((v+w, v+w) - (v, v) - (w, w))$.)

The group of symmetries of a figure in $\mathbb{R}^n$ is a subgroup of $\mathbf{O}(n)$. If the figure is a polytope, that subgroup will be finite. If the polytope is regular, its symmetry group is generated by reflections. For example, the group Dih_n of symmetries of a regular n -gon is generated by any two adjacent reflections. Thus we are led to our geometric question: study/classify finite subgroups of $\mathbf{O}(n)$ or $\mathbf{O}(V)$ generated by reflections.

Perhaps we should have recalled the notion of a reflection. First, a **hyperplane** (through the origin) in V is a subspace W with $\dim V/W = 1$. The **reflection about W** is the linear map that fixes W pointwise and takes any $v \in W^\perp$ to $-v$. ([Bourbaki 4-6] calls these **orthogonal reflections**.) We leave it to the reader to verify that the map described is unique, has order 2, and preserves the inner product. The fixed hyperplane is called the **mirror** of the reflection and we say the reflection is **through** or **about** the hyperplane and **along** any perpendicular vector.

If $v \in V$ is nonzero, we define the **reflection determined by v** to be the map $\sigma_v(w) = w - \frac{2(w,v)}{v,v}v$. This is clearly a linear transformation, and its set of fixed points is $v^\perp$ and $\sigma_v(v) = -v$. Thus σ_v is the reflection about the hyperplane $v^\perp$, since both fix $v^\perp$ pointwise and take v to $-v$. One can also verify directly that σ_v has order 2 and preserves the inner product. It also follows that $\sigma_v = \sigma_{v'}$ if and only if v, v' are multiples of each other. (We urge the reader to verify this directly.)

We call v a **root** of the reflection σ_v . We've just remarked that any nonzero multiple of v is also a root of the same reflection. Some sources restrict v , say by requiring its norm of v to be 1. That restriction allows only two roots for each reflection, v and $-v$. We do not make this restriction, but we may single out particular roots for each reflection.

Let $\phi \in \mathbf{O}(V)$ be an isometry and let $v \in V$. Then (using the isometry property of ϕ in the penultimate step)

$$\phi \circ \sigma_v \circ \phi^{-1}(w) = \phi \left(\phi^{-1}(w) - \frac{2(\phi^{-1}(w), v)}{(v, v)}v \right) = w - \frac{2(w, \phi(v))}{(\phi(v), \phi(v))}\phi(v) = \sigma_{\phi(v)}(w).$$

Thus $\phi \circ \sigma_v \circ \phi^{-1} = \sigma_{\phi(v)}$. It follows that if a subgroup G of $\mathbf{O}(V)$ is generated by reflections and Φ is the set of roots of the reflections in G , then $\sigma_v(\Phi) = \Phi$ for every $v \in \Phi$. Because G is generated by reflections, it follows that $\phi(\Phi) = \Phi$ for all $\phi \in G$.

Our next step is to define a root system, and there are multiple ways to do this. Some of the multiplicity comes from the properties one wishes to require, but some comes from a concrete versus abstract approach.

We will start with a concrete approach. If G is a subgroup of $\mathbf{O}(V)$ generated by reflections, we define the **unreduced root system of G** to be the set of roots of reflections in G and we denote it by $\Phi(G)$ or just Φ . That is, $\Phi(G) = \{v \in V \mid \sigma_v \in G\}$. This set has two key properties. One is that if $v \in \Phi$, then $\lambda v \in \Phi$ for all nonzero scalars λ . The second is that if $v, w \in \Phi$, then $\sigma_v(w) \in \Phi$. Some sources don't like the redundancy of allowing so many multiples of v , so they only select a special v (e.g., a v of length 1). Thus a **reduced root system of G** is a collection Φ of roots such that (a) every reflection in G has a root in Φ , (b) for each $v \in \Phi$, the only multiples of v in Φ are $\pm v$, and (c) $\sigma_v(w) \in \Phi$ for all $v, w \in \Phi$. (Condition (c) is not automatic for reduced root systems: we have to choose our roots so that it holds.)

Warning. The notation for a root system and a base (defined below) differ greatly in different sources. [Humphreys] uses Φ for the root system and Δ for a base, just as we do. [Jacobson] does not introduce abstract root systems; he uses π for a base (the set of simple roots). [Bourbaki 4-6] and [Erdmann-Wildon] use R for the root system and B for a base (which [Bourbaki 4-6] calls a basis). [Fulton-Harris] uses R for the root system and does not seem to use a symbol for a base. [Grove-Benson] and [Knapp] use Δ for the root system and Π for a base.

There are other things we desire. For example, if G is finite, then Φ must consist of multiples of a finite subset. For an arbitrary finite Φ , the group generated by σ_v with $v \in \Phi$ need not be finite. As an example of this, take $\Phi = \{(1, 0), (\cos 1, \sin 1)\} \subseteq \mathbb{R}^2$ (or even $\Phi = \{1, \pi\} \subseteq \mathbb{R}^1$). But if we impose the condition $\sigma_v(w) \in \Phi$ for $v, w \in \Phi$, then G will be finite. To see this, note that G acts on the finite set Φ , which induces a homomorphism $G \rightarrow \text{Sym}_n$ where $n = |\Phi|$. Let W be the subspace spanned by Φ . Then each σ_v , $v \in \Phi$, fixes $W^\perp$ pointwise, whence all elements of G fix $W^\perp$ pointwise. If $\phi \in G$ also fixes each $v \in \Phi$, then ϕ fixes $W + W^\perp = V$, so $\phi = \text{id}$. Thus the kernel of the action of G on Φ is trivial, and so G embeds in Sym_n .

Another pair of conditions that might interest us turn out to be equivalent. If Φ is the root system of a group G generated by reflections, then Φ spans V if and only if G leaves no point of V , except 0, unmoved. ([Grove-Benson, p. 36] call G **effective** when it has this property.) To see this, first suppose Φ spans V and let $w \in V$. If $\sigma_v(w) = w$ for all $v \in \Phi$, then $(w, v) = 0$ for all $v \in \Phi$. Since Φ spans V , this implies $w = 0$.

For the converse, suppose the span of Φ is U , a proper subspace of V . Then there exists a nonzero $w \in U^\perp$, and for this w , we have $\sigma_v(w) = w$ for all $v \in V$. This implies $\phi(w) = w$ for all $\phi \in G$, so G is not effective.

If $U = \{v \in V \mid g(v) = v \text{ for all } g \in G\}$, we can simply regard G as a subgroup of $\mathbf{O}(U^\perp)$ and replace V by $U^\perp$. This makes G effective. Likewise, if Φ spans a proper subspace W of V , we can replace V by W , in which case Φ is a spanning set. Thus we amend our definition and only consider effective groups and (reduced or unreduced) root systems that span V . We will also require finiteness.

Thus we define a concrete **Coxeter group** to be a finite effective subgroup of $\mathbf{O}(V)$ that is generated by reflections. We define an abstract **unreduced root system** in V to be a finite subset Φ not containing 0 such that the following conditions hold:

- (1) Every nonzero multiple of an element of Φ is in Φ .
- (2) Φ spans V .
- (3) If $v, w \in \Phi$, then $\sigma_v(w) \in \Phi$.

An abstract **reduced root system** is a finite subset Φ of V that satisfies (1'), (2), (3):
 (1') If $v \in \Phi$, then $\lambda v \in \Phi$ if and only if $\lambda = \pm 1$.

Here is a nonstandard definition: if Φ satisfies (2),(3), we define the **closure** of Φ to be the set of nonzero multiples of elements of Φ , and we denote it $\bar{\Phi}$. The set $\bar{\Phi}$ is then an abstract unreduced root system.

This is not the definition in [Humphreys, §9, p.42]: he requires that root systems also satisfy the crystallographic condition that we will discuss in Section 5 (The Crystallographic Condition, p.157).

Let us sum up our results.

Proposition 5.1.1 (The root system of a Coxeter group).

Let V be a finite-dimensional vector space over $\mathbb{R}$ with inner product $(,)$. Define concrete Coxeter group and unreduced root system as before the proposition.

- (1) *If G is a concrete finite Coxeter group, then the root system Φ of G is an abstract unreduced root system, and G is generated by the reflections σ_v for $v \in \Phi$.*
- (2) *If Φ is an abstract unreduced root system in V , then the subgroup G of $\mathbf{O}(V)$ generated by the reflections σ_v for $v \in \Phi$ is a concrete Coxeter group, and its root system is Φ .*

Is this true?

- (3) *If Φ is an abstract reduced root system in V , then the subgroup G of $\mathbf{O}(V)$ generated by the reflections σ_v for $v \in \Phi$ is a concrete Coxeter group, and its root system is the closure $\bar{\Phi}$.*

Check (2), (3).

PROOF. (1) This is contained in the discussion prior to the proposition.

(2), (3) We need to prove one more thing for this: the roots of G consist only of (the multiples of) the elements of Φ . **Later.** ■

2. Two-dimensional Root Systems

Notation: We use Dih_n to denote the dihedral group of degree n and order $2n$, that is, the symmetries of a regular n -gon. (This is the notation for the *group* Dih_n ; we will later use the notation Dih_n for a certain Coxeter graph/Dynkin diagram and the corresponding root system.) We say two nonzero vectors v, w are proportional if there is a scalar λ with $v = \lambda w$ and nonproportional otherwise. We say an angle is a multiple of another angle if it is an integral multiple.

When $V = \mathbb{R}^1$, the only finite subgroups of $\mathbf{O}(V)$ are $\{1\}$ and $\{-1, 1\}$. Both are Coxeter groups if we allow the root system to be empty. In the case of $G = \{-1, 1\}$, the unique unreduced root system is $\mathbb{R} \setminus \{0\}$ and reduced root systems have the form $\{-a, a\}$ for an arbitrary positive real number a .

Before we consider the two dimensional case, let us make an observation. Given $v, w \in V$, we can define an unsigned angle θ between them by the equation $\cos \theta = \frac{(v, w)}{\|v\|\|w\|}$, which we usually take to be between 0 and π inclusive. We can then add a sign to the angle to indicate counterclockwise by a positive angle and clockwise by a negative angle. (We have to establish an orientation on the plane v, w span for the positive/negative part.)

Since we are interested in groups generated by reflections, it is useful to summarize some facts about products of reflections. In a finite Coxeter group, all such products have finite order. This places strong limitations on the angles between roots.

Proposition 5.2.1 (Products of two reflections).

Let V be a finite-dimensional vector space over $\mathbb{R}$ with inner product $(,)$. Let $v, w \in V$ be nonzero and let θ be the angle from w to v ; assume $-\pi < \theta \leq \pi$. If v, w are nonproportional, let P be the plane through the origin that they span.

- (1) *If v, w are nonproportional, then $\sigma_v \sigma_w$ leaves $P^\perp$ pointwise fixed and is a rotation by angle 2θ in P .*
- (2) *If the order of $\sigma_v \sigma_w$ is a finite number b , then $\theta = \pm \frac{a}{b}\pi$ with $a, b \in \mathbb{N}$, $\gcd(a, b) = 1$, $0 \leq a \leq b$.*
- (3) *σ_v, σ_w commute if and only if θ is a multiple of $\frac{\pi}{2}$.*

PROOF. (1) This is well-known; we urge the reader to check it.

(2) The order of rotation by 2θ is the smallest positive integer b such that $b \cdot 2\theta = 2\pi a$ for some integer a . This implies $\theta = \pm \frac{a}{b}\pi$ and $\gcd(a, b) = 1$, $0 \leq a \leq b$.

(3) σ_v, σ_w commute if and only if the angles 2θ and -2θ give the same rotation. This happens if and only if 2θ is a multiple of π , and that leads to the claim. ■

Now suppose $V = \mathbb{R}^2$ and Φ is the root system of a(n effective) Coxeter group G in $\mathbf{O}(V)$. It is well-known that the only finite subgroups of $\mathbf{O}(V)$ are the dihedral groups Dih_n of symmetries of a regular n -gon and their subgroups Cyc_n of rotational symmetries. The group Cyc_n contains no reflections, so it is not a Coxeter group (except in a trivial case mentioned below).

We say $\text{Cyc}_1 = \text{Dih}_1$ is the trivial symmetry group. It can be regarded as the Coxeter group with empty root system, but it's not effective.

We say Dih_2 is the symmetry group of a line segment, say $[-1, 1]$ (or the set $\{-1, 1\}$), so it consists of I , rotation by π , and reflections about the x -axis and the y -axis. The roots are $a(1, 0)$ and $b(0, 1)$ and $\{(1, 0), (0, 1)\}$ is a reduced root system.

Suppose now $n \geq 3$ and, when convenient, identify $\mathbb{R}^2$ with $\mathbb{C}$. If our n -gon has the x -axis as a central axis, then its roots described as complex numbers have the form $a_k e^{k\pi i/n}$ for $0 \leq k \leq n-1$ and real numbers a_k . What restrictions must we put on the a_k to obtain a reduced root system?

We know that G permutes the roots, in either a reduced or unreduced root system. Let $\tau_\phi \in G$ represent rotation by ϕ , so $\phi = \frac{2\pi j}{n}$ for some j . If we start with the root $a e^{k\pi i/n}$, applying the various τ_ϕ gives us roots $a e^{(k+2j)\pi i/n}$ as j ranges over $0, \dots, n-1$. These images are all distinct and so they give us n roots. Thus the orbit of a given root under the action of all of G has size either n or $2n$.

Since $k + 2j \equiv k \pmod{2}$, the roots we obtain from rotations are those of the form $a e^{\ell\pi i/n}$ with ℓ of the same parity as k . If n is even, then k and $k \pm n$ have the same parity, so $a e^{(k+n)\pi i/n} = -a e^{k\pi i/n}$ is in the same orbit as $a e^{k\pi i/n}$.

Now we consider the action of the reflections on our root $a e^{k\pi i/n}$. Certainly $-a e^{k\pi i/n}$ is obtained from $a e^{k\pi i/n}$ via the reflection with root $a e^{k\pi i/n}$. This means that if n is even, $a e^{k\pi i/n}$ cannot have an orbit of size $2n$. Thus when n is even, the action of G on the roots has two orbits. In particular, we can choose, if we wish, two different lengths for our roots in a reduced root system.

When n is odd, one can show the size of the orbit is $2n$, i.e., G acts transitively on the roots. This means that there can be only one root length in a reduced root system.

Details!

If we take the roots 1 and $e^{(n-1)\pi i/n}$, then every root is a linear combination of these two with either all nonnegative coefficients (those in the upper half plane) or all nonpositive coefficients (those in the lower half plane plus -1). **Details.**

3. Positive Roots and Bases

At the moment we are mainly concerned about whether the angle between v and w is acute, which equates to $(v, w) > 0$, or obtuse, which equates to $(v, w) < 0$. The following result says that if some vectors have pairwise obtuse angles and all lie on the same side of a hyperplane, those vectors are linearly independent.

Suppose W is a hyperplane in V and $x \in W^\perp$ is nonzero. Then every $v \in V \setminus W$ satisfies either $(v, x) < 0$ or $(v, x) > 0$. The set of v with $(v, x) < 0$ will be denoted H_- and the set of v with $(v, x) > 0$ will be denoted H_+ . We call H_- and H_+ the two **sides** of H . Note that 0 is on every hyperplane and hence is not on any side of a hyperplane. We also see that if $v, w \in V$ are nonzero, there is a hyperplane such that v, w lie on opposite sides if and only if w is not a positive multiple of v .

Proposition 5.3.1 (Obtuse angles implies linearly independent).

Let V be a finite-dimensional vector space over $\mathbb{R}$ and let $(,)$ be an inner product on V . Suppose $v_1, \dots, v_k \in V$ all lie on the same side of some hyperplane. If $(v_i, v_j) \leq 0$ for all $i \neq j$, then $v_1, \dots, v_k$ are linearly independent.

REFERENCES AND REMARKS. See [Humphreys, §10.1, proof of Theorem' (3), p. 48; Grove-Benson, Proposition 4.1.6, p. 39]. **Other references?**

This statement is false without the “same side” hypothesis. For a trivial example, take any nonzero v_1 and set $v_2 = -v_1$. For another example, consider $V = \mathbb{C} = \mathbb{R}^2$, let $\omega = e^{2\pi i/3}$ be a primitive root of unity and consider $1, \omega, \omega^2 = \bar{\omega}$. These vectors have angles of 120° between them but they are not linearly independent (e.g., they add up to 0).

A related result is the following: if S is a finite set of vectors and $(v, w) < 0$ whenever $v, w \in S, v \neq w$, then any proper subset of S is linearly independent. We cannot weaken the condition to $(v, w) \leq 0$; as an example, take S to be the set of unit vectors in any $\mathbb{R}^n$.

PROOF. If $v_1, \dots, v_k$ are linearly dependent, then some nontrivial linear combination is 0 . If we renumber the v_i and move the negative coefficients to the other side in the corresponding equation, we get a relation $\sum_{i=1}^r \lambda_i v_i = \sum_{i=r+1}^k \lambda_i v_i$ for some r and some nonnegative scalars $\lambda_1, \dots, \lambda_k$. (We can have $r = 0$ or $r = k$, i.e., one sum could be void.) Taking the norm-squared of the left hand side, we have

$$0 \leq \left\| \sum_{i=1}^r \lambda_i v_i \right\|^2 = \left(\sum_{i=1}^r \lambda_i v_i, \sum_{i=r+1}^k \lambda_i v_i \right) = \sum_{\substack{1 \leq i \leq r \\ r+1 \leq j \leq k}} \lambda_i \lambda_j (v_i, v_j).$$

All of the inner products are nonpositive and all of the scalars are nonnegative, so the final sum is nonpositive, but at least 0 . This forces the final sum to be 0 , so $\sum_{i=1}^r \lambda_i v_i = 0$.

Now suppose the points all lie on the same side of the hyperplane $x^\perp$, so that (x, v_i) has the same sign for all i . Replacing x by $-x$ if necessary, we can assume (x, v_i) is always positive. If we take the inner product of x with the equation $\sum_{i=1}^r \lambda_i v_i = 0$, we see $\sum_{i=1}^r \lambda_i (x, v_i) = 0$. Since each $(x, v_i) > 0$, this forces $\lambda_i = 0$ for $1 \leq i \leq r$. The same argument shows $\lambda_i = 0$ for $r + 1 \leq i \leq k$. This proves $v_1, \dots, v_k$ are linearly independent. ■

We will divide roots into two types: positive and negative. For that, we need a notion of positive root. There are several definitions of positive. We will choose one below, but let us list the requirements for such a division. We require three things.

- (1) If the sum of two positive roots is a root, then it is positive.
- (2) If a positive multiple of a positive root is a root, then it positive.
- (3) If α is a root, then exactly one of $\alpha, -\alpha$ is positive.

For reduced root systems, condition (2) is of course redundant and for unreduced root systems, (2) becomes “a positive multiple of a positive root is a positive root”. In either case, (1) and (2) imply that any linear combination of positive roots with nonnegative coefficients (not all 0) is a positive root if it is a root. Moreover, such a combination cannot be 0 unless all coefficients are 0. To see this, note that if such a combination existed with some coefficient positive, there would be positive roots $\alpha, \beta_1, \dots, \beta_k$ with $-\alpha = \sum_{i=1}^k \lambda_i \beta_i$ with each $\lambda_i \geq 0$. By the observation above, this would imply $-\alpha$ is a positive root, violating (3).

We can define a partial order on a reduced root system by defining $\alpha > \beta$ if $\alpha - \beta$ is a positive root. This will rarely be a total order.

One way to define positive roots is to define a partial order on $\mathbb{R}^n$ that satisfies conditions (1) and (2) for its positive elements and then restrict the order to roots.

Check. We can do this, for example, by using a lexicographic order on $\mathbb{R}^n$, which in fact is a total order.

We will use another means to define positive roots. We define an element $x \in V$ to be **regular** with respect to the root system Φ if $(x, v) \neq 0$ for any $v \in \Phi$. Thus an element is regular if it is not in $\cup_{v \in \Phi} v^\perp$. This union is a finite union of hyperplanes, and it is well-known that a finite union of hyperplanes cannot equal a nonzero vector space over a field of characteristic 0. Thus regular elements always exist. (In fact, most elements of V are regular.) Fix a regular element $t \in V$ and let $H = t^\perp$ be the hyperplane perpendicular to t . Then we define $\Phi_t^+ = H_+ \cap \Phi$ and we call elements of Φ_t^+ **t -positive roots**. When t is understood, we refer simply to Φ^+ and positive roots. We can likewise define Φ_t^- and negative roots. Since t is regular, $\Phi = \Phi^+ \cup \Phi^-$, and it is obvious that $\Phi^- = -\Phi^+$.

Warning: The division into positive and negative roots definitely depends on t . As a simple example, if we replace t by $-t$, we switch the positive and negative roots. More subtle changes are also possible. This warning applies to definitions we will make below, such as the definition of a base.

A subset Δ of Φ^+ is called a **t -base** of Φ if it is linearly independent and every element of Φ^+ can be written as a linear combination of elements of Δ with nonnegative coefficients. We will drop t when it is understood. Elements of a t -base are called **t -simple roots** or just **simple roots**. (Some sources call a base a **simple system**.) Since Φ spans V , any base is a basis of V . The reflections σ_v for $v \in \Delta$ are called **simple reflections**. Some sources refer to **fundamental roots** and **fundamental reflections**.

Proposition 5.3.2 (Existence and uniqueness of bases).

Let V be a finite-dimensional vector space over $\mathbb{R}$ and let $(,)$ be an inner product on V . Let Φ be a (reduced or unreduced) root system in V and let $t \in V$ be regular. Then there is a unique t -base Δ of V . Moreover, $(v, w) \leq 0$ for any distinct $v, w \in \Delta$.

Uniqueness in the unreduced case means that if Δ' is another t -base, its elements are positive scalar multiples of the elements of Δ .

REFERENCES AND REMARKS. See [Humphreys, §10.1, Theorem, pp. 48-51; Jacobson, §IV.3, pp. 131-133; Bourbaki 1-3, §VI.1.5-6, pp. 165-169; Erdmann-Wildon, Theorem 11.10, p. 116; Grove-Benson, §4.1, Theorem 4.17 & Proposition 4.1.8, pp. 37-40].

PROOF. We prove existence below by showing that a minimal nonnegative spanning set is a base.

Uniqueness: **Later.** ■

We make a nonstandard definition for temporary use. We say a subset S of Φ^+ is **nonnegatively spanning** if every element of Φ^+ is a linear combination of elements of S with nonnegative coefficients. Such subsets exist since Φ^+ itself is nonnegatively spanning. It's easy to see that if S is a nonnegatively spanning set, then every negative element of Φ is a linear combination of elements of S with nonpositive coefficients. Since Φ^+ is finite, there exist minimal nonnegatively spanning sets. Clearly a base is a minimal nonnegatively spanning set; we will prove any such set is a base for Φ .

We begin by proving a few properties of minimal nonnegatively spanning sets.

Proposition 5.3.3 (Mixed combination of base roots is not a root).

Let V be a finite-dimensional vector space over $\mathbb{R}$ and let $(,)$ be an inner product on V . Let Φ be an abstract root system in V with a set Φ^+ of positive roots and let Δ be a minimal nonnegatively spanning subset of Φ^+ . If $v, w \in \Delta$ and λ, μ are positive real numbers, then $\lambda v - \mu w$ is not a root (not in Φ).

REFERENCES AND REMARKS. See [Humphreys, §10.1, Lemma, p. 47; Jacobson, §IV.3, XV(i), p. 120; Grove-Benson, Proposition 4.1.4, p. 38].

PROOF. If $\lambda v - \mu w$ is a root, it is either positive or negative. Suppose first $\lambda v - \mu w$ is positive. Then we have $\lambda v - \mu w = av + bw + \sum_{\substack{u \in \Delta \\ u \neq v, w}} c_u u$ for nonnegative a, b, c_u . If $\lambda > a$, set $d = (\lambda - a)^{-1}$. Then we have $v = d(\mu + b)w + \sum_{\substack{u \in \Delta \\ u \neq v, w}} dc_u u$. Since $\mu + b > 0$, this violates minimality of Δ (since we don't need v).

If $\lambda \leq a$, we can write $0 = (a - \lambda)v + (b + \mu)w + \sum_{\substack{u \in \Delta \\ u \neq v, w}} c_u u$. Since all the coefficients on the right hand side are nonnegative and $b + \mu > 0$, this is impossible (the right hand side is positive). Thus we get a contradiction in either case.

If $\lambda v - \mu w$ is negative, then $\mu w - \lambda v$ is positive and we can apply the above argument to show this is impossible. ■

Proposition 5.3.4 (Base root angles are obtuse).

Let V be a finite-dimensional vector space over $\mathbb{R}$ and let $(,)$ be an inner product on V . Let Φ be an abstract root system in V with a set Φ^+ of positive roots and let Δ be a minimal nonnegatively spanning subset of Φ^+ . If $v, w \in \Delta$, $w \neq v$, then the following statements hold.

- (1) $\sigma_v(w) \in \Phi^+$.
- (2) $(v, w) \leq 0$.
- (3) $(v, w) = -\|v\|\|w\| \cos \frac{\pi}{m(v, w)}$, where $m(v, w)$ is the order of $\sigma_v \sigma_w$.

REFERENCES AND REMARKS. See [Humphreys, §10.1, Lemma, p. 47; §10.2, Lemma B, p. 50; Jacobson, §IV.3, XV(ii), p. 120; Bourbaki 4-6, §VI.5.1, Remark 3, p. 166; Erdmann-Wildon, §11.3, Exercise 11.3, p. 115; Grove-Benson, Proposition 4.1.5, p. 38].

A stronger statement than (1) holds: see Proposition 5.3.6 (Simple reflections almost preserve positive roots, p. 151).

PROOF. (1), (2) We know $w - \frac{2(w, v)}{(v, v)}v = \sigma_v(w) \in \Phi$ is a root. By Proposition 5.3.3 (Mixed combination of base roots is not a root, p. 149), this can only happen if $\frac{2(w, v)}{(v, v)}$ is not positive, i.e., if $(w, v) \leq 0$. In this case, $\sigma_v(w) \in \Phi$ is a positive root.

(3) **Details.** ■

Proposition 5.3.5 (Minimal nonnegatively spanning set is a base).

Let V be a finite-dimensional vector space over $\mathbb{R}$ and let $(,)$ be an inner product on V . Let Φ be an abstract root system in V with a set Φ^+ of positive roots and let Δ be a minimal nonnegatively spanning subset of Φ^+ . Then Δ is a base.

REFERENCES AND REMARKS. See [Grove-Benson, Theorem 4.1.7, p. 39].

PROOF. By definition, all elements of a base lie on the same side of a hyperplane. Therefore, by Proposition 5.3.4 (Base root angles are obtuse, p. 150) and Proposition 5.3.1 (Obtuse angles implies linearly independent, p. 147), Δ is linearly independent. Therefore Δ is a base. ■

Proposition 5.3.6 (Simple reflections almost preserve positive roots).

Let V be a finite-dimensional vector space over $\mathbb{R}$ and let $(,)$ be an inner product on V . Let Φ be an abstract root system in V with a set Φ^+ of positive roots and base Δ . If $v \in \Delta$, then $\sigma_v(\Phi^+ \setminus \{v\}) = \Phi^+ \setminus \{v\}$.

REFERENCES AND REMARKS. See [Humphreys:Coxeter, §1.4, Proposition, p. 10].

PROOF. Let $w \in \Phi^+$, $w \neq v$, and write $w = \sum_{u \in \Delta} \lambda_u u$, where each $\lambda_u \geq 0$. Since $w \neq v$, we cannot have $\lambda_u = 0$ for all $u \neq v$. Thus there is a u' with $\lambda_{u'} > 0$. Now $\sigma_v(w) = w - \mu v = (\lambda_v - \mu)v + \sum_{u \in \Delta, u \neq v} \lambda_u u$ for some scalar μ . If $\sigma_v(w)$ is negative, then $-\sigma_v(w) = (-\lambda_v + \mu)v + \sum_{u \in \Delta, u \neq v} -\lambda_u u$ is positive. But this expression for $-\sigma_v(w)$ has a negative coefficient for u' , contradicting the fact that Δ is nonnegatively spanning (and in fact, a basis by Proposition 5.3.5 (Minimal nonnegatively spanning set is a base, p. 150)). Thus we must have $\sigma_v(w)$ positive. If $\sigma_v(w) = v$, then we must have $w = -v$, which would make w negative. Thus $\sigma_v(w) \neq v$.

We've shown $\sigma_v(\Phi^+ \setminus \{v\}) \subseteq \Phi^+ \setminus \{v\}$. Since σ_v is a bijection and Φ is finite, this implies equality holds. ■

Proposition 5.3.7 (Conjugacy of positive roots and bases).

Let V be a finite-dimensional vector space over $\mathbb{R}$ and let $(,)$ be an inner product on V . Let Φ be an abstract root system in V and let G be the corresponding Coxeter group.

- (1) *If Φ^+ and Φ^{++} are sets of positive roots in Φ , then there is a $\sigma \in G$ with $\sigma(\Phi^+) = \Phi^{++}$.*
- (2) *If Δ and Δ' are bases of Φ , then there is a $\sigma \in G$ with $\sigma(\Delta) = \Delta'$.*

REFERENCES AND REMARKS. See [Humphreys:Coxeter, §1.4, Theorem, p. 10].

PROOF. **Later.** ■

4. Coxeter Graphs and Matrices

An **abstract Coxeter graph** is a finite unoriented graph whose edges are labelled with integers that are at least 3. When an edge has a label of 3, we omit the number when drawing the graph. This is a slightly non-standard definition in that most sources allow real numbers that are greater than 2. Let G be a finite Coxeter group with root system Φ

and base Δ and let $m(v, w)$ be the order of the product $\sigma_v \sigma_w$. We define the **Coxeter graph** or **Coxeter diagram** of G (really of Δ) to be a graph whose vertices (also called nodes) are the elements of Δ , with edges between v and w whenever $m(v, w) > 2$. We label the edge between v and w with $m(v, w)$; we only draw the label if $m(v, w) > 3$. This means that two simple roots are connected in the Coxeter graph if and only if they are not perpendicular, and two roots have an edge with no number if and only if the angle between them is 120° .

Warning: [Humphreys, §11.2, p. 56] gives a different definition of Coxeter graph, one that more closely resembles a Dynkin diagram. In place of an edge labelled by a number $r > 3$, he use $r/2$ edges to join two vertices. Since he only deals with crystallographic Coxeter groups, the values of r that appear for him are 4 and 6, so $r/2$ is always an integer.

Let Γ be a Coxeter graph with set of vertices Δ and labels $m(v, w)$ for the edge between v and w (and set $m(v, v) = 1$, $m(v, w) = 2$ if there is no edge between v and w , and $m(v, w) = 3$ if the edge is unlabelled) We define the **Schläfli matrix** of Γ to be the $\Delta \times \Delta$ matrix whose v, w entry is $-\cos \frac{\pi}{m(v, w)}$. Proposition 5.3.4 (Base root angles are obtuse, p. 150)(3) tells us that this entry is just $\frac{(v, w)}{\|v\|\|w\|}$.

When $v_1, \dots, v_k$ are vectors in an inner product space, the **Gram matrix** of these vectors is the $k \times k$ matrix whose i, j entry is (v_i, v_j) . When our root vectors have norm 1, we see that the Schläfli matrix of Δ is the same as the Gram matrix of Δ .

Proposition 5.4.1 (Coxeter graphs are positive definite).

Let V be a finite-dimensional vector space over $\mathbb{R}$ with inner product $(,)$ and let G be a finite Coxeter group in $\mathbf{O}(V)$ with root system Φ , base Δ , and Schläfli matrix A . Then A is positive definite.

REFERENCES AND REMARKS. See [Grove-Benson, Theorem 5.1.3, p. 56].

Proposition 5.4.4 (Irreducible Coxeter graphs, p. 154) and its proof tell us that in fact a Schläfli matrix is the matrix of a Coxeter group if and only if it is positive definite.

Check.

PROOF. Let $\Delta = \{v_1, \dots, v_r\}$, set $m(i, j) = m(v_i, v_j)$. As noted above, A is the matrix with i, j entry $-\cos(\pi/m(i, j)) = \frac{(v_i, v_j)}{\|v_i\|\|v_j\|}$. Since A is symmetric, there is a symmetric bilinear $\langle, \rangle$ form associated to A , given by $\langle x, y \rangle = \sum_{i,j} a_{ij} x_i y_j$, where $x = \sum_i x_i v_i$, $y = \sum_j y_j v_j$. Set $w = \sum_i x_i v_i / \|v_i\|$. If $x \neq 0$, then $w \neq 0$ and

$$\langle x, x \rangle = \sum_{i,j} a_{ij} x_i x_j = \sum_{i,j} \frac{(v_i, v_j)}{\|v_i\|\|v_j\|} x_i x_j = \left(\sum_i x_i v_i / \|v_i\|, \sum_j x_j v_j / \|v_j\| \right) = (w, w) = \|w\|^2 > 0.$$

This proves $\langle, \rangle$ is an inner product and so A is positive definite. ■

The following technical result is also useful. If Γ is a Coxeter graph, we say Γ' is a **subgraph** of Γ if Γ' contains a subset of the vertices of Γ and some of the edges between those vertices, and if the labels on the edges in Γ' are less than or equal to the labels on the same edges on Γ . Clearly Γ' is another Coxeter graph.

Proposition 5.4.2 (Subgraphs are positive definite).

Suppose Γ is a positive definite Coxeter graph and Γ' is a nonempty subgraph of Γ . Then Γ' is positive definite.

REFERENCES AND REMARKS. See [Grove-Benson, Proposition 5.1.6, p. 61].

PROOF. Suppose the vertices of Γ are $v_1, \dots, v_n$ and those of Γ' are $v_1, \dots, v_k$, and regard them as a basis of some V . Let the label on the edge between v_i and v_j be m_{ij} and the label on the edge between v'_i and v'_j be m'_{ij} . If an edge is missing from Γ' , we set $m'_{ij} = 2$. Let A, A' be the Schläfli matrices of A, A' respectively, so $a_{ij} = -\cos \frac{\pi}{m_{ij}} \leq -\cos \frac{\pi}{m'_{ij}} = a'_{ij} \leq 0$. If Γ' is not positive definite, there is a nonzero $x = \sum_{i=1}^k \lambda_i v_i$ with $\sum_{1 \leq i, j \leq k} \lambda_i \lambda_j a'_{ij} < 0$. Set $\mu_i = |\lambda_i|$ for $1 \leq i \leq k$ and $\mu_i = 0$ for $k+1 \leq i \leq n$, and set $y = \sum_{i=1}^n \mu_i v_i$. Then

$$\sum_{1 \leq i, j \leq n} \mu_i \mu_j a_{ij} = \sum_{1 \leq i, j \leq k} |\lambda_i| |\lambda_j| a_{ij} \leq \sum_{1 \leq i, j \leq k} \lambda_i \lambda_j a'_{ij} < 0.$$

This shows Γ is not positive definite. ■

This result is used in the contrapositive: if we suspect a Coxeter graph is not positive definite, it is enough to find a subgraph that is not positive definite.

We say a root system Φ is **reducible** if it is the union of two proper subsets that are mutually orthogonal, that is, $\Phi = \Phi_1 \cup \Phi_2$ with Φ_1, Φ_2 nonempty and $(v, w) = 0$ for all $v \in \Phi_1, w \in \Phi_2$. We say a root system is **irreducible** if it is nonempty and not reducible. We make the same definition for bases of root systems. More generally, any root system or base can be written uniquely as a union of finitely many mutually orthogonal irreducible subsets. We call these the **components** of the root system or base.

We say a nonempty graph is **connected** if there is a path of edges between any two vertices. Any finite graph is a unique union of maximal connected subgraphs, called its **components**.

We say a Coxeter group G is **irreducible** if we cannot write G as an internal direct sum of two nontrivial Coxeter subgroups (i.e., subgroups that are generated by reflections). This is different from G being indecomposable. The dihedral group Dih_6 , symmetries of a regular hexagon, is decomposable as HN where H is generated by the rotation by 180° and N is the group of symmetries of the triangle formed by alternating vertices of the hexagon. However, Dih_6 is irreducible as a Coxeter group, since none of its root pairs are orthogonal.

Proposition 5.4.3 (Connected equivalent to irreducible).

Let V be a finite-dimensional vector space over $\mathbb{R}$ with inner product $(,)$ and let G be a finite Coxeter group in $\mathbf{O}(V)$ with root system Φ , base Δ , and Coxeter graph Γ . Then the following statements are equivalent.

- (1) Φ is irreducible.
- (2) Δ is irreducible.
- (3) Γ is connected.
- (4) G is irreducible.

REFERENCES AND REMARKS. See [Humphreys, §10.4, p. 52; Erdmann-Wildon, Lemma 11.8, p. 114; Exercise 11.7, p. 121; Grove-Benson, Proposition 5.1.4, p. 57].

PROOF. **Later.** ■

Proposition 5.4.3 (Connected equivalent to irreducible, p. 154) tells us that to classify (finite) Coxeter groups or root systems, it is enough to classify the irreducible ones. In the next result we list the irreducible root systems, and we describe the Coxeter groups they generate, both as a symmetry group and an abstract group, when possible. In describing the abstract groups, we use the following notation. We let Dih_n denote the dihedral group of order $2n$, i.e., the symmetries of a regular n -gon. We let Sym_n denote the symmetric group on n letters and Alt_n denote the alternating group on n letters.

We let $\text{Sym}_n^\pm$ denote the group of signed permutations. The group $\text{Sym}_n^\pm$ can be realized in several ways. One is as signed permutation matrices, i.e., matrices with exactly one nonzero entry in each row and column, with that entry being ± 1 . A second description is as the permutations σ of the set $\{\pm 1, \pm 2, \dots, \pm n\}$ with the property that $\sigma(-i) = -\sigma(i)$. A third description is as the set of functions $\sigma : \{1, 2, \dots, n\} \rightarrow \{\pm 1, \pm 2, \dots, \pm n\}$ with the property that $i \neq j$ implies $|\sigma(i)| \neq |\sigma(j)|$. The group $\text{Sym}_n^\pm$ is isomorphic to the semidirect product $(\mathbb{Z}_2)^n \rtimes \text{Sym}_n$ and has order $2^n n!$. We also call $\text{Sym}_n^\pm$ the **hyperoctahedral group**; it is the symmetry group of the n -cube. We need an index 2 subgroup $W(D_n)$ of $\text{Sym}_n^\pm$ as well, consisting of either signed permutation matrices with an even number of -1 's or signed permutations σ such that $\sigma(i)$ is negative for an even number of $i = 1, \dots, n$.

Proposition 5.4.4 (Irreducible Coxeter graphs).

The following is a complete list of Coxeter graphs of irreducible Coxeter groups. We use the notation defined prior to the proposition for the description of the Coxeter groups

generated. (There is some redundancy: see the comments.)

Type	Coxeter graph	Coxeter group	
		Symmetry group of	Abstract group
A_ℓ		regular ℓ -simplex	$\text{Sym}_{\ell+1}$
B_ℓ		ℓ -cube	$\text{Sym}_\ell^\pm$
D_ℓ		ℓ -demicube	$W(D_\ell)$
E_6			$\text{SO}(5, \mathbb{F}_3)$
E_7			order $2903040 = 2^{10}3^45^17^1$
E_8			order $696729600 = 2^{14}3^55^27^1$
F_4		regular 24-cell in 4D	order $1152 = 2^73^2$
G_2		regular hexagon	Dih_6
H_2		regular pentagon	Dih_5
H_3		regular dodecahedron or icosahedron	$\text{Alt}_5 \times \mathbb{Z}_2$
H_4		regular 120-cell or 600-cell in 4D	order $14400 = 2^63^25^2$
$I_2(n)$		regular n -gon in 2D	Dih_n

REFERENCES AND REMARKS. See [Bourbaki 4-6, §VI.4.1, Theorem 1, pp. 205-206; Grove-Benson, Figure 5.3, p. 57, Theorem 5.1.7, p. 62; Wikipedia].

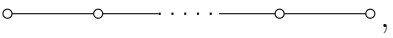
Some of these graphs are the same for different types. For example, H_2 and $I_2(5)$ are the same, as are G_2 and $I_2(6)$, and as are A_3 and D_3 .

To avoid the above equalities and some degenerate or undefined cases, we usually require: $\ell \geq 2$ for B_ℓ ; $\ell \geq 4$ for D_ℓ ; $n = 5$ or $n \geq 7$ for $I_2(n)$.

Warning: The letters for Coxeter graphs may vary in some sources. In particular, what we denote $I_2(n)$ is denoted H_2^n in [Grove-Benson]. Likewise, [Grove-Benson]

denotes by I_k what we have denoted H_k for $k = 3, 4$. Some sources also denote our B_ℓ by BC_ℓ or $B_\ell = C_\ell$.

PROOF. The proof is in three parts. The first is showing the diagrams above have positive definite Schläfli matrices. The second is showing that no other Coxeter graphs have positive definite Schläfli matrices. The third is showing that Coxeter groups/root systems exist with each of the given Coxeter graphs.

We give an example of the first and third step. The graph for A_n , , has Schläfli matrix
$$\begin{pmatrix} 1 & -\frac{1}{2} & \cdots & 0 \\ -\frac{1}{2} & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$
 Call this matrix A_n . Then it's easy to see that

$\det A_n = \det A_{n-1} - \frac{1}{4} \det A_{n-2}$ and $\det A_1 = 1$, $\det A_2 = \frac{3}{4}$. This leads to the general formula $\det A_n = \frac{n+1}{2^n}$. The principal minors for A_n are the $\det A_i$ for $1 \leq i \leq n$, and so A_n is positive definite.

Inspired by the roots of the Lie algebra $\mathfrak{sl}(n+1)$, we consider the root system consisting of the vectors $v_{ij} = \mathbf{e}_i - \mathbf{e}_j \in \mathbb{R}^{n+1}$ for $1 \leq i, j \leq n$, $i \neq j$. The positive roots are those v_{ij} with $i < j$. A base for this root system consists of the vectors $v_i = v_{i,i+1}$ for $1 \leq i \leq n$. This root system has Coxeter graph A_n : the simple roots are perpendicular, except successive simple roots have angles of 120° .

We give an example of the second step: we show that no positive definite Coxeter graph can contain a (labelled) cycle. If a positive definite graph contained a labelled cycle, Proposition 5.4.2 (Subgraphs are positive definite, p. 153) implies that a cycle (with all labels 3) would be positive definite. It is not hard to show that the columns (or rows) of the Schläfli matrix of a cycle add up to 0 and so the matrix has determinant 0. Thus the matrix cannot be positive definite. For example, if Γ is a triangle , then

the Schläfli matrix is
$$\begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & 1 \end{pmatrix}.$$
 If Γ is a square , then the Schläfli matrix is

$$\begin{pmatrix} 1 & -\frac{1}{2} & 0 & -\frac{1}{2} \\ -\frac{1}{2} & 1 & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{2} & 0 & -\frac{1}{2} & 1 \end{pmatrix}.$$
 Other Coxeter graphs are handled by reducing to some minimal

set of subgraphs and then showing the matrices from that minimal set are not positive definite.

More details?

■

5. The Crystallographic Condition

A **lattice** in a vector space V is the $\mathbb{Z}$ -linear span of a basis. The canonical example is $\mathbb{Z}^n \subseteq \mathbb{R}^n$, but the lattice need not be right-angled or with sides of even length. Another example is $\{(n\pi, (n-m)e) \mid n, m \in \mathbb{Z}\} \subseteq \mathbb{R}^2$, which is the $\mathbb{Z}$ -linear span of (π, e) and $(0, -e)$. A subgroup G of $\mathbf{O}(V)$ that leaves a lattice in V invariant is said to be a **crystallographic group** or to satisfy the **crystallographic condition** (also called the **crystallographic restriction**). We call a finite crystallographic Coxeter group a **Weyl group**. When studying Lie algebras and their root systems, we usually prefer the term “Weyl group” to “Coxeter group”.

Proposition 5.5.1 (Crystallographic condition for root systems).

Let V be a finite-dimensional vector space over $\mathbb{R}$ with inner product $(\cdot, \cdot)$. Let G be a concrete Coxeter group with reduced root system Φ and base Δ . Then the following conditions are equivalent.

- (1) G is a crystallographic group.
- (2) For every $v, w \in \Phi$, the scalar $\frac{2(v, w)}{(v, v)}$ is an integer.
- (3) The only angles between distinct simple roots are $90^\circ, 120^\circ, 135^\circ, 150^\circ$.

REFERENCES AND REMARKS. See [Humphreys, §III.9.2, p. 42; Bourbaki 4-6, §VI.2.5, Remark 2, pp. 193-194; Grove-Benson, Proposition 5.2.1, p. 63; Wikipedia].

We could add a fourth condition: the crystallographic condition is equivalent to the condition that the only angles between nonproportional roots are $30^\circ, 45^\circ, 60^\circ, 90^\circ, 120^\circ, 135^\circ, 150^\circ$. For example, among symmetry groups of regular polygons, only those of the triangle, the square, and the hexagon appear.

PROOF. (1) $\Rightarrow$ (3) Suppose G fixes a lattice L and let $v, w \in \Delta$. Set $m = m(v, w)$. With respect to an appropriate basis, the matrix of $\sigma_v \sigma_w$ is

$$\begin{pmatrix} \cos \frac{2\pi}{m} & -\sin \frac{2\pi}{m} & 0 \\ \sin \frac{2\pi}{m} & \cos \frac{2\pi}{m} & 0 \\ 0 & 0 & I \end{pmatrix},$$

where I is the $(n-2) \times (n-2)$ identity matrix. Thus the trace of $\sigma_v \sigma_w$ is $2 \cos \frac{2\pi}{m} + n - 2$. With respect to the lattice basis of L , the matrix of $\sigma_v \sigma_w$ has integer entries, so the trace is an integer. Thus $\cos \frac{2\pi}{m}$ is an integer or half-integer. This can only happen if $m = 1, 2, 3, 4, 6$. The angle between v and w is $(1 - \frac{1}{m})\pi$, and so (3) follows.

$-\cos \frac{\pi}{m}$, and so (3) follows. **Check.**

More later. ■

Because of this proposition, we say a reduced root system Φ satisfies the **crystallographic condition** or **integrality condition** if $\frac{2(v, w)}{(v, v)}$ is an integer for every $v, w \in \Phi$. We frequently shorten this to “ Φ is crystallographic”.

Proposition 5.5.2 (Base for crystallographic root systems).

Let V be a finite-dimensional vector space over $\mathbb{R}$ with inner product $(,)$. Let G be a concrete Coxeter group with reduced root system Φ satisfying the crystallographic condition and base Δ .

- (1) *Every element of Φ^+ can be written as a linear combination of elements of Δ with nonnegative integer coefficients.*
- (2) *A positive root is in Δ if and only if it cannot be written as the sum of two positive roots.*

REFERENCES AND REMARKS. The condition in (2) is used to define a base in many sources. See [Humphreys, §10.1, pp.47-48].

PROOF. **Later.** ■

Proposition 5.5.3 (Irreducible crystallographic Coxeter graphs).

The following is a complete list of Coxeter graphs of irreducible crystallographic Coxeter groups (i.e., Weyl groups): $A_\ell, B_\ell, D_\ell, E_6, E_7, E_8, F_4, G_2$.

REFERENCES AND REMARKS. See [Humphreys, §11.4, Theorem, pp.57-58; Jacobson, §IV.5, Theorem 5, p.134; Bourbaki 1-3; Erdmann-Wildon, Theorem 13.1, p.142; Bourbaki 4-6, §VI.4.2, Theorem 3, p.209; Grove-Benson, Theorem 5.2.2, p.64; Wikipedia].

To avoid equalities and some degenerate or undefined cases, we usually require: $\ell \geq 2$ for B_ℓ and $\ell \geq 4$ for D_ℓ .

When we discuss simple Lie algebras, we will have the same classification, but we will use Dynkin diagrams instead of Coxeter graphs.

PROOF. **Later maybe.** ■

CHAPTER 6

Classification of Semisimple Lie Algebras

This chapter is just a sketch. In Section 1 (Objects associated to Semisimple Lie Algebras, p. 159), we show how to turn Coxeter graphs into Dynkin diagrams. These diagrams are what is used in the classification of semisimple Lie algebras. We also discuss Cartan matrices. In Section 2 (Classification of Split Semisimple Lie Algebras, p. 164), we show that split simple Lie algebras correspond bijectively to connected Dynkin diagrams. In particular the split simple Lie algebras are precisely the classical Lie algebras $\mathfrak{sl}(n)$, $\mathfrak{so}(n)$, $\mathfrak{sp}(n)$ and the five exceptional Lie algebras E_6, E_7, E_8, F_4, G_2 . Since semisimple Lie algebras are direct sums of simple Lie algebras and Dynkin diagrams are unions of connected Dynkin diagrams, this establishes a bijection between split semisimple Lie algebras and Dynkin diagrams.

Terminology: Remember that what we call a maximal toral subalgebra, most sources call a Cartan subalgebra.

1. Objects associated to Semisimple Lie Algebras

If L is a semisimple Lie algebra with split maximal toral subalgebra H and set of roots Φ , then Proposition 4.2.5 (Properties of the root space decomposition, p. 135) and Proposition 4.2.6 (Properties of the dual Killing form, p. 138) tell us that Φ forms a reduced crystallographic root system. We wish to show this root system completely determines the Lie algebra. Proposition 5.5.3 (Irreducible crystallographic Coxeter graphs, p. 158) classifies irreducible crystallographic root systems in terms of Coxeter graphs. Using this, we will obtain a complete classification of split semisimple Lie algebras.

Semisimple Lie algebras are typically categorized in terms of Dynkin diagrams or Cartan matrices rather than Coxeter graphs. A Dynkin diagram is another way of writing a Coxeter graph, with a little more information. We only use Dynkin diagrams for reduced Crystallographic root systems. Before we describe the connection between the two types of diagrams, we need to make one observation. (See [Fulton-Harris, §21.1, Table 21.4, p. 321] for more.) Suppose α, β are roots and let θ be the angle between them. If $\theta = 90^\circ$, then the lengths of α, β do not have a general correlation. If $\theta = 60^\circ$ or 120° , then α, β have the same length. If $\theta = 45^\circ$ or 135° , then the ratio of the length of the longer root

to the length of the shorter root is $\sqrt{2}$. If $\theta = 30^\circ$ or 150° , then the ratio of the length of the longer root to the length of the shorter root is $\sqrt{3}$.

To turn a crystallographic Coxeter graph into a Dynkin diagram, we turn any edge with a label of $n > 3$ into $\frac{n}{2}$ edges with no label. In this case ($n = 4$ or $n = 6$), we also place an arrow on these edges pointing from the longer root to the shorter root. This occurs in types B_ℓ, C_ℓ, F_4, G_2 . This arrow is what differentiates types B_ℓ and C_ℓ , which have the same Coxeter graph. For other types, the Dynkin diagram and the Coxeter graph are the same. Diagrams with only single edges are called **simply laced** and are also said to be of type **ADE**.

Here is a list of connected Dynkin diagrams and the corresponding Coxeter graphs. More graphs and diagrams with various embellishments can be seen in Section 14 (Dynkin Diagrams and Coxeter Graphs, p. 193) in Chapter 8 (Miscellaneous Calculations and Observations, p. 173).

Type	Lie algebra	Dynkin diagram	Coxeter graph
A_ℓ	$\mathfrak{sl}(\ell + 1)$		
B_ℓ	$\mathfrak{so}(2\ell + 1)$		
C_ℓ	$\mathfrak{sp}(2\ell)$		
D_ℓ	$\mathfrak{so}(2\ell)$		
E_6			
E_7			
E_8			
F_4			
G_2			

Root systems can also be described by Cartan matrices. If Φ is a root system with base Δ , we define its **Cartan matrix** to be the $\Delta \times \Delta$ matrix A whose α, β -entry is $\frac{2(\alpha, \beta)}{(\beta, \beta)}$. When the root system is written as a union of irreducible components, the Cartan matrix has a block diagonal form with the Cartan matrices of the irreducible components as its blocks. A Cartan matrix always has 2's on the diagonal and integers from 0, -1, -2, -3 off-diagonal.

Cartan matrices and Dynkin diagrams convey the same information. To obtain a Cartan matrix from the Dynkin diagram, first put 2's on the diagonal. Then put 0 in the i, j entry whenever there is no edge between vertex i and vertex j . If there is one edge joining vertex i and vertex j , put a -1 in both the i, j and j, i positions. If there are $m > 1$ edges joining vertex i and vertex j and there is an arrow pointing from vertex i to vertex j , put a -1 in the j, i position and a $-m$ in the i, j position.

To obtain a Dynkin diagram from a Cartan matrix, connect vertex i and vertex j with an edge whenever the i, j (and hence also the j, i) entry is nonzero. If either the i, j or j, i entry is $-m$ for $m > 1$, connect vertices i and j with m edges and draw an arrow pointing from i to j if the i, j entry is $-m$.

Here is a list of Cartan matrices for the classical Lie algebras.

Type	Lie algebra	Cartan Matrix
A_ℓ	$\mathfrak{sl}(\ell+1)$	$\begin{pmatrix} 2 & -1 & 0 & \dots & 0 & 0 & 0 \\ -1 & 2 & -1 & \dots & 0 & 0 & 0 \\ 0 & -1 & 2 & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 2 & -1 & 0 \\ 0 & 0 & 0 & \dots & -1 & 2 & -1 \\ 0 & 0 & 0 & \dots & 0 & -1 & 2 \end{pmatrix}$
B_ℓ	$\mathfrak{so}(2\ell+1)$	$\begin{pmatrix} 2 & -1 & 0 & \dots & 0 & 0 & 0 \\ -1 & 2 & -1 & \dots & 0 & 0 & 0 \\ 0 & -1 & 2 & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 2 & -1 & 0 \\ 0 & 0 & 0 & \dots & -1 & 2 & -2 \\ 0 & 0 & 0 & \dots & 0 & -1 & 2 \end{pmatrix}$
C_ℓ	$\mathfrak{sp}(2\ell)$	$\begin{pmatrix} 2 & -1 & 0 & \dots & 0 & 0 & 0 \\ -1 & 2 & -1 & \dots & 0 & 0 & 0 \\ 0 & -1 & 2 & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 2 & -1 & 0 \\ 0 & 0 & 0 & \dots & -1 & 2 & -1 \\ 0 & 0 & 0 & \dots & 0 & -2 & 2 \end{pmatrix} = B_\ell^T$
D_ℓ	$\mathfrak{so}(2\ell)$	$\begin{pmatrix} 2 & -1 & 0 & \dots & \dots & 0 & 0 & 0 \\ -1 & 2 & -1 & \dots & \dots & 0 & 0 & 0 \\ 0 & -1 & 2 & \ddots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 2 & -1 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & -1 & 2 & -1 & -1 \\ 0 & 0 & 0 & \dots & 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & \dots & 0 & -1 & 0 & 2 \end{pmatrix}$

Here is a list of Cartan matrices for the exceptional Lie algebras.

Type	Lie Algebra	Cartan Matrix
E_6		$\begin{pmatrix} 2 & 0 & -1 & 0 & 0 & 0 \\ 0 & 2 & 0 & -1 & 0 & 0 \\ -1 & 0 & 2 & -1 & 0 & 0 \\ 0 & -1 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & -1 & 2 \end{pmatrix}$
E_7		$\begin{pmatrix} 2 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & 2 & -1 & 0 & 0 & 0 \\ 0 & -1 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 \end{pmatrix}$
E_8		$\begin{pmatrix} 2 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & -1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 2 \end{pmatrix}$
F_4		$\begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -2 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}$
G_2		$\begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}$

Now we list some numbers associated to simple Lie algebras.

Proposition 6.1.1 (Numerical invariants of simple Lie algebras).

Let L be a finite-dimensional split simple Lie algebra over a field of characteristic 0, H be

a split maximal toral subalgebra, Φ be its set of roots, Φ^+ be its set of positive roots, Δ be its set of simple roots (base), and W be its Weyl group.

Type	Lie algebra	$\dim L$	$ \Phi $	$ \Phi^+ $	$ \Delta = \dim H$	$ W $
A_ℓ	$\mathfrak{sl}(\ell + 1)$	$\ell(\ell + 2)$	$\ell(\ell + 1)$	$\frac{1}{2}\ell(\ell + 1)$	ℓ	$(\ell + 1)!$
B_ℓ	$\mathfrak{so}(2\ell + 1)$	$\ell(2\ell + 1)$	$2\ell^2$	ℓ^2	ℓ	$2^\ell \ell!$
C_ℓ	$\mathfrak{sp}(2\ell)$	$\ell(2\ell + 1)$	$2\ell^2$	ℓ^2	ℓ	$2^\ell \ell!$
D_ℓ	$\mathfrak{so}(2\ell)$	$\ell(2\ell - 1)$	$2\ell(\ell - 1)$	$\ell(\ell - 1)$	ℓ	$2^{\ell-1} \ell!$
E_6		78	72	36	6	$51,840 = 2^7 3^4 5^1$
E_7		133	126	63	7	$2,903,040 = 2^{10} 3^4 5^1 7^1$
E_8		248	240	120	8	$696,729,600 = 2^{14} 3^5 5^2 7^1$
F_4		52	48	24	4	$1152 = 2^7 3^2$
G_2		14	12	6	2	$12 = 2^2 3^1$

REFERENCES AND REMARKS. See [Humphreys, §12.2, Table 1, p. 66].

PROOF. **Omitted.** ■

2. Classification of Split Semisimple Lie Algebras

Proposition 6.2.1 (Classification of split semisimple Lie algebras).

Let $\mathbb{F}$ be a field of characteristic 0.

- (1) Every split semisimple Lie algebra over $\mathbb{F}$ has a Dynkin diagram that is a disjoint union of copies of A_ℓ , B_ℓ ($\ell \geq 2$), C_ℓ ($\ell \geq 3$), D_ℓ ($\ell \geq 4$), E_6 , E_7 , E_8 , F_4 , G_2 .
- (2) Any two split semisimple Lie algebras over $\mathbb{F}$ with the same Dynkin diagram are isomorphic.
- (3) There is a split simple Lie algebra with each of the Dynkin diagrams listed in (1).

REFERENCES AND REMARKS. See [Humphreys, §18.4, Theorem, p. 101; Jacobson, §IV.4, Theorem 3, p. 127, §IV.6, pp. 135-146; Bourbaki 7-9, §VIII.4.3, Theorem 1, p. 102, §VIII.4.4, Theorem 2, p. 105; Erdmann-Wildon, Theorem 14.6, p. 157; Fulton-Harris, §21.2, Theorem 21.11, pp. 325-326, §21.3, pp. 330-338; Wikipedia; Wikipedia].

PROOF. **Omitted.** ■

CHAPTER 7

More Representation Theory

This chapter is only a sketch at this point. We discuss the classification of (finite-dimensional) irreducible representations of semisimple Lie algebras. In the Section 1 (Basics of Weights, p. 165), we define the fundamental weights and show their nonnegative integral linear combinations form a special class known as dominant integral weights. In the Section 2 (The Theorem of the Highest Weight, p. 168), we state the theorem of the highest weight: the finite-dimensional irreducible representations of a semisimple Lie algebra are in bijective correspondence with the set of dominant integral weights. In Section 3 (Representations of $\mathfrak{sl}(n)$, p. 169), we specialize this representation theory to the case of $L = \mathfrak{sl}(n)$.

1. Basics of Weights

Throughout this chapter, L is a semisimple Lie algebra with split maximal toral subalgebra (i.e., split Cartan subalgebra) H and set of roots Φ relative to H . We know that Φ spans H^* , so it must contain a basis. It turns out that it contains a subset, called a base, with a far stronger property. A **base** is a set of roots with the property that every root is either a linear combination of elements of the base with nonnegative integer coefficients or a linear combination of elements of the base with nonnegative integer coefficients. Elements of a base are called **simple roots**. This is covered in Chapter 5 (Root Systems, p. 141).

For the reader who skipped that chapter, we recall that roots can be divided disjointly into positive and negative roots (the negative roots being the additive inverses of the positive ones). We denote the set of positive roots Φ^+ . Once that is done, we can choose a base Δ of Φ with the property that every positive root is a linear combination of elements of the base with nonnegative integer coefficients. Thus Δ is a basis for H^* . The elements of Δ are called **simple roots**. The set Φ depends on the choice of H and the set Δ depends on the choice of which roots are positive. Once we choose H and Φ^+ , everything else is uniquely determined. Alternatively, the positive roots are determined once we choose Δ .

For example, if $L = \mathfrak{sl}(n)$, we typically take H to be the set of diagonal matrices in L , in which case the set of roots is $\Phi = \{ \pm \alpha_{ij} \mid 1 \leq i < j \leq n \}$. If we declare the positive roots to be the α_{ij} with $i < j$, then the base is $\Delta = \{ \alpha_{i,i+1} \mid 1 \leq i \leq n-1 \}$.

For any $\alpha \in \Phi$, let S_α be the copy of $\mathfrak{sl}(2)$ spanned by $e_\alpha, h_\alpha, f_\alpha$, as described in Proposition 4.2.5 (Properties of the root space decomposition, p. 135)(4,5). We use the notation of that proposition. In this chapter, we will denote the dual Killing form by $(,)$. That is, if $\alpha, \beta \in H^*$, then $(\alpha, \beta) = \kappa^*(\alpha, \beta) = \alpha(t_\beta) = \beta(t_\alpha)$. (Here, κ^* should really be $\kappa|_{H^*}$.)

If V is a (finite-dimensional) representation of L , then V is a representation of each S_α . By the $\mathfrak{sl}(2)$ -representation theory, the action of h_α on V is diagonalizable. Since H is commutative and spanned by the h_α for $\alpha \in \Phi$, the action of H on V is simultaneously diagonalizable. In particular, $V = \bigoplus_{\lambda \in H^*} V_\lambda$, where $V_\lambda = \{ v \in V \mid h.v = \lambda(h)v \text{ for all } h \in H \}$. We call the nonzero spaces V_λ the **weight spaces** of V and the corresponding λ 's the **weights** of V . Just as with root spaces, we have $\alpha \in \Phi$, then $L_\alpha.V_\lambda \subseteq V_{\alpha+\lambda}$. To see this, let $h \in H, x \in L_\alpha, v \in V_\lambda$. Then $h.(x.v) = x.(h.v) + [h, x].v = x.\lambda(h)v + \alpha(h)x.v = (\lambda + \alpha)(h)x.v$.

As we did in Section 3 (Positive Roots and Bases, p. 147) in Chapter 5 (Root Systems, p. 141) for roots, we can define a partial order on weights. We can also define a **highest weight vector** for the representation V to be a nonzero $v \in V_\lambda$ for some weight λ such that $L_\alpha.v = 0$ for all positive roots α . We call the corresponding λ a **highest weight**. We will show that every (nonzero, finite-dimensional) representation has a highest weight vector and every irreducible representation has a unique (up to scalar multiples) highest weight vector whose weight uniquely determines the representation.

For $\alpha \in H^*$, we define the **coroot** of α to be $\alpha^\vee = \frac{2\alpha}{(\alpha, \alpha)} = \frac{2\alpha}{\kappa(t_\alpha, t_\alpha)}$. Just as t_α is dual to α in the sense that $\alpha(h) = \kappa(t_\alpha, h)$ for $h \in H$, the element h_α is dual to $\alpha^\vee$ in the sense that $\alpha^\vee(h) = \kappa(h_\alpha, h)$ for $h \in H$. (Some sources call h_α the coroot of α .) The $\alpha^\vee$ with $\alpha \in \Delta$ form a basis $\Delta^\vee$ for H^* . Since the dual Killing form $(,)$ is nondegenerate, there is a basis $\{ \omega_\alpha \mid \alpha \in \Delta \}$ of H^* that is dual to $\Delta^\vee$. That is, $(\alpha^\vee, \omega_\beta) = \delta_{\alpha, \beta}$ for all $\alpha, \beta \in \Delta$. The set $\{ \omega_\alpha \mid \alpha \in \Delta \}$ is called the set of **fundamental weights** for L (relative to Δ).

Warning: As usual, there are notational variations. [Humphreys, §13.1, p. 67] denotes the fundamental weights by λ , as do [Jacobson, §VII.4, Proof, p. 221] and [Erdmann-Wildon, §15.1.1, p. 167]. [Bourbaki 4-6, §VI.1.10, p. 180] and [Bourbaki 7-9, §VIII.7.3, Proposition 8, p. 131] use ϖ . This character may look like $\overline{\omega}$ (`\overline{\omega}`), but it is actually `\varpi` in L^AT_EX.

If $\lambda \in H^*$, we say λ is **integral** if $\frac{2(\lambda, \alpha)}{(\alpha, \alpha)} = \lambda(h_\alpha)$ is an integer for every $\alpha \in \Phi$. In fact, it is enough that $\frac{2(\lambda, \alpha)}{(\alpha, \alpha)}$ is an integer for every $\alpha \in \Delta$. We say λ is **dominant** if $\frac{2(\lambda, \alpha)}{(\alpha, \alpha)} \geq 0$ for every $\alpha \in \Delta$ and **strongly dominant** if $\frac{2(\lambda, \alpha)}{(\alpha, \alpha)} > 0$ for every $\alpha \in \Delta$.

Proposition 7.1.1 (Fundamental and dominant integral weights).

Let L be a finite-dimensional semisimple Lie algebra over a field of characteristic 0 with a split maximal toral subalgebra H . Let Φ be the set of roots, Φ^+ the set of positive roots, and Δ a base for Φ .

- (1) *The integral weights are precisely the integral linear combinations of fundamental weights.*
- (2) *The dominant integral weights are precisely the linear combinations of fundamental weights with nonnegative integer coefficients.*

PROOF. Let ω_α be the fundamental weight dual to $\alpha \in \Delta$. If λ is any weight, set $m_\alpha = \frac{2(\lambda, \alpha)}{(\alpha, \alpha)} = \lambda(h_\alpha)$. Since $(\omega_\beta, \beta) = \frac{(\beta, \beta)}{2}$, we have $(\lambda - \sum_{\alpha \in \Delta} m_\alpha \omega_\alpha, \beta) = (\lambda, \beta) - (\lambda, \beta) = 0$ for all $\beta \in \Delta$. Since $(,)$ is nondegenerate, this implies $\lambda = \sum_{\alpha \in \Delta} m_\alpha \omega_\alpha$.

(1) If λ is an integral weight, then each $m_\alpha \in \mathbb{Z}$, and so the calculation above proves (1).

(2) If λ is a dominant integral weight, then each m_α is a nonnegative integer. The calculation above thus proves (2). ■

Another way to state the result of the last proposition is that the integral weights form a lattice with basis the fundamental weights; we call this lattice the **weight lattice** of L . The lattice of $\mathbb{Z}$ -linear combinations of the (simple) roots is called the **root lattice** and has the simple roots as its basis. Both lattices depend in some sense on the choice of H and Δ . It's easy to see that all roots are integral weights, so the root lattice is a sublattice of the weight lattice. Roots need not be dominant integral weights. In fact, if $L \neq \mathfrak{sl}(2)$ is simple, no simple root is a dominant integral weight.

If $\alpha, \beta \in \Delta$, we know $a_{\alpha\beta} = \frac{2(\beta, \alpha)}{(\beta, \beta)} = \alpha(h_\beta)$ is an integer by Proposition 4.2.5 (Properties of the root space decomposition, p. 135)(6). These integers are called the **Cartan integers** of L , and the $\Delta \times \Delta$ matrix with α, β entry $a_{\alpha\beta}$ is called the **Cartan matrix** of L (relative to Δ). Its diagonal entries are 2's and its off-diagonal entries are nonpositive integers.

Proposition 7.1.2 (Simple roots and fundamental weights).

Let L be a finite-dimensional semisimple Lie algebra over a field of characteristic 0 with a

split maximal toral subalgebra H . Let Φ be the set of roots, Φ^+ the set of positive roots, and Δ a base for Φ . Let ω_α be the fundamental weights corresponding to the simple roots in Δ and let $A = (a_{\alpha\beta})$ be the Cartan matrix defined above. Let α be the column vector of simple roots and ω be the column vector of fundamental weights. Then $\alpha = A\omega$.

REFERENCES AND REMARKS. See [Humphreys, §13.1, p. 68].

PROOF. We know the fundamental weights form a basis of H^* , so we can write $\alpha = \sum_{\beta \in \Delta} c_{\alpha,\beta} \omega_\beta$ for some scalars $c_{\alpha,\beta}$. For any $\gamma \in \Delta$, we have $\frac{2(\alpha,\gamma)}{(\gamma,\gamma)} = \sum_{\beta \in \Delta} c_{\alpha,\beta} \frac{2(\omega_\beta,\gamma)}{(\gamma,\gamma)} = c_{\alpha,\gamma}$. Thus $\alpha = \sum_{\beta \in \Delta} \frac{2(\alpha,\beta)}{(\beta,\beta)} \omega_\beta = \sum_{\beta \in \Delta} a_{\alpha\beta} \omega_\beta$. In matrix notation, this says $\alpha = A\omega$. ■

2. The Theorem of the Highest Weight

Proposition 7.2.1 (Theorem of the highest weight).

Let L be a finite-dimensional semisimple Lie algebra over a field of characteristic 0 with a split maximal toral subalgebra H . Let Φ be the set of roots, Φ^+ the set of positive roots, and Δ a base for Φ .

- (1) Every finite-dimensional irreducible representation of L has a highest weight that is a dominant integral weight, and two finite-dimensional irreducible representations with the same highest weight are isomorphic.
- (2) For every dominant integral weight λ , there is a finite-dimensional irreducible module with highest weight λ .

REFERENCES AND REMARKS. See [Humphreys, §21.1, Theorem, p. 112, §21.2, Theorem & Corollary, p. 113; Jacobson, §VII.3, Theorem 3, p. 215, §VII.3, Theorem 4, p. 219; Bourbaki 7-9, §VIII.6.3, Theorem 1, p. 120 and §VIII.7.2, Theorem 1, p. 126; Erdmann-Wildon, Theorem 15.5, p. 167; Wikipedia].

PROOF. **Later.** ■

Vista. For every weight $\lambda \in H^*$, there is an infinite-dimensional module with highest weight λ , called the **Verma module** of weight λ , and commonly denoted $M(\lambda)$. The Verma module has a unique irreducible factor module, which is commonly denoted $L(\lambda)$. [Humphreys] calls $M(\lambda)$ the **standard cyclic module** of weight λ and denotes it $Z(\lambda)$. He also uses the notation $V(\lambda)$ instead of $L(\lambda)$. [Bourbaki 7-9] also denotes $M(\lambda)$ by $Z(\lambda)$ and denotes $L(\lambda)$ by $E(\lambda)$. [Erdmann-Wildon] denotes $L(\lambda)$ by $V(\lambda)$. Wikipedia uses W_λ to denote $M(\lambda)$. In an afterword in later editions, [Humphreys] indicates that the notation we use has become standard.

The irreducible quotient $L(\lambda)$ is finite-dimensional if and only if λ is dominant integral. In “most” cases, $L(\lambda) = M(\lambda)$, but the precise conditions for this equality take a little more effort.

All finite-dimensional irreducible representations are isomorphic to some $L(\lambda)$, but not all infinite-dimensional irreducible representations are.

The module $M(\lambda)$ is defined as follows. We start with $B = H \oplus_{\alpha \in \Phi^+} L_\alpha$; this is the Borel subalgebra of L containing H . It's not hard to see B is solvable. We can define a 1-dimensional representation $\mathbb{F}v$ of B via $h.v = \lambda(h)v$ for $h \in H$ and $x.v = 0$ for $x \in L_\alpha$, $\alpha \in \Phi^+$. We define $M(\lambda)$ to be the L -module induced from B to L . The clearest way to define this is to use the universal enveloping algebra: then $M(\lambda) = U(L) \otimes_{U(B)} \mathbb{F}v$. This is a highest weight module with highest weight vector $1 \otimes v$ and highest weight λ .

3. Representations of $\mathfrak{sl}(n)$

Let $L = \mathfrak{sl}(n)$. As we've seen before, if we take H to be the set of diagonal matrices in L , then the set of roots is $\Phi = \{\pm\alpha_{ij} \mid 1 \leq i < j \leq n\}$. Let us define $\alpha_i = \alpha_{i,i+1}$. If we declare the positive roots to be the α_{ij} with $i < j$, then the base is $\Delta = \{\alpha_i \mid 1 \leq i \leq n-1\}$.

Note that $\alpha_{ij}^\vee = \frac{2\alpha_{ij}}{\kappa(t_{ij}, t_{ij})} = 2n\alpha_{ij}$. Define $\omega_i(\text{diag}(d_1, \dots, d_n)) = d_1 + \dots + d_i$. We claim $\omega_1, \dots, \omega_{n-1}$ is the set of fundamental weights for L . To see this, note $(\alpha_{i,i+1}^\vee, \omega_j) = \omega_j(2n\frac{1}{2n}h_{i,i+1}) = \omega_j(h_{i,i+1})$. If $j < i$, then $\omega_j(h_{i,i+1}) = 0$, while $\omega_i(h_{i,i+1}) = 1$. If $j > i$, then $\omega_j(h_{i,i+1}) = 1 + -1 = 0$. Thus the ω_j are dual to the simple roots $\alpha_i^\vee$ relative to $(,)$.

Let's compute the Cartan matrix A . We have

$$a_{ij} = a_{\alpha_i, \alpha_j} = \alpha_{i,i+1}(h_{j,j+1}) = \begin{cases} -1, & \text{if } j = i \pm 1 \\ 2, & \text{if } j = i \\ 0, & \text{otherwise.} \end{cases}$$

For example, if $j = i - 1$, then $a_{ij} = \alpha_{i,i+1}(h_{i-1,i}) = -1 - 0 = -1$, while if $j = i + 1$, then $a_{ij} = \alpha_{i,i+1}(h_{i+1,i+2}) = 0 - 1 = -1$.

$$\text{Thus } A = \begin{pmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ 0 & -1 & 2 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 2 \end{pmatrix}.$$

It is not hard to see directly that

$$\alpha_{12} = 2\omega_1 - \omega_2, \quad \alpha_{n-1,n} = -\omega_{n-2} + 2\omega_{n-1} \quad \text{and} \quad \alpha_{i,i+1} = -\omega_{i-1} + 2\omega_i - \omega_{i+1} \quad \text{for } 2 \leq i \leq n-2.$$

This agrees with Proposition 7.1.2 (Simple roots and fundamental weights, p. 167), given the Cartan matrix above. If d_n is the determinant of the Cartan matrix for $\mathfrak{sl}(n)$, then expanding along the first row of the Cartan matrix, one gets the relation $d_n = 2d_{n-1} - d_{n-2}$ (where $d_0 = 0, d_1 = 1$). From this, it's easy to deduce that $d_n = n$. Note that $\mathfrak{sl}(n)$ is

often expressed by its Dynkin diagram of type A, whence we often write $\mathfrak{sl}(\ell + 1) = A_\ell$. We this notation, we would have $d_\ell = \ell + 1$.

Let's see how these results compare with our previous results for $\mathfrak{sl}(2)$ and $\mathfrak{sl}(3)$ in Section 3 (Representations of $\mathfrak{sl}(2)$, p. 103) and Section 4 (Representations of $\mathfrak{sl}(3)$, p. 107) in Chapter 3 (Representations, p. 83).

For $\mathfrak{sl}(2)$, there is only one positive root, $\alpha = \alpha_{12}$ and the Cartan matrix is $A = (2)$. Thus the fundamental weight is $\omega = \frac{1}{2}\alpha$. The corresponding irreducible representation is the natural representation of $\mathfrak{sl}(2)$ on $\mathbb{F}^2$, with weights $-1, 1$.

For $\mathfrak{sl}(3)$, the simple roots are $\alpha_1 = \alpha_{12}$, $\alpha_2 = \alpha_{23}$, and the Cartan matrix is $A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$. For a diagonal matrix D , the fundamental weights are given by $\omega_1(D) = d_{11}$ and $\omega_2(D) = d_{11} + d_{22} = -d_{33}$. In the notation of Section 4 (Representations of $\mathfrak{sl}(3)$, p. 107) in Chapter 3 (Representations, p. 83), the notation for the α 's is the same, and $\omega_1 = \beta_1$, $\omega_2 = \beta_1 + \beta_2 = -\beta_3$. Just as in that section, we have $\alpha_1 = 2\omega_1 - \omega_2$, $\alpha_2 = -\omega_1 + 2\omega_2$, $\omega_1 = \frac{2}{3}\alpha_1 + \frac{1}{3}\alpha_2$, $\omega_2 = \frac{1}{3}\alpha_1 + \frac{2}{3}\alpha_2$. The weight ω_1 is the highest weight of the natural representation $\mathbb{F}^3$ of $\mathfrak{sl}(3)$; this is the irreducible module $\Gamma_{1,0}$ in the notation of Proposition 3.4.1 (Irreducible representations of $\mathfrak{sl}(3)$, p. 113). The weight ω_2 is the highest weight of the irreducible module $\Gamma_{0,1}$, which is the dual of $\Gamma_{1,0}$. In fact, $\Gamma_{a,b}$ is the irreducible module with highest weight $a\omega_1 + b\omega_2$.

Let's look at some small Cartan matrices. When $n = 2$, we have $A = (2)$, $A^{-1} = (\frac{1}{2})$. When $n = 3$, we have $A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$, $A^{-1} = \frac{1}{3} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. When $n = 4$, we have $A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$, $A^{-1} = \frac{1}{4} \begin{pmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{pmatrix}$.

4. The Weyl Character Formula

Let Λ be the root lattice of L , that is, the $\mathbb{Z}$ -linear span of the fundamental weights. We will define the group ring $\mathbb{Z}[\Lambda]$ in a slightly unusual way. The elements of $\mathbb{Z}[\Lambda]$ are sums $\sum_{\lambda \in \Lambda} n_\lambda e^\lambda$, where the n_λ are integers, only finitely many of which are nonzero. Multiplication in $\mathbb{Z}[\Lambda]$ is defined via the distributive law and the fundamental formula $e^\lambda e^\mu = e^{\lambda+\mu}$. This ring is isomorphic to the polynomial ring $\mathbb{Z}[x_1, \dots, x_n]$, where the variables x_i are the elements e^{ω_i} for the fundamental weights $\omega_1, \dots, \omega_n$.

The **formal character** of a (finite-dimensional) representation V of L is $\text{ch } V = \sum_{\lambda \in \Lambda} \dim V_\lambda e^\lambda \in \mathbb{Z}[\Lambda]$. The dimensions $\dim V_\lambda$ are the multiplicities of the weight λ in V .

Before we state Weyl's formula for characters of irreducible representations, we need two more objects. First, define $\rho = \frac{1}{2} \sum_{\alpha \in \Phi^+} \alpha$. This is called (naturally) the **half-sum of the positive roots**. ([Humphreys, §13.3, p. 70] uses the letter δ instead but ρ has become the common symbol.) The weight ρ plays a very important part in many formulas as an offset or correction factor. As the next result shows, ρ is a dominant integral weight; however, it need not be a root.

Proposition 7.4.1 (Half-sum of positive roots).

Let L be a finite-dimensional semisimple Lie algebra over a field of characteristic 0 with a split maximal toral subalgebra H . Let Φ be the set of roots, Φ^+ the set of positive roots, and Δ a base for Φ . Let ρ be the half-sum of the positive roots.

- (1) ρ is a strictly dominant integral weight.
- (2) ρ is equal to the sum of the fundamental weights.

REFERENCES AND REMARKS. See [Humphreys, §13.3, Lemma A, p. 70; Fulton-Harris, §24.1, p. 400-401]

PROOF. One can check that $\frac{2(\rho, \alpha)}{(\alpha, \alpha)} = 1$ for all simple roots α . This implies in turn that ρ is the sum of the fundamental weights and is strictly dominant integral. *More details.* ■

One can check that $\frac{2(\rho, \alpha)}{(\alpha, \alpha)} = 1$ for all simple roots α .

For the reader who skipped Chapter 5 (Root Systems, p. 141), we recall that the **Weyl group** of a semisimple Lie algebra L is the subgroup W of $\mathbf{GL}(n, \mathbb{R})$ generated by the reflections about the roots in the root system of L . For each α in the root system (or in $\mathbb{R}^n \setminus \{0\}$ generally), let σ_α denote the reflection about the hyperplane (through the origin) perpendicular to α . The key fact is that $\sigma_\alpha(\Phi) = \Phi$: thus every member of the Weyl group takes roots to roots.

Each element $\sigma \in W$ has a **sign** $\text{sgn } \sigma$. This is defined to be $(-1)^\ell$, where ℓ is the number of simple reflections in any expression for σ as a product of reflections about simple roots.

Now we are ready to state Weyl's formula.

Proposition 7.4.2 (Weyl character formula).

Let L be a finite-dimensional semisimple Lie algebra over a field of characteristic 0 with a split maximal toral subalgebra H . Let Φ be the set of roots, Φ^+ the set of positive roots, Δ a base for Φ , and W the Weyl group. Let ρ be the half-sum of the positive roots. For any weight μ , define $A_\mu = \sum_{\sigma \in W} \text{sgn } \sigma e^{\sigma(\mu)}$. For a dominant integral weight λ , let $L(\lambda)$ be the irreducible representation of L with highest weight λ .

- (1) If λ is dominant integral, then $\text{ch } L(\lambda) = \frac{A_{\lambda+\rho}}{A_\rho}$.
- (2) $A_\rho = \prod_{\alpha \in \Phi^+} (e^{\alpha/2} - e^{-\alpha/2}) = e^{-\rho} \prod_{\alpha \in \Phi^+} (e^\alpha - 1)$.

REFERENCES AND REMARKS. See [**Humphreys**, §24.3, Theorem, p. 139; **Jacobson**, §VIII.3, Weyl's Theorem (35) & (36), p. 255; **Bourbaki 7-9**, §VIII.9.1, Theorem 1, p. 154; **Fulton-Harris**, §24.1, Theorem 24.2, p. 400; **Wikipedia**]

PROOF. Not given. ■

Proposition 7.4.3 (Weyl dimension formula).

Let L be a finite-dimensional semisimple Lie algebra over a field of characteristic 0 with a split maximal toral subalgebra H . Let Φ be the set of roots, Φ^+ the set of positive roots, Δ a base for Φ , and W the Weyl group. Let ρ be the half-sum of the positive roots. Let λ be a dominant integral weight and let $L(\lambda)$ be the irreducible representation of L with highest weight λ . Then $\dim L(\lambda) = \frac{\prod_{\alpha \in \Phi^+} (\lambda + \rho, \alpha)}{\prod_{\alpha \in \Phi^+} (\rho, \alpha)}$.

REFERENCES AND REMARKS. See [**Humphreys**, §24.3, Corollary, p. 139; **Jacobson**, §VIII.3, (40), p. 257; **Bourbaki 7-9**, §VIII.9.2, Theorem 2, pp. 154-155; **Fulton-Harris**, §24.1, Corollary 24.4, p. 402; **Wikipedia**]

PROOF. Not given. ■

CHAPTER 8

Miscellaneous Calculations and Observations

In this chapter, we collect some calculations, examples, and observations. There is no rhyme or reason to the ordering of the sections.

1. Squarefree Polynomials and the Derivative

Proposition 8.1.1 (A polynomial is squarefree iff it is relatively prime to its derivative). *Assume $\mathbb{F}$ is perfect, e.g., of characteristic 0. Let $f \in \mathbb{F}[x]$ and let f' be the formal derivative of f .*

- (1) *If $p \in \mathbb{F}[x]$ is irreducible and $p \mid \gcd(f, f')$, then $p^2 \mid f$.*
- (2) *f is squarefree if and only if $\gcd(f, f') = 1$.*

PROOF. (1) Write $f = qp^2 + r$ where $\deg r < \deg p^2$. Note that $p \mid r$. Write $r = ps$, so that $f = p(qp + s)$ and $\deg s < \deg p$. Differentiate this to get $f' = p'(qp + s) + p(q'p + qp' + s')$. Since $p \mid f'$, this implies $p \mid p's$. As p is irreducible, we see either $p \mid p'$ or $p \mid s$. By considering degrees, we see both of these are impossible unless $s = 0$. Thus $r = 0$, and so $p^2 \mid f$.

(2) Suppose f is squarefree and $\gcd(f, f') \neq 1$. Then there is an irreducible p with $p \mid \gcd(f, f')$. By (1), this implies $p^2 \mid f$, which is a contradiction.

Suppose $p^2 \mid f$, say $f = gp^2$. Then $f' = g'p^2 + 2gpp'$ is divisible by f . Thus $\gcd(f, f') \neq 1$. ■

Proposition 8.1.2 (Field extensions don't change being squarefree).

Assume $\mathbb{F}, \mathbb{E}$ are perfect, e.g., of characteristic 0. Let $\mathbb{E}$ be an extension field of $\mathbb{F}$ and let $f \in \mathbb{F}[x]$. Then f is squarefree in $\mathbb{F}[x]$ if and only if it is squarefree in $\mathbb{E}[x]$.

PROOF. This is immediate from Proposition 8.1.1 (A polynomial is squarefree iff it is relatively prime to its derivative, p. 173). ■

Suppose $F = \mathbb{F}_p(t)$ for a prime p . Then $f = x^p - t \in \mathbb{F}[x]$ is irreducible and in particular is squarefree. However, $\gcd(f, f') = 0$. Moreover, if $\mathbb{E} = \mathbb{F}[t^{1/p}]$, then $f = (x - t^{1/p})^p$ is not squarefree.

2. Multiplication and Commutator of Triangular Matrices

We will prove some formulas for the multiplication and commutators of certain subspaces of the space of upper triangular matrices.

Let T_0 be the space of $n \times n$ upper triangular matrices over $\mathbb{F}$ and let D_0 be the space of diagonal matrices. For $k > 0$, let D_k be set of matrices whose nonzero entries are concentrated on the k^{th} diagonal above the main diagonal and let T_k be the space of upper triangular matrices such that all entries of the main diagonal and the first $k - 1$ diagonals above the main diagonal are all 0. Thus $A \in D_k$ if and only if $a_{ij} = 0$ unless $j = i + k$ and $A \in T_k$ if and only if $a_{ij} = 0$ whenever $j < i + k$. For example, T_1 is the space of strictly upper triangular matrices and D_1 contains the off-diagonal entries of every Jordan block. We declare $D_k = 0$ and $T_k = 0$ for $k \geq n$. Note that $T_i = \sum_{k \geq i} D_k$.

In the proposition below, given subspaces U, W of $\mathfrak{gl}(n) = M_n(\mathbb{F})$, we let UW and $[U, W]$ denote the linear span of the products uw and commutators $[u, w] = uw - wu$, respectively, for $u \in U, w \in W$. We leave it to the reader to decide if we actually need linear combinations.

Proposition 8.2.1 (Multiplication of triangular matrices).

Let n be a positive integer and let T_i and D_i be as defined just before the proposition. Let i, j be nonnegative integers.

- (1) $D_i D_j = D_{i+j}$ and $T_i T_j = T_{i+j}$.
- (2) If $(i, j) \neq (0, 0)$, then $[D_i, D_j] = D_{i+j}$ and $[T_i, T_j] = T_{i+j}$.
- (3) $[D_0, D_0] = 0$ and $[T_0, T_0] = T_1$.
- (4) Any product of n strictly upper triangular matrices is 0, and thus if A is strictly upper triangular, $A^n = 0$. If A, B are upper triangular, then $AB - BA$ is strictly upper triangular and hence nilpotent.

PROOF. We first note that a basis for D_i is the set of matrix units $e_{r, r+i}$ with $1 \leq r \leq n-i$. We will prove the equalities for matrix units in the case of the D_i and then prove the results for the T_i .

(1) Since matrix unit products are zero unless they are of the form $e_{rs}e_{st} = e_{rt}$, it's enough to consider such products. If $e_{rs} \in D_i$, $e_{st} \in D_j$, then $s = r+i$, $t = s+j = r+i+j$, so $e_{rs}e_{st} = e_{r, r+i+j} \in D_{i+j}$. Moreover, every matrix unit in D_{i+j} will occur in such a product, so $D_i D_j = D_{i+j}$.

(2) If $t > r$, then $[e_{rs}, e_{st}] = e_{rs}e_{st}$. Thus if $i+j \neq 0$, the calculations in (1) show $[D_i, D_j] = D_{i+j}$.

(3) This is obvious for D_0 since diagonal matrices commute.

Now we use the equality $T_i = \sum_{k \geq i} D_k$ to prove the results for the T_i . Thus

$$T_i T_j = \sum_{\substack{k \geq i \\ \ell \geq j}} D_k D_\ell = \sum_{\substack{k \geq i \\ \ell \geq j}} D_{k+\ell} = \sum_{r \geq i+j} D_r = T_{i+j}.$$

Likewise, as long as $(i, j) \neq (0, 0)$, we have

$$[T_i, T_j] = \sum_{\substack{k \geq i \\ \ell \geq j}} [D_k, D_\ell] = \sum_{\substack{k \geq i \\ \ell \geq j}} D_{k+\ell} = \sum_{r \geq i+j} D_r = T_{i+j}.$$

Finally, $[T_0, T_0] = \sum_{\ell \geq 0} [D_\ell, D_\ell] = \sum_{r \geq 0} D_r + [D_0, D_0] = T_1$.

(4) For the fact about the product of strictly upper triangular matrices, use (1) and induction. The fact that $AB - BA$ is upper triangular is part of (3). ■

Recall that if S is a subspace of L , we define $C(S) = \{x \in L \mid [x, s] = 0 \text{ for all } s \in S\}$ and $N(S) = \{x \in L \mid [x, s] \in S \text{ for all } s \in S\}$.

Proposition 8.2.2 (Centralizers and normalizers for triangular matrices).

Let n be a positive integer and let T_i and D_i be as defined just before the proposition. Let i, j be nonnegative integers.

- (1) $C(D_i) = C(T_i) = T_{n-i}$ if $i > 0$.
- (2) $C(D_0) = D_0$ and $C(T_0) = 0$.
- (3) $N(D_i) = D_0 + T_{n-i}$ and $N(T_i) = L$ (i.e., T_i is an ideal of $\mathfrak{t}(n)$) for all $i \geq 0$.

PROOF. We apply the results of Proposition 8.2.1 (Multiplication of triangular matrices, p. 174).

(1) If $i > 0$, then $[D_i, D_j] = 0$ if and only if $i+j \geq n$, i.e., $j \geq n-i$. Thus $C(D_i) = T_{n-i}$, and this implies $C(T_i) \subseteq T_{n-i}$. But $[T_i, T_{n-i}] = 0$, so equality holds.

(2) This is an easy calculation.

(3) $[D_i, D_j] \subseteq D_i$ if and only if $j = 0$ or $[D_i, D_j] = 0$. Thus $N(D_i) = D_0 + T_{n-i}$ by (1). We always have $[T_i, T_j] \subseteq T_i$. ■

3. Matrices for the Orthogonal and Symplectic Groups

We explicitly compute the elements of $\mathfrak{gl}(n, M)$ for the matrices M used by [Humphreys, §1.2, pp. 2-3] in defining his version of the orthogonal and symplectic groups.

Suppose $n = 2m$ is even and $M = \begin{pmatrix} 0 & I_m \\ I_m & 0 \end{pmatrix}$ and suppose $X = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$. Then $X^T M = \begin{pmatrix} A^T & C^T \\ B^T & D^T \end{pmatrix} \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} = \begin{pmatrix} C^T & A^T \\ D^T & B^T \end{pmatrix}$, while $MX = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} C & D \\ A & B \end{pmatrix}$. Thus $X^T M = -MX$ if and only if $C^T = -C$, $A^T = -D$, $D^T = -A$, $B^T = -B$. The two

conditions involving A, D are equivalent, so dropping one of them, we get the condition to be in this version of $\mathfrak{o}(n)$. This matrix representation shows again that $\mathfrak{o}(n) \subseteq \mathfrak{sl}(n)$.

When $n = 2m + 1$ is odd, let $M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & I_m \\ 0 & I_m & 0 \end{pmatrix}$ (where the 1 is 1×1 identity).

Suppose $X = \begin{pmatrix} r & s & t \\ v & A & B \\ w & C & D \end{pmatrix}$, where A, B, C, D are $m \times m$, r is 1×1 , s, t are $1 \times m$, and

v, w are $m \times 1$. Then $X^T M = \begin{pmatrix} r & v^T & w^T \\ s^T & A^T & C^T \\ t^T & B^T & D^T \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & I \\ 0 & I & 0 \end{pmatrix} = \begin{pmatrix} r & w^T & v^T \\ s^T & C^T & A^T \\ t^T & D^T & B^T \end{pmatrix}$, while

$MX = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & I \\ 0 & I & 0 \end{pmatrix} \begin{pmatrix} r & s & t \\ v & A & B \\ w & C & D \end{pmatrix} = \begin{pmatrix} r & s & t \\ w & C & D \\ v & A & B \end{pmatrix}$. Thus $X^T M = -MX$ if and only if

$r = -r$, $w^T = -s$, $v^T = -t$, $s^t = -w$, $C^T = -C$, $A^T = -D$, $t^T = -v$, $D^T = -A$, $B^T = -B$. This implies $r = 0$ and (eliminating redundancies) $w = -s^T$, $v = -t^T$, $A^T = -D$, $B^T = -B$, $C^T = -C$. Once again, we see trace of X is 0.

Now consider the symplectic algebra with $m = 2n$ and $M = \begin{pmatrix} 0 & I_m \\ -I_m & 0 \end{pmatrix}$. Suppose $X = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$. Then $X^T M = \begin{pmatrix} A^T & C^T \\ B^T & D^T \end{pmatrix} \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} = \begin{pmatrix} -C^T & A^T \\ -D^T & B^T \end{pmatrix}$, while $MX = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} C & D \\ -A & -B \end{pmatrix}$. Thus $X^T M = -MX$ if and only if $C^T = C$, $A^T = -D$, $-D^T = A$, $B^T = B$. Eliminating the redundant A, D conditions leads to $A^T = -D$, $B^T = B$, $C^T = C$. This implies $\text{tr } X = 0$.

4. The a, b, c Lie Algebras

Let $a, b, c \in \mathbb{F}$ and consider the Lie algebra $L = L_{a,b,c}$ with basis x, y, z and brackets $[x, y] = cz$, $[y, z] = ax$, $[z, x] = by$. It is enough to check the Jacobi identity on a basis, and when two of the elements are equal, the identity always holds. Thus it's enough to check it for x, y, z , and in this case, all summands are 0.

4.1. The four cases and the derived subalgebra. This class of algebras divides into four broad cases, and some of these cases may have subcases. The broad cases are defined by how many of a, b, c are nonzero. Thus case 0 is the case where L is Abelian.

Case 1 is the case where one scalar is nonzero; by a change of variables, we can assume $a = b = 0, c = 1$. This gives us the Heisenberg algebra, which is nilpotent but not Abelian, since $L^1 = \mathbb{F}z$ and $L^2 = 0$.

Case 2 is the case where one scalar is 0. Again, by a change of variables, we can assume $c = 1, a \neq 0, b = 0$. In this case, $L^1 = \mathbb{F}z + \mathbb{F}x$ and $L^2 = L^1$, while $L^{(2)} = 0$. Thus L is solvable but not nilpotent.

In case 3, $L^1 = L$, so L is not solvable. In fact, L is simple (and so semisimple), as we now show. Let I be an ideal of L and let $v \in I$ be nonzero, say $v = \lambda x + \mu y + \nu z$. Then I contains $[x, v] = \mu[x, y] + \nu[x, z] = -b\nu y + c\mu z$. Thus I contains $[[x, v], y] = c\mu[z, y] = -ac\mu x$ and $[[x, v], z] = -b\nu[y, z] = -ab\nu x$. It follows that $x \in I$ unless $\mu = \nu = 0$. But if $\mu = \nu = 0$, then $v = \lambda x$ and so $x \in I$ in this case too. It's now easy to see that $x, y, z \in I$ and hence $I = L$.

The calculations above show that L is a case n algebra if and only if $n = \dim[L, L]$. One can show that all 3-dimensional Lie algebras L with $\dim[L, L] = 0$ or $\dim[L, L] = 3$ are a, b, c Lie algebras for appropriate choice of a, b, c . When $\dim[L, L] = 1$, there is one Lie algebra that is not an a, b, c Lie algebra. If K is the 2-dimensional non-Abelian Lie algebra and $\mathbb{F}$ is the 1-dimensional Lie algebra, then the other Lie algebra with $\dim[L, L] = 1$ is $L = K \oplus \mathbb{F}$. This can be described as having basis x, y, z with $[x, y] = x, [y, z] = [x, z] = 0$.

When $\dim[L, L] = 2$, there are many Lie algebras that are not a, b, c Lie algebras. In fact, if $\mathbb{F}$ is infinite, there are infinitely many non-isomorphic such Lie algebras. They are all solvable but not nilpotent. All 3-dimensional Lie algebras are described in [**Jacobson**, §I.4, pp. 11-14; **Bourbaki 1-3**, §I.6, Exercise 23, p. 107; **Erdmann-Wildon**, (when $\mathbb{F} = \mathbb{C}$) §3.2, pp. 20-25; **Wikipedia** (when $\mathbb{F} = \mathbb{R}$)]

4.2. Change of variables and isomorphisms. Let us now consider what we can do with a simple change of variables in cases 2 and 3. It's easy to see that we can assume $c = 1$. If we set $X = \alpha x, Y = \beta y, Z = \gamma z$, we see $[X, Y] = \alpha\beta\gamma^{-1}Z$. To retain the condition $c = 1$, we need $\gamma = \alpha\beta$.

First we consider case 2. We still have $[Z, X] = \gamma\alpha[z, x] = 0$ and we have $[Y, Z] = \beta\gamma\alpha^{-1}aX$. As $\beta\gamma\alpha^{-1} = \beta^2$, we see that we change a by multiplying it by a square. Thus a is determined up an element of $\mathbb{F}^\times/(\mathbb{F}^\times)^2$. (Of course, there might be further changes possible with other isomorphisms.) This means that if $\mathbb{F}$ is algebraically closed, there is only one case 2 algebra, given by $[x, y] = z, [y, z] = x, [z, x] = 0$. If $\mathbb{F} = \mathbb{R}$, there are two case 2 algebra, given by $[x, y] = z, [y, z] = \pm x, [z, x] = 0$.

Let us consider this change of variables in case 3, again assuming $c = 1$. As above, we need $\gamma = \alpha\beta$. We have $[Z, X] = \gamma\alpha\beta^{-1}bY = \alpha^2bY$ and $[Y, Z] = \beta\gamma\alpha^{-1}aX = \beta^2aX$. Thus

we can change a, b independently by multiplying by squares, and so at worst they are determined as elements of $\mathbb{F}^\times/(\mathbb{F}^\times)^2$. If $\mathbb{F}$ is algebraically closed, there is only one case 3 algebra, given by $[x, y] = z$, $[y, z] = x$, $[z, x] = y$. This is isomorphic to both $\mathfrak{o}(3)$ and $\mathfrak{sl}(2)$. If $\mathbb{F} = \mathbb{R}$, there are conceivably four case 3 algebras, given by the choices $[x, y] = z$, $[y, z] = \pm x$, $[z, x] = \pm y$.

In fact, when $\mathbb{F} = \mathbb{R}$, if a, b, c have the same sign, then $L_{a,b,c} \cong \mathfrak{o}(3)$ and if they do not all have the same sign, then $L_{a,b,c} \cong \mathfrak{sl}(2)$. To see this, we can use the above change of variables and assume $c = 1$. First suppose $a = -1, b = 1$ and set $e = x + y$, $f = -x + y$, $h = 2z$. Then $[e, f] = [x, y] - [y, x] = 2z = h$ and $[h, e] = 2[z, x] + 2[z, y] = 2y + 2x = 2e$ and $[h, f] = -2[z, x] + 2[z, y] = -2y + 2x = -2f$.

Next suppose $a = 1, b = -1$ and set $e = x - y$, $f = x + y$, $h = 2z$. Then $[e, f] = [x, y] - [y, x] = 2z = h$ and $[h, e] = 2[z, x] - 2[z, y] = -2y + 2x = 2e$ and $[h, f] = 2[z, x] + 2[z, y] = -2y - 2x = -2f$.

Finally, suppose $a = b = -1$ and set $e = y + z$, $f = y - z$, $h = 2x$. Then $[e, f] = [z, y] - [y, z] = 2x = h$ and $[h, e] = 2[x, y] + 2[x, z] = 2z + 2y = 2e$ and $[h, f] = 2[x, y] - 2[x, z] = 2z - 2y = -2f$.

Note that when $a = b = c = 1$, the linear representation of L below is $x = e_{32} - e_{23}$, $y = e_{13} - e_{31}$, $z = e_{21} - e_{12}$, and this is a basis of the space of skew symmetric matrices, i.e., $\mathfrak{o}(3)$. Thus $L_{1,1,1} \cong \mathfrak{o}(3)$ over any field.

We'll now show $L_{1,1,1}(\mathbb{R}) \cong \mathfrak{su}(2, \mathbb{R})$. A basis of $\mathfrak{su}(2)$ is given by:

$$x = \frac{1}{2} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \quad y = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad z = \frac{1}{2} \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

Moreover,

$$\begin{aligned} [x, y] &= \frac{1}{4} \left(\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} - \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} \right) = \frac{2}{4} \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} = z, \\ [y, z] &= \frac{1}{4} \left(\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} - \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \right) = \frac{2}{4} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = x, \\ [z, x] &= \frac{1}{4} \left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right) = \frac{2}{4} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = y. \end{aligned}$$

This shows $\mathfrak{su}(2) = L_{1,1,1}$.

4.3. The adjoint representation and linear realizations. Let us compute the adjoint representation of L . A simple calculation shows

$$\begin{aligned}\operatorname{ad} x &= ce_{32} - be_{23} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -b \\ 0 & c & 0 \end{pmatrix} \\ \operatorname{ad} y &= ae_{13} - ce_{31} = \begin{pmatrix} 0 & 0 & a \\ 0 & 0 & 0 \\ -c & 0 & 0 \end{pmatrix} \\ \operatorname{ad} z &= be_{21} - ae_{12} = \begin{pmatrix} 0 & -a & 0 \\ b & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.\end{aligned}$$

As we noted before, when $Z(L) = 0$ this provides us with a linear realization of L . That is, in cases 2 and 3, we can use the above formulas with x, y, z in place of $\operatorname{ad} x, \operatorname{ad} y, \operatorname{ad} z$ and realize L as a subalgebra of $\mathfrak{sl}(3)$. (As a sanity check, note that these formulas give linear independent matrices x, y, z precisely when at most one of a, b, c is zero.)

The subalgebra of diagonal matrices in $\mathfrak{gl}(3)$ is isomorphic to L in case 0, and we can explicitly take $x = e_{11}, y = e_{22}, z = e_{33}$ for a linear realization. In this case, the concrete trace form is the dot product form, with matrix I , while the Killing form is identically 0.

All other cases can be realized as subalgebras of $\mathfrak{sl}(3)$. In case 1, we can see the strictly upper triangular matrices form a subalgebra isomorphic to the Heisenberg algebra L . If $a = b = 0 \neq c$, we can take $x = ce_{12}, y = ce_{23}, z = ce_{13}$ for a linear presentation. The concrete trace form and the Killing form are identically 0 in this case.

4.4. The Killing form and trace form. In cases 2 and 3, the Killing form and the concrete trace form will be identical because we chose the linear representation to be the adjoint representation. To compute either form, we only need the product xy or $\operatorname{ad} x \operatorname{ad} y$, since yx or $\operatorname{ad} y \operatorname{ad} x$ gives us the same trace. However, we will need all of the products below, so we compute them.

Then $x^2 = -bce_{22} - bce_{33}$, $xy = bce_{21}$, $xz = bce_{31}$, $yx = ace_{12}$, $y^2 = -ace_{11} - ace_{33}$, $yz = ace_{32}$, $zx = abe_{13}$, $zy = abe_{23}$, $z^2 = -abe_{11} - abe_{22}$. (To confirm these matrices represent L , note $xy - yx = c(be_{21} - ae_{12}) = cz$, $yz - zy = a(ce_{32} - be_{23}) = ax$, $zx - xz = b(ae_{13} - ce_{31}) = by$.)

Displaying the products, we see

$$\begin{aligned}
 x^2 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & -bc & 0 \\ 0 & 0 & -bc \end{pmatrix} & xy &= \begin{pmatrix} 0 & 0 & 0 \\ bc & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & xz &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ bc & 0 & 0 \end{pmatrix} \\
 yx &= \begin{pmatrix} 0 & ac & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & y^2 &= \begin{pmatrix} -ac & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -ac \end{pmatrix} & yz &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & ac & 0 \end{pmatrix} \\
 zx &= \begin{pmatrix} 0 & 0 & ab \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & zy &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & ab \\ 0 & 0 & 0 \end{pmatrix} & z^2 &= \begin{pmatrix} -ab & 0 & 0 \\ 0 & -ab & 0 \\ 0 & 0 & 0 \end{pmatrix}
 \end{aligned}$$

Reviewing the above calculations, we see $\text{tr } xy = \text{tr } yx = 0$, $\text{tr } yz = \text{tr } zy = 0$, $\text{tr } zx = \text{tr } xz = 0$, $\text{tr } x^2 = -2bc$, $\text{tr } y^2 = -2ac$, $\text{tr } z^2 = -2ab$. Thus the matrix of the concrete trace form and the Killing form is $\text{diag}(-2bc, -2ac, -2ab)$. This is nondegenerate if and only if a, b, c are all nonzero, that is, in case 3.

Suppose we are in case 2 with $c = 1$, $a \neq 0$. Then the matrix is $\text{diag}(0, -2a, 0)$, which has rank 1. The radical of this form is $\mathbb{F}x + \mathbb{F}z = L^1$. In case 3 with $\mathbb{F} = \mathbb{R}$, we see that for $\mathfrak{o}(3)$ (with $a = b = c = 1$), the concrete trace form is negative definite. For $\mathfrak{sl}(2)$ (e.g., $a = b = 1, c = -1$), the form has signature $(2, 1)$, and so it is indefinite.

4.5. The Jordan-Chevalley decomposition. Let $v = \alpha x + \beta y + \gamma z \in L$. Then

$$\begin{aligned}
 \det(\lambda I - v) &= \det \begin{pmatrix} \lambda & \gamma a & -\beta a \\ -\gamma b & \lambda & \alpha b \\ \beta c & -\alpha c & \lambda \end{pmatrix} = \lambda(\lambda^2 + \alpha^2 bc) + \gamma b(\gamma a \lambda - \alpha \beta ac) + \beta c(\alpha \gamma ab + \beta a \lambda) = \\
 &= \lambda^3 + (\alpha^2 bc + \gamma^2 ab + \beta^2 ac)\lambda = [\text{if } abc \neq 0] \quad \lambda \left(\lambda^2 + abc \left(\frac{\alpha^2}{a} + \frac{\beta^2}{b} + \frac{\gamma^2}{c} \right) \right).
 \end{aligned}$$

If exactly one of a, b, c is zero, say $a = 0$, then the characteristic polynomial of v is $\lambda(\lambda^2 + \alpha bc)$. Thus v is nilpotent if $\alpha = 0$ and v is semisimple if $\alpha \neq 0$.

Suppose now none of a, b, c are 0. If $\mathbb{F}$ is an ordered field (e.g., $\mathbb{F} = \mathbb{R}$) and a, b, c have the same sign, then $\frac{\alpha^2}{a} + \frac{\beta^2}{b} + \frac{\gamma^2}{c}$ is not 0 unless $\alpha = \beta = \gamma = 0$, i.e., unless $v = 0$. This means $\det(\lambda I - A)$ has three distinct roots in $\overline{\mathbb{F}}$, so every element of L is semisimple. However, L is not Abelian, or even solvable. (We've seen L is simple.) This shows that we need to include "Abelian" in the definition of a toral subalgebra over non-algebraically-closed fields. For a general $\mathbb{F}$ and L in case 3, the nonzero element v of L is semisimple if and only if $\frac{\alpha^2}{a} + \frac{\beta^2}{b} + \frac{\gamma^2}{c} \neq 0$; otherwise v is nilpotent. Moreover, v is diagonalizable if and only if $-abc(\frac{\alpha^2}{a} + \frac{\beta^2}{b} + \frac{\gamma^2}{c})$ has a square root in $\mathbb{F}$.

4.6. The Casimir element. Let's calculate the (universal) Casimir element in case (3). If we take x, y, z as our basis, the dual basis is $-\frac{1}{2bc}x, -\frac{1}{2ac}y, -\frac{1}{2ab}z$. Thus the Casimir element is $-\frac{1}{2bc}x^2 - \frac{1}{2ac}y^2 - \frac{1}{2ab}z^2 = -\frac{ax^2+by^2+cz^2}{2abc}$. Any scalar multiple will do as well, so we may use $ax^2 + by^2 + cz^2$ if we prefer. The matrix representation of $ax^2 + by^2 + cz^2$ is $\text{diag}(-2abc, -2abc, -2abc)$ and that of $-\frac{ax^2+by^2+cz^2}{2abc}$ is I .

5. Killing and Trace Form Computations

5.1. Forms for $\mathfrak{sl}(2)$. Consider $L = \mathfrak{sl}(2)$ with basis e, h, f , where $e = e_{12}$, $h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = e_{11} - e_{22}$, $f = e_{21}$. Then $[e, f] = h$, $[h, e] = 2e$, $[h, f] = -2f$. To compute the concrete trace form we need the product of basis vectors, so $e^2 = 0$, $eh = -e$, $ef = e_{11}$, $he = e$, $h^2 = I$, $hf = -f$, $fe = e_{22}$, $fh = f$, $f^2 = 0$. The only product with nonzero

trace are ef, h^2, fe , and so the matrix of the concrete trace form is $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 0 \end{pmatrix}$. This is

nonsingular with eigenvalues $-1, 1, 2$. As a real matrix it has signature $(2, 1)$ and is an indefinite form.

Now we consider the Killing form, the trace form for $\text{ad } L$; we use the ordered basis e, h, f . We use $\text{ad } e(e) = 0$, $\text{ad } e(h) = -2e$, $\text{ad } e(f) = h$, $\text{ad } h(e) = 2e$, $\text{ad } h(h) = 0$, $\text{ad } h(f) = -2f$, $\text{ad } f(e) = -h$, $\text{ad } f(h) = 2f$, $\text{ad } f(f) = 0$. This gives us the matrices

$$\text{ad } e = \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \text{ad } h = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \text{ad } f = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}.$$

$$\text{gives } (\text{ad } e)^2 = \begin{pmatrix} 0 & 0 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, (\text{ad } e)(\text{ad } h) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix}, (\text{ad } e)(\text{ad } f) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$(\text{ad } h)^2 = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 4 \end{pmatrix}, (\text{ad } h)(\text{ad } f) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -4 & 0 \end{pmatrix}, (\text{ad } f)^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -2 & 0 & 0 \end{pmatrix}.$$

The matrix of the Killing form, obtained by taking the traces, is $\begin{pmatrix} 0 & 0 & 4 \\ 0 & 8 & 0 \\ 4 & 0 & 0 \end{pmatrix}$. This is 4 times the matrix

of the concrete trace form, and so it has the same properties. For example, over the reals, the form is indefinite.

5.2. Examples where the Killing form is zero. Let a be a nonzero scalar and let L be the Lie algebra spanned by x, y, z with brackets $[x, y] = x$, $[y, z] = az$, $[z, x] = 0$. Then L^1 is spanned by x, z , whence $L^2 = L^1$ and $L^{(2)} = 0$, so L is solvable but not nilpotent. We can easily see that $Z(L) = 0$. We have $\text{tr}(\text{ad } y)^2 = a^2 + 1$, and in fact, if

$a^2 = -1$, then the Killing form on L is identically zero, as we now show. Relative to the basis x, y, z , the matrices for the various adjoints are as follows. $\text{ad } x = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$,

$$\text{ad } y = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & a \end{pmatrix}, \text{ad } z = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -a & 0 \end{pmatrix}. \text{ Thus } (\text{ad } x)^2 = \text{ad } x \circ \text{ad } y = \text{ad } x \circ \text{ad } z = 0,$$

$$(\text{ad } y)^2 = \text{diag}(1, 0, a^2), \text{ad } y \circ \text{ad } z = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -a^2 & 0 \end{pmatrix}, (\text{ad } z)^2 = 0. \text{ This means the matrix}$$

$$\text{of the Killing form is } \begin{pmatrix} 0 & 0 & 0 \\ 0 & a^2 + 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \text{ If } a^2 = -1, \text{ the Killing form is identically zero.}$$

For an example over an arbitrary field (of characteristic 0), let b, c be nonzero scalars and let L be the Lie algebra spanned by x, y, z, w with brackets $[x, y] = x$, $[y, z] = bw$, $[y, w] = cz$, $[z, x] = [w, x] = [z, w] = 0$. Then L^1 is spanned by x, z, w , whence $L^2 = L^1$ and $L^{(2)} = 0$, so L is solvable but not nilpotent. We can easily see that $Z(L) = 0$. We have $\text{tr}(\text{ad } y)^2 = 2bc + 1$, and in fact, if $2bc = -1$, then the Killing form on L is identically zero, as we now show.

We can always choose such b, c (e.g., $b = 1, c = -\frac{1}{2}$), so we can create a non-nilpotent Lie algebra with zero Killing form. Relative to the basis x, y, z, w , the matrices

$$\text{for the various adjoints are as follows. } \text{ad } x = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \text{ad } y = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c \\ 0 & 0 & b & 0 \end{pmatrix},$$

$$\text{ad } z = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -b & 0 & 0 \end{pmatrix}, \text{ad } w = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -c & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \text{ Thus } (\text{ad } x)^2 = \text{ad } x \text{ad } y = \text{ad } x \text{ad } z =$$

$$\text{ad } x \text{ad } w = 0, (\text{ad } y)^2 = \text{diag}(1, 0, bc, bc), \text{ad } y \text{ad } z = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -bc & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \text{ad } y \text{ad } w =$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -bc & 0 & 0 \end{pmatrix}, (\text{ad } z)^2 = \text{ad } z \text{ad } w = (\text{ad } w)^2 = 0. \text{ This means the matrix of the Killing}$$

form is $\text{diag}(0, 2bc + 1, 0, 0)$ and thus the Killing form is identically 0 if $2bc = -1$. We can always find b, c making this true; for example, $b = 1, c = -\frac{1}{2}$ gives a zero Killing form.

6. Bases of L and $\text{ad } L$

The next result is really trivial, but we write it out.

Proposition 8.6.1 (bases for L and $\text{ad } L$).

Let $x_1, \dots, x_n \in L$ and let $\bar{x}_1, \dots, \bar{x}_n$ denote their cosets in $\bar{L} = L/Z(L)$.

- (1) $\bar{x}_1, \dots, \bar{x}_n$ are linearly independent if and only if $\text{ad } x_1, \dots, \text{ad } x_n$ are linearly independent.
- (2) $\bar{x}_1, \dots, \bar{x}_n$ span $\bar{L}$ if and only if $\text{ad } x_1, \dots, \text{ad } x_n$ span $\text{ad } L$.
- (3) $\bar{x}_1, \dots, \bar{x}_n$ is a basis of $\bar{L}$ if and only if $\text{ad } x_1, \dots, \text{ad } x_n$ is a basis of $\text{ad } L$.

PROOF. (1) Suppose first that $\sum_i \lambda_i \text{ad } x_i = 0$. This means that $\sum_i \lambda_i x_i \in \ker \text{ad} = Z(L)$, and so $\sum_i \lambda_i \bar{x}_i = 0$. If the $\bar{x}_i$ are linearly independent, this implies each $\lambda_i = 0$. If the $\bar{x}_i$ are not linearly independent, then we can choose not-all-zero λ_i making this true, and this in turn shows $\text{ad } x_1, \dots, \text{ad } x_n$ are linearly dependent.

(2) First suppose $\bar{x}_1, \dots, \bar{x}_n$ span $\bar{L}$ and $x \in L$. Then $\bar{x} = \sum_i \lambda_i \bar{x}_i$ for some scalars λ_i , and so $x - \sum_i \lambda_i x_i \in Z(L) = \ker \text{ad}$. Thus $\text{ad } x = \sum_i \lambda_i \text{ad } x_i$.

Suppose conversely that $\text{ad } x_1, \dots, \text{ad } x_n$ span $\text{ad } L$ and let $x \in L$. Then $\text{ad } x = \sum_i \lambda_i \text{ad } x_i$ for some scalars λ_i . Thus $x - \sum_i \lambda_i x_i \in \ker \text{ad} = Z(L)$, and so $\bar{x} = \sum_i \lambda_i \bar{x}_i$.

(3) This follows immediately from (1) and (2). ■

Note that if $Z(L) = 0$ and we realize L as a linear Lie algebra via a basis of L , then the concrete trace form of this realization is the Killing form on L . **Verify!**

As an example, let $L = \mathfrak{sl}(2)$ with the usual basis e, h, f . Then $\text{ad } L \cong L$ has basis $\text{ad } e = -2e_{12} + e_{31} = \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, $\text{ad } h = 2e_{11} - 2e_{33} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}$, $\text{ad } f = -e_{21} - 2e_{32} = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}$. This realizes $\mathfrak{sl}(2)$ as a subalgebra of $\mathfrak{sl}(3)$.

7. More on Bilinear Forms

7.1. Degenerate forms. Let V be a finite-dimensional vector space. We give some results for degenerate forms generalizing the results of Proposition 1.4.1 (Orthogonal complements, p. 35).

Let $\langle, \rangle$ be a symmetric or skew-symmetric bilinear form on V with radical K . Then $\dim W + \dim W^\perp = \dim V + \dim(W \cap K)$ and $(W^\perp)^\perp = W + K$. If $W \cap W^\perp = 0$, we

still have $V = W \oplus W^\perp$. In general, if the form restricted to W is nonsingular, we have $\dim W + \dim W^\perp = \dim V$. **Check.**

7.2. Invariant forms on some algebras. What are the invariant symmetric forms on the 2-dimensional non-Abelian Lie algebra L (basis x, y with $[x, y] = x$)? Let $\langle, \rangle$ be such a form. Then $0 = \langle [x, x], y \rangle = \langle x, [x, y] \rangle = \langle x, x \rangle$ and $0 = \langle [y, y], x \rangle = \langle y, [y, x] \rangle = -\langle y, x \rangle$. Thus we have $\langle ax + by, cx + dy \rangle = bd\lambda$, where $\lambda = \langle y, y \rangle$. This does yield an invariant symmetric form and if $\lambda \neq 0$, then the radical of $\langle, \rangle$ is $\mathbb{F}x$.

Note that for any $L, \langle, \rangle$, if $z \in Z(L), x, y \in L$, then $0 = \langle x, [y, z] \rangle = \langle [x, y], z \rangle$. This shows $Z(L) \subseteq (L^1)^\perp$.

If L is the Heisenberg Lie algebra (z central, $[x, y] = z$) and $\langle, \rangle$ is an invariant symmetric form on L , then the above shows $\langle z, z \rangle = 0$. We guess that $\langle, \rangle$ can be defined arbitrarily on $\mathbb{F}x + \mathbb{F}y$. **Check.**

Is there an example of a Lie algebra L with a nondegenerate symmetric invariant form $\langle, \rangle$ such that L is not semisimple? This would necessarily imply L has a nonzero Abelian ideal.

8. Derivations of the 2-dimensional and the Heisenberg Lie Algebras

If L is an Abelian Lie algebra, then any linear operator on L is a derivation, and only 0 is inner.

We have shown that all derivations of semisimple Lie algebras are inner.

Let L be the 2-dimensional non-Abelian Lie algebra with basis x, y and $[x, y] = x$. Then L is not semisimple, but we will show all derivations of L are inner.

Suppose $\delta : L \rightarrow L$ is a derivation and $\delta(x) = ax + cy, \delta(y) = bx + dy$. Then $ax + cy = \delta([x, y]) = [\delta(x), y] + [x, \delta(y)] = ax - dx$. Thus $c = d = 0$. Now $[bx - ay, x] = -a[y, x] = ax = \delta(x)$ and $[bx - ay, y] = bx = \delta(y)$. Thus $\delta = \text{ad}(bx - ay)$ is inner.

Let L be the Heisenberg Lie algebra with basis x, y, z, z central, and $[x, y] = z$. Then $\delta(z) = \delta([x, y]) = [x, \delta(y)] + [\delta(x), y]$. Since only multiples of z can come out of the bracket, it follows that $\delta(z) = \lambda z$ for some $\lambda \in \mathbb{F}$. Moreover, $0 = \delta([x, z]) = [\delta(x), z] + [x, \lambda z] = 0 + 0$ holds whatever $\delta(x)$ is. A similar calculation works for y , so $\delta(x), \delta(y)$ can be anything with one condition. Namely, if $\delta(x) = ax + by + cz, \delta(y) = dx + ey + fz$, we must have $a + d = \lambda$, i.e., $\delta(z) = (a + d)z$. Such a derivation is inner if and only if $a = b = d = e = 0$, and in fact, any inner derivation has the (unique) form $\text{ad}(fx - cy)$ for some scalars c, f .

9. Derivation Formulas

In this section, we state and verify some formulas involving derivations that are used elsewhere.

9.1. Powers of a derivation on a commutator. Let δ be a derivation on the Lie algebra L , let n be a positive integer, and let $x, y \in L$. Then

$$\delta^n([x, y]) = \sum_{i=0}^n \binom{n}{i} [\delta^i(x), \delta^{n-i}(y)].$$

We prove this by induction, the case $n = 1$ being the definition of the derivative. The proof is very much like the proof of the binomial theorem.

Suppose the formula holds for n . Then

$$\begin{aligned} \delta^{n+1}([x, y]) &= \sum_{i=0}^n \binom{n}{i} \delta([\delta^i(x), \delta^{n-i}(y)]) = \sum_{i=0}^n \binom{n}{i} [\delta^{i+1}(x), \delta^{n-i}(y)] + \sum_{i=0}^n \binom{n}{i} [\delta^i(x), \delta^{n-i+1}(y)] \\ &= (\text{put } j = i + 1 \text{ in first sum}) \sum_{j=1}^{n+1} \binom{n}{j-1} [\delta^j(x), \delta^{n+1-j}(y)] + \sum_{j=0}^n \binom{n}{j} [\delta^j(x), \delta^{n+1-j}(y)] \\ &= [x, \delta^{n+1}(y)] + \sum_{j=1}^n \left(\binom{n}{j-1} + \binom{n}{j} \right) [\delta^j(x), \delta^{n+1-j}(y)] + [\delta^{n+1}(x), y] \\ &= \sum_{j=0}^{n+1} \binom{n+1}{j} [\delta^j(x), \delta^{n+1-j}(y)]. \end{aligned}$$

9.2. Powers of a derivation minus scalars on a commutator. This is a generalization of the previous formula; it is needed in a particular argument.

Let δ be a derivation on the Lie algebra L , let n be a positive integer, let $x, y \in L$, let $\lambda, \mu \in \mathbb{F}$, and let I denote the identity operator on L . Then

$$(\delta - (\lambda + \mu)I)^n([x, y]) = \sum_{i=0}^n \binom{n}{i} [(\delta - \lambda I)^i(x), (\delta - \mu I)^{n-i}(y)].$$

When $n = 1$, this says $(\delta - (\lambda + \mu)I)([x, y]) = [(\delta - \lambda I)(x), y] + [x, (\delta - \mu I)(y)]$, which is true. The rest of the proof is like that of the previous formula, by induction.

9.3. Powers of ad in $\mathfrak{gl}(V)$. Let $L \subseteq \mathfrak{gl}(V)$ be a linear Lie algebra, let n be a nonnegative integer, and let $x, y \in L$. Then products involving x, y make sense as composition or matrix multiplication. We have

$$(\text{ad}_L x)^n(y) = \sum_{i=0}^n (-1)^{n-i} \binom{n}{i} x^i y x^{n-i}.$$

When $n = 0$, this says $\text{id}(y) = y$ and when $n = 1$, this says $\text{ad } x(y) = -yx + xy$. Both are correct. The general case is proved by induction on n , and we extend $\text{ad } x$ to $\mathfrak{gl}(V)$ with

the definition $(\operatorname{ad} x)(z) = xz - zx$.

$$\begin{aligned}
 (\operatorname{ad} x)^{n+1}(y) &= \sum_{i=0}^n (-1)^{n-i} \binom{n}{i} [x, x^i y x^{n-i}] = \sum_{i=0}^n (-1)^{n-i} \binom{n}{i} x^i [x, y] x^{n-i} \\
 &= \sum_{i=0}^n (-1)^{n-i} \binom{n}{i} x^i (xy - yx) x^{n-i} = \sum_{i=0}^n (-1)^{n-i} \binom{n}{i} x^{i+1} y x^{n-i} - \sum_{i=0}^n (-1)^{n-i} \binom{n}{i} x^i y x^{n-i+1} \\
 &= (\text{put } j = i + 1 \text{ in first sum}) \sum_{j=1}^{n+1} (-1)^{n+1-j} \binom{n}{j-1} x^j y x^{n+1-j} + \sum_{j=0}^n (-1)^{n+1-j} \binom{n}{j} x^j y x^{n+1-j} \\
 &= (-1)^{n+1} y x^{n+1} + \sum_{j=1}^n (-1)^{n+1-j} \left(\binom{n}{j-1} + \binom{n}{j} \right) x^j y x^{n+1-j} + x^{n+1} y \\
 &= \sum_{j=0}^{n+1} (-1)^{n+1-j} \binom{n+1}{j} x^j y x^{n+1-j}.
 \end{aligned}$$

10. Jordan-Chevalley Computations

If $x = [a] \in \mathfrak{gl}(1)$, then x is always semisimple and only nilpotent if $a = 0$, so $x_s = x$, $x_n = 0$. On the other hand, $\operatorname{ad} x = 0$ is both semisimple and nilpotent for any a . It seems the semisimple and nilpotent parts corresponding to $\operatorname{ad} x$ don't make sense. Semisimple and diagonalizable are the same in this case.

Now let $G = \mathfrak{gl}(2)$, $x = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. We know x is nilpotent if and only if $d = -a$ and $a^2 + bc = 0$. We know x is semisimple if and only if the discriminant of its characteristic polynomial is nonzero. This discriminant is $(\operatorname{tr} A)^2 - 4 \det A = (a + d)^2 - 4(ad - bc) = (a - d)^2 + 4bc$. Thus if $(a - d)^2 \neq -4bc$, we have $x_s = x$, $x_n = 0$.

Suppose $(a - d)^2 = -4bc$ and set $x_s = \begin{pmatrix} \frac{a+d}{2} & 0 \\ 0 & \frac{a+d}{2} \end{pmatrix}$, $x_n = \begin{pmatrix} \frac{a-d}{2} & b \\ c & -\frac{a-d}{2} \end{pmatrix}$. Plainly $x_s + x_n = x$ and x_s is a scalar matrix, so x_s, x_n commute and x_s is semisimple. To see x_n is nilpotent, note $\operatorname{tr} x_n = 0$ and $\det x_n = 0$, whence $x_n^2 = 0$. Note that the semisimple and nilpotent parts do not seem to behave continuously in the transition between semisimple and non-semisimple, although perhaps there is another way of thinking of continuity here.

Is there?

Let us take the basis $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ of $\mathfrak{gl}(2)$. Note e, h, f form a basis of $S = \mathfrak{sl}(2)$. As I is central and $\operatorname{ad} x$ maps $S = \mathfrak{sl}(2)$ into itself, the matrix of $\operatorname{ad}_G(x) = \begin{pmatrix} 0 & 0 \\ 0 & M \end{pmatrix}$, where M is the matrix of ad_S . Thus we just need to compute the matrix of ad_S to determine if $\operatorname{ad} x$ is semisimple or nilpotent.

$$\begin{aligned}\operatorname{ad} x(e) &= xe - ex = \begin{pmatrix} 0 & a \\ 0 & c \end{pmatrix} - \begin{pmatrix} c & d \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} -c & a-d \\ 0 & c \end{pmatrix} = (a-d)e - ch \\ \operatorname{ad} x(h) &= xh - hx = \begin{pmatrix} a & -b \\ c & -d \end{pmatrix} - \begin{pmatrix} a & b \\ -c & -d \end{pmatrix} = \begin{pmatrix} 0 & -2b \\ 2c & 0 \end{pmatrix} = -2be + 2cf \\ \operatorname{ad} x(f) &= xf - fx = \begin{pmatrix} b & 0 \\ d & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ a & b \end{pmatrix} = \begin{pmatrix} b & 0 \\ d-a & -b \end{pmatrix} = bh - (a-d)f.\end{aligned}$$

$$\text{Thus } \operatorname{ad}_S x = \begin{pmatrix} a-d & -2b & 0 \\ -c & 0 & b \\ 0 & 2c & -(a-d) \end{pmatrix} \text{ and so the characteristic polynomial of } \operatorname{ad}_S x$$

is

$$\begin{aligned}& \begin{vmatrix} \lambda - (a-d) & 2b & 0 \\ c & \lambda & -b \\ 0 & -2c & \lambda + (a-d) \end{vmatrix} = (\lambda - (a-d))(\lambda^2 + (a-d)\lambda - 2bc) - 2bc(\lambda + a-d) = \\ &= \lambda^3 - (a-d)\lambda^2 + (a-d)\lambda^2 - (a-d)^2\lambda - 2bc\lambda + 2bc(a-d) - 2bc\lambda - 2bc(a-d) \\ &= \lambda^3 - ((a-d)^2 + 4bc)\lambda = \lambda(\lambda^2 - ((a-d)^2 + 4bc)).\end{aligned}$$

It follows that $\operatorname{ad}_S x$ (and hence $\operatorname{ad}_G x$) is semisimple if and only if $(a-d)^2 + 4bc \neq 0$ and nilpotent if and only if $(a-d)^2 + 4bc = 0$. This is the same condition as the one for the semisimplicity of x , so $\operatorname{ad}_S x$, $\operatorname{ad}_G x$, x are semisimple for the same x . In addition, we see that $\operatorname{ad} x$ diagonalizable if and only if $(a-d)^2 + 4bc$ has a nonzero square root in $\mathbb{F}$ or $x = 0$. These are precisely the conditions under which x is diagonalizable. On the other hand, x is nilpotent if and only if $d = -a$ and $a^2 + bc = 0$, which translates to $a + d = 0$ and $(a-d)^2 + 4bc = 0$. Thus x nilpotent implies $\operatorname{ad} x$ nilpotent, but it is possible to have $\operatorname{ad} x$ nilpotent but x not nilpotent. Further analysis shows $\operatorname{ad} x$ is nilpotent if and only if $x = \lambda I + y$ for y nilpotent. If $x \in \mathfrak{sl}(2)$, then $\operatorname{ad} x$ is nilpotent if and only if x is nilpotent.

We wish to write $x = s + n$, where $\operatorname{ad} s$ semisimple, $\operatorname{ad} n$ nilpotent, $\operatorname{ad} s$, $\operatorname{ad} n$ commuting (preferably, s, n commuting). When x is semisimple, we can take $s = x$, $n = 0$. But we can also write $s = x - \lambda I$, $n = \lambda I$ and all conditions, including $sn = ns$, are satisfied. When x is nilpotent, we can take $s = 0$, $n = x$ or more generally $s = \lambda I$, $n = x - \lambda I$. If we want to choose $s, n \in \mathfrak{sl}(2)$, then s, n are unique. **Check!**

In general, we can write $x = x_s + x_n$ with x_s semisimple and x_n nilpotent and $[x_s, x_n] = 0$ as in previous paragraphs, and then $\operatorname{ad} x_s$ is semisimple and $\operatorname{ad} x_n$ nilpotent.

11. Lie algebra Plus its Dual

Let L be a Lie algebra and let L^* denote its dual space. We can make L^* into an L -module via $x.\phi = -\phi \circ \text{ad } x$, that is $(x.\phi)(y) = -\phi([x, y])$. This is a representation, since $x.(y.\phi) - y.(x.\phi) = -(y.\phi) \circ \text{ad } x + (x.\phi) \circ \text{ad } y = \phi \circ \text{ad } y \circ \text{ad } x - \phi \circ \text{ad } x \circ \text{ad } y = \phi \circ ([\text{ad } y, \text{ad } x]) = -\phi \circ \text{ad } [x, y] = [x, y].\phi$.

As in Proposition 3.1.11 (Extension of a Lie algebra by a module, p.95), we can make $M = L \oplus L^*$ into a Lie algebra by defining $[(x, \phi), (y, \psi)] = ([x, y], x.\psi - y.\phi)$. We will identify L^* with $0 \oplus L^*$ and L with $L \oplus 0$ henceforth.

We can define a symmetric bilinear form $\langle, \rangle$ on M by $\langle (x, \phi), (y, \psi) \rangle = \phi(y) + \psi(x)$. This form is invariant:

$$\begin{aligned} \langle [(x, \phi), (y, \psi)], (z, \theta) \rangle &= \langle ([x, y], x.\psi - y.\phi), (z, \theta) \rangle = x.\psi(z) - y.\phi(z) + \theta([x, y]) = \\ &= -\psi([x, z]) + \phi([y, z]) + \theta([x, y]) = \theta([x, y]) + \phi([y, z]) + \psi([z, x]) \\ \langle (x, \phi), [(y, \psi), (z, \theta)] \rangle &= \langle (x, \phi), ([y, z], y.\theta - z.\psi) \rangle = (y.\theta)(x) - (z.\psi)(x) + \phi([y, z]) = \\ &= -\theta([y, x]) + \psi([z, x]) + \phi([y, z]) = \theta([x, y]) + \phi([y, z]) + \psi([z, x]). \end{aligned}$$

We will now show the form is nondegenerate by computing the orthogonal complements of L^* and L . First note $(x, \phi) \in (L^*)^\perp$ if and only if for any $\psi \in L^*$, we have $0 = \langle (x, \phi), (0, \psi) \rangle = \psi(x)$. This means $(x, \phi) \in (L^*)^\perp$ if and only if $x = 0$, i.e., $(L^*)^\perp = L$. Next note that $(x, \phi) \in L^\perp$ if and only if for any $y \in L$, we have $0 = \langle (x, \phi), (y, 0) \rangle = \phi(y)$. This means $(x, \phi) \in L^\perp$ if and only if $\phi = 0$, i.e., $L^\perp = L$. It follows that the radical of $\langle, \rangle$ is $(L^*)^\perp \cap L^\perp = 0$. Thus $\langle, \rangle$ is an invariant nondegenerate symmetric form on M .

Another way to see the form is invariant is as follows. Let $x_1, \dots, x_n$ be a basis for L and let $X_1, \dots, X_n$ be the dual basis for L^* , i.e., $X_i(x_j) = \delta_{ij}$. Identify X_i with $(0, X_i)$ and x_j with $(x_j, 0)$, so that $\langle X_i, x_j \rangle = \delta_{ij}$, $\langle X_i, X_j \rangle = 0$, $\langle x_i, x_j \rangle = 0$ for all i, j . If we take $X_1, \dots, X_n, x_1, \dots, x_n$ as a basis for M , then the matrix of $\langle, \rangle$ with respect to this basis is $\begin{pmatrix} 0 & I_n \\ I_n & 0 \end{pmatrix}$. Thus $\langle, \rangle$ is nondegenerate. Moreover, $(L^*)^\perp = L$ and $L^\perp = L$.

Note that L^* is an ideal of M , while L is only a subalgebra. In fact, L is an ideal if and only if L is Abelian, in which case M is Abelian. Note that $(x, \phi) \in Z(M)$ if and only if $(0, 0) = ([x, y], x.\psi - y.\phi)$ for all $y \in L, \psi \in L^*$. This requires $x \in Z(L)$, which implies $x.\psi = 0$. We have $y.\phi = 0$ for all y if and only if ϕ vanishes on $[L, L]$. Thus $Z(M) = \{ (x, \phi) \mid x \in Z(L) \text{ and } \phi([L, L]) = 0 \}$. Thus $Z(M) \cong (L/[L, L])^* \oplus Z(L)$ and $M/Z(M) \cong (L/[L, L])^* \oplus L/Z(L)$. **Check - is it $[L, L]^*$?**

We compute $[M, M]$ in pieces. First, we note $[L^*, L^*] = 0$ and $[L, L]$ in M is $[L, L] \oplus 0$, and $[L^*, L]$ is spanned by elements of the form $[(0, \phi), (y, 0)] = -y.\phi$. Thus $[L^*, L]$ is spanned by elements of the form $(0, \phi \circ \text{ad } y)$, where $\phi \in L^*, y \in L$. How can we describe

the space of morphisms $\phi \circ \text{ad } y$? **Can we?** When L is Abelian, this is only the 0 morphism. In the first example below, this space is all of L^* . In the second example, this space is spanned by X, Y , so it is $Z(M) \cap L^* = [M, M] \cap L^*$.

We consider two specific examples. First take L be the 2-dimensional non-Abelian Lie algebra with basis x, y and $[x, y] = x$. Define X, Y in L^* to be the dual basis, so $X(x) = 1, X(y) = 0$ and $Y(x) = 0, Y(y) = 1$. Identify X with $(0, X)$, x with $(x, 0)$, and so on. Then $M = L^* \oplus L$ is a 4-dimensional Lie algebra with basis X, Y, x, y . To compute the Lie bracket, we note $x.X(x) = -X([x, x]) = 0, x.X(y) = -X([x, y]) = -1$, so $x.X = -Y$. Likewise, $x.Y = 0, y.X = X, y.Y = 0$. We have $[X, Y] = 0$ and $[x, y] = x$. Further, $[X, x] = (0, -x.X) = Y, [X, y] = (0, -y.X) = -X, [Y, x] = -(0, x.Y) = 0$, and $[Y, y] = (0, -y.Y) = 0$.

Thus we see $[M, M]$ is spanned by x, X, Y . Thus $M^2 = [M, M]$, while $M^{(2)}$ is spanned by Y . It follows that M is solvable but not nilpotent. The center of M is $\mathbb{F}Y$; to see this, let $z = \alpha X + \beta Y + \gamma x + \delta y \in Z(M)$. Then $0 = [X, z] = \gamma Y - \delta X$, so $\gamma = \delta = 0$. Also $0 = [x, z] = -\alpha Y + \delta x$, so $\alpha = 0$. It's not hard to see that $Z_2(M) = Z(M)$. Clearly $[M, M]^\perp = M^{(2)}$.

Next take L to be the Heisenberg Lie algebra with basis x, y, z , z central and $[x, y] = z$. Let X, Y, Z be the basis of L^* dual to x, y, z , and use the identification above to make X, Y, Z, x, y, z the basis of $M = L^* \oplus L$. The brackets of X, Y, Z are all zero, as are those of x, y, z except $[x, y] = z$. To compute the other brackets, we note that the span of $\text{ad } x$ and $\text{ad } y$ are $\mathbb{F}z$. Thus $x.X = y.X = x.Y = y.Y = 0$, while $x.Z(y) = -Z([x, y]) = -1$, so $x.Z = -Y$. Likewise $y.Z = X$ and z being central implies $z.X = z.Y = z.Z = 0$. Thus $[X, x] = [Y, x] = [X, y] = [Y, y] = 0$ and $[Z, x] = Y$ and $[Z, y] = -X$. Moreover $[X, z] = [Y, z] = [Z, z] = 0$. It follows that X, Y, z are all central, and in fact they span $Z(M)$. To see this, let $r = \alpha x + \beta y + \gamma Z \in Z(M)$. Then $0 = [Z, r] = -\alpha Y - \beta X$, so $\alpha = \beta = 0$ and $0 = [x, r] = \beta z - \gamma Y$, so $\gamma = 0$.

We also have that $[M, M]$ is spanned by X, Y, z , and so $[M, M] = Z(M)$. Thus M is nilpotent. Further, $[M, M]^\perp = [M, M]$.

12. The Transporter

If A, B are subspaces of the Lie algebra L , we define $(A : B) = \{x \in L \mid [x, B] \subseteq A\}$. This construction is common in commutative ring theory, where A, B are usually ideals of a ring. It is called the **ideal quotient** of A, B or the **colon ideal**. The related construction where A, B are rings and $A \subseteq B$ and A replaces L is called a the **conductor** of A in B . For Lie algebras, [Bourbaki 1-3, Chapter 1, Exercise 1, p. 73] calls $(A : B)$ the **transporter** of B into A . We will use this the term “transporter”.

Many constructions are special cases of the transporter. For example, $Z(L) = (0 : L)$. We can generalize this in two directions: in the ascending central series, $Z_{k+1} = (Z_k : L)$ and for any subspace B , its centralizer $C(B)$ is $(0 : B)$. The normalizer of any subalgebra A is $N(A) = (A : A)$, and A is an ideal if and only if $(A : A) = L$.

What kind of object is $(A : B)$? Clearly it is a subspace. Consider the 2-dimensional Lie algebra L with $[x, y] = x$ and let $A = \mathbb{F}y$. This is a subalgebra (but not an ideal) of L . Clearly $(A : 0) = L$, $(A : A) = A$, $(A : L) = 0$. In particular, $(A : A)$ is a subalgebra but not an ideal. On the other hand, if $A = \mathbb{F}x$, then $(A : 0) = (A : A) = (A : L) = L$.

Next, consider $L = \mathfrak{sl}(2)$ with the usual basis e, h, f . Set $A = \mathbb{F}h$ and $B = \mathbb{F}e + \mathbb{F}f$. Note that A is a subalgebra and B is not, and $(A : B) = B$ is therefore not a subalgebra.

The next proposition lists some properties of the transporter construction and gives some sufficient conditions for $(A : B)$ to be a subalgebra or an ideal.

Proposition 8.12.1 (Transporter properties).

Let L be a Lie algebra and let A, B be subspaces of L .

- (1) *If A', B' are subspaces of L with $A' \subseteq A$ and $B \subseteq B'$, then $(A' : B') \subseteq (A : B)$.*
- (2) *If C is a subspace of L , then $(A : B + C) = (A : B) \cap (A : C)$ and $(A \cap C : B) = (A : B) \cap (C : B)$.*
- (3) *If B is an ideal of L or $A \subseteq B$, then $(A : B)$ is a subalgebra of L .*
- (4) *If A, B are ideals of L , then $(A : B)$ is an ideal of L and $A \cap B \subseteq (A : B)$.*

PROOF. (1) Let $x \in (A' : B')$. Then $[x, B] \subseteq [x, B'] \subseteq A' \subseteq A$.

(2) Suppose $[x, B + C] \subseteq A$: then certainly $[x, B] \subseteq A$ and $[x, B] \subseteq C$, so $(A : B + C) \subseteq (A : B) \cap (A : C)$. Suppose $x \in (A : B) \cap (A : C)$: then $[x, B + C] = [x, B] + [x, C] \subseteq A + A = A$, so $(A : B) \cap (A : C) \subseteq (A : B + C)$.

Suppose $x \in (A \cap C : B)$, i.e., $[x, B] \subseteq A \cap C$. Certainly, then, $x \in (A : B) \cap (C : B)$. Conversely, suppose $x \in (A : B) \cap (C : B)$, so $[x, B] \subseteq A$ and $[x, B] \subseteq C$. Then $[x, B] \subseteq A \cap C$, so $x \in (A \cap C : B)$.

(3) & (4). Let $x, y \in L, b \in B$. Then $[[x, y], b] = [[x, b], y] + [x, [y, b]]$.

(3) If B is an ideal of L , then $[x, b], [y, b] \in B$, so if $x, y \in (A : B)$, we have $[[x, y], b] \in A$. If $A \subseteq B$ and $x, y \in (A : B)$, then $[x, b], [y, b] \in A \subseteq B$, so $[[x, b], y], [x, [y, b]] \in A$. Thus $[[x, y], b] \in A$.

(4) Now suppose $x \in L, y \in (A : B)$, so $[y, b] \in A$. If A is an ideal of L , then $[x, [y, b]] \in A$. If B is an ideal of L , then $[x, b] \in B$ and so $[[x, b], y] \in A$. Thus $[[x, y], b] \in A$. ■

We can generalize this notion. Let V be an L -module and A, B be subspaces of V . We define $(A : B) = \{x \in L \mid x.b \subseteq A \text{ for all } b \in B\}$. When A, B are subspaces of L , the old definition of $(A : B)$ agrees with the new one when we regard L as an $\text{ad } L$ -module.

Proposition 8.12.2 (Transporter properties for modules).

Let L be a Lie algebra, let V be an L -module, and let A, B be subspaces of V .

- (1) *If A', B' are subspaces of V with $A' \subseteq A$ and $B \subseteq B'$, then $(A' : B') \subseteq (A : B)$.*
- (2) *If C is a subspace of V , then $(A : B + C) = (A : B) \cap (A : C)$ and $(A \cap C : B) = (A : B) \cap (C : B)$.*
- (3) *If B is a submodule of V or $A \subseteq B$, then $(A : B)$ is a subalgebra of L .*
- (4) *If A, B are submodules of V , then $(A : B)$ is an ideal of L .*

REFERENCES AND REMARKS. We would like an example of a Lie algebra L and a subalgebra S such that $(S : L)$ is not an ideal. Is there such an example? ??

PROOF. (1), (2) These are proved just as (1) & (2) of Proposition 8.12.1 (Transporter properties, p. 190).

(3) & (4) Let $x, y \in L, b \in B$. Then $[x, y].b = x.y.b - y.x.b$.

(3) If B is a submodule of V , then $x.b, y.b \in B$. Thus if $x, y \in (A : B)$, we have $y.x.b, x.y.b \in A$, and so $x.y.b - y.x.b \in A$. If $A \subseteq B$ and $x, y \in (A : B)$, then $x.b, y.b \in A \subseteq B$, so $y.x.b, x.y.b \in A$. Thus $x.y.b - y.x.b \in A$.

(4) Now suppose $x \in L, y \in (A : B)$, so $y.b \in A$. If A is a submodule of V , then $x.y.b \in A$. If B is a submodule of V , then $x.b \in B$ and so $y.x.b \in A$. Thus $x.y.b - y.x.b \in A$. ■

13. $\mathfrak{sl}(2)$ Representations

We give some examples in small dimension and then we fill in some details from Section 3 (Representations of $\mathfrak{sl}(2)$, p. 103) in Chapter 3 (Representations, p. 83).

13.1. Small dimension $\mathfrak{sl}(2)$ representations. Let $L = \mathfrak{sl}(2)$. Since $L = [L, L]$, we know the only 1-dimensional representation is the trivial one. Consider the usual basis $e = e_{12}$, $h = \text{diag}(1, -1)$, $f = e_{21}$ of L and let $V = \mathbb{F}^2$ with basis $v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $w = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Then V is an L -module (the natural representation of L) with the following relations:

$$e.v = 0 \quad e.w = v \quad h.v = v \quad h.w = -w \quad f.v = w \quad f.w = 0.$$

This is the 2-dimensional irreducible representation of $\mathfrak{sl}(2)$, as described in Section 3 (Representations of $\mathfrak{sl}(2)$, p. 103) in Chapter 3 (Representations, p. 83). Note that v, w are h -eigenvectors with eigenvalues $1, -1$ respectively. Note e kills v and f kills w .

Next we consider the 2-dimensional contragredient (dual) representation V^* , which will turn out to be isomorphic to V . For V^* , we have a basis ϕ, ψ , where $\phi(v) = -1$, $\phi(w) = \psi(v) = 0$, $\psi(w) = 1$. We wish to compute the actions of elements of L on ϕ, ψ . For example $e.\phi(v) = -\phi(e.v) = 0$, $e.\phi(w) = -\phi(e.w) = -\phi(v) = 1$, whence $e.\phi = \psi$. Likewise $h.\psi(v) = -\psi(h.v) = -\psi(v) = 0$, $h.\psi(w) = -\psi(h.w) = 1$, whence $h.\psi = \psi$. Also, $f.\psi(v) = -\psi(f.v) = -\psi(w) = 0$, $f.\psi(w) = -\psi(f.w) = -\psi(w) = -1$, whence $f.\psi = \phi$. The full relations are:

$$e.\psi = 0 \quad e.\phi = \psi \quad h.\psi = \psi \quad h.\phi = -\phi \quad f.\psi = \phi \quad f.\phi = 0.$$

These are the same relations (replacing v by ψ and w by ϕ) as above, so $V^* \cong V$.

We next consider the L -module $V \otimes V$ with basis $a = v \otimes v$, $b = v \otimes w + w \otimes v$, $c = v \otimes w - w \otimes v$, $d = w \otimes w$. We wish to compute the actions of elements of L on a, b, c, d . For example, $e.a = e.v \otimes v + v \otimes e.v = 0$, $e.b = e.v \otimes w + v \otimes e.w + e.w \otimes v + w \otimes e.v = 2v \otimes v = 2a$, $e.c = e.v \otimes w + v \otimes e.w - e.w \otimes v - w \otimes e.v = 0$, $e.d = e.w \otimes w + w \otimes e.w = v \otimes w + w \otimes v = b$. Likewise, $h.a = h.v \otimes v + v \otimes h.v = 2v \otimes v = 2a$, $h.b = h.v \otimes w + v \otimes h.w + h.w \otimes v + w \otimes h.v = v \otimes w - v \otimes w - w \otimes v + w \otimes v = 0$, $h.c = 0$, $h.d = h.w \otimes w + w \otimes h.w = -2w \otimes w = -2d$. Finally, $f.a = f.v \otimes v + v \otimes f.v = w \otimes v + v \otimes w = b$, $f.b = f.v \otimes w + v \otimes f.w + f.w \otimes v + w \otimes f.v = w \otimes w + w \otimes w = 2d$, $f.c = 0$, $f.d = f.w \otimes w + w \otimes f.w = 0$.

Thus c is killed by all of e, h, f , so $\mathbb{F}c$ is a 1-dimensional trivial submodule. The relations for a, b, d are:

$$e.a = 0 \quad e.b = 2a \quad e.d = b \quad h.a = 2a \quad h.b = 0 \quad h.d = -2d \quad f.a = b \quad f.b = 2d \quad f.d = 0.$$

This is a 3-dimensional submodule of $V \otimes V$. If we order the basis d, b, a , the basis vectors are h -eigenvectors with eigenvalues $-2, 0, 2$ respectively. Since the eigenvalues are distinct, it isn't hard to see that this submodule is irreducible. (See Subsection 13.3 (Irreducibility of V_n , p. 193) below.) Note that acting by f moves the basis vectors down ($a \mapsto b$, $b \mapsto 2d$, $d \mapsto 0$), while the action of e moves the basis vectors up ($d \mapsto b$, $b \mapsto 2a$, $a \mapsto 0$).

This 3-dimensional representation is isomorphic to the adjoint representation of $\mathfrak{sl}(2)$. To see this, regard a as $-e$, b as h , and d as f and see that the relations above correspond to the adjoint representation (e.g., $e.b = 2a$ becomes $\text{ad } e(h) = -2e = 2a$).

13.2. E, F, H relations. Recall from Section 3 (Representations of $\mathfrak{sl}(2)$, p. 103) that $E = x \frac{\partial}{\partial y}$, $H = x \frac{\partial}{\partial x} - y \frac{\partial}{\partial y}$, $F = y \frac{\partial}{\partial x}$ are linear operators acting on $\mathbb{F}[x, y]$. We claim $EF - FE = H$, $HE - EH = 2E$, $HF - FH = -2F$.

We will show this in two different ways. First note that $x, \frac{\partial}{\partial y}$ commute as do $y, \frac{\partial}{\partial x}$, while the product rule tells us $\frac{\partial}{\partial x}x = x \frac{\partial}{\partial x} + 1$ and $\frac{\partial}{\partial y}y = y \frac{\partial}{\partial y} + 1$.

Thus

$$\begin{aligned}
EF - FE &= x \frac{\partial}{\partial y} y \frac{\partial}{\partial x} - y \frac{\partial}{\partial x} x \frac{\partial}{\partial y} = x(y \frac{\partial}{\partial y} + 1) \frac{\partial}{\partial x} - y(x \frac{\partial}{\partial x} + 1) \frac{\partial}{\partial y} \\
&= xy \frac{\partial^2}{\partial x \partial y} + x \frac{\partial}{\partial x} - xy \frac{\partial^2}{\partial x \partial y} - y \frac{\partial}{\partial y} = x \frac{\partial}{\partial x} - y \frac{\partial}{\partial y} = H \\
HE - EH &= x \frac{\partial}{\partial x} x \frac{\partial}{\partial y} - y \frac{\partial}{\partial y} x \frac{\partial}{\partial y} - \left(x \frac{\partial}{\partial y} x \frac{\partial}{\partial x} - x \frac{\partial}{\partial y} y \frac{\partial}{\partial y} \right) = \\
&= x^2 \frac{\partial^2}{\partial x \partial y} + x \frac{\partial}{\partial y} - xy \frac{\partial^2}{\partial y^2} - x^2 \frac{\partial^2}{\partial x \partial y} + xy \frac{\partial^2}{\partial y^2} + x \frac{\partial}{\partial y} = 2x \frac{\partial}{\partial y} = 2E \\
HF - FH &= x \frac{\partial}{\partial x} y \frac{\partial}{\partial x} - y \frac{\partial}{\partial y} y \frac{\partial}{\partial x} - \left(y \frac{\partial}{\partial x} x \frac{\partial}{\partial x} - y \frac{\partial}{\partial x} y \frac{\partial}{\partial y} \right) = \\
&= xy \frac{\partial^2}{\partial x^2} - y \frac{\partial}{\partial x} - y^2 \frac{\partial^2}{\partial x \partial y} - y \frac{\partial}{\partial x} - xy \frac{\partial^2}{\partial x^2} + y^2 \frac{\partial^2}{\partial x \partial y} = -2y \frac{\partial}{\partial x} = -2F
\end{aligned}$$

For the second proof, note

$$E(x^i y^j) = j x^{i+1} y^{j-1} \quad H(x^i y^j) = (i - j) x^i y^j \quad F(x^i y^j) = i x^{i-1} y^{j+1}.$$

Thus

$$\begin{aligned}
(EF - FE)(x^i y^j) &= (j + 1) i x^i y^j - (i + 1) j x^i y^j = (i - j) x^i y^j = H(x^i y^j) \\
(HE - EH)(x^i y^j) &= j((i + 1) - (j - 1)) x^{i+1} y^{j-1} - (i - j) j x^{i+1} y^{j-1} = 2j x^{i+1} y^{j-1} = 2E(x^i y^j) \\
(HF - FH)(x^i y^j) &= i(i - 1 - (j + 1)) x^{i-1} y^{j+1} - i(i - j) x^{i-1} y^{j+1} = -2i x^{i-1} y^{j+1} = -2F(x^i y^j)
\end{aligned}$$

13.3. Irreducibility of V_n . Recall that V_n is the span of $x^n, x^{n-1}y, \dots, y$, with E, F, H acting in their defined way. We now prove that V_n is irreducible. In this proof, “monomial” will mean a scalar multiple of $x^i y^j$ for some i, j . This monomial is a h -eigenvector with eigenvalue $i - j$. Since $i + j = n$, these eigenvalues are distinct. Suppose W is a nonzero submodule of V_n : by Proposition 1.1.5 (Consequences of distinct eigenvalues, p. 22), W must contain a monomial $x^i y^j$. Repeatedly applying the operator E to $x^i y^j$ will give all monomials with smaller j and repeatedly applying the operator F will give all monomials with larger j . Thus repeatedly applying E and F will give us any monomial in V_n , and so v will generate V_n . Thus we see $W = V_n$.

14. Dynkin Diagrams and Coxeter Graphs

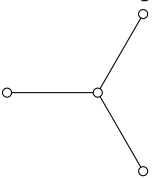
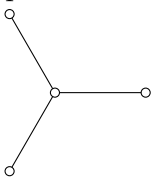
In this section, we list the standard Dynkin diagrams and Coxeter graphs for a general number of nodes, and then we list the Dynkin diagrams for small numbers of nodes.

For the L^AT_EX code used in creating these diagrams, see Subsection 14.2 (The LaTeX code, p. 202). They were all created using the package `dynkin-diagrams`.

14.1. Dynkin and Coxeter diagrams. These are the standard irreducible Dynkin diagrams and Coxeter graphs. For the $\text{\LaTeX}$ code used in creating these diagrams, see the next subsection. They were all created using the package `dynkin-diagrams`.

We begin by listing the general diagrams for each type, and then we list the Dynkin diagrams for small numbers of nodes. In all cases, ℓ is the number of nodes (vertices, simple roots).

Some Coxeter graphs do not occur as Dynkin diagrams (for simple root systems, such as types H and I). There are many variations of diagrams that do not change their meaning. For example G_2 can appear as both $\circ \rightleftharpoons \circ$ and $\circ \leftrightharpoons \circ$, and D_4 can appear

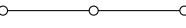
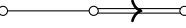
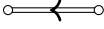
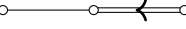
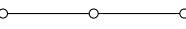
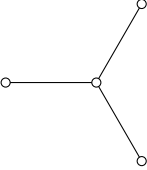
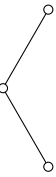
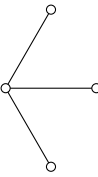
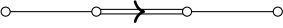
as both  and . There are many other variations on this theme. Note that H_2 and $I_2(5)$ are the same Coxeter graph, as are G_2 and $I_2(6)$.

Type	Dynkin Diagram	Coxeter Diagram
A_ℓ		
B_ℓ		
C_ℓ		
D_ℓ		
E_6		
E_7		
E_8		
F_4		
G_2		
H_2		
H_3		
H_4		
$I_2(n)$		

We next give the Dynkin diagrams with $\ell = 1, 2, 3, 4$ nodes. We give two examples of the diagrams of type D : the diagrams may look different, but they are isomorphic. The diagram for D_2 is the only one that is not connected; it is isomorphic to $A_1 \times A_1$ (or $D_1 \times D_1$). (The `dynkin-diagram` package gives an incorrect diagram for D_2 .)

Several types give the same graph for a small number of nodes. Clearly $A_1 = B_1 = C_1 = D_1$. Also B_2 and C_2 are isomorphic, as are A_3 and D_3 . Because of these isomorphisms, it is common to require $\ell \geq 2$ for B_ℓ , $\ell \geq 3$ for C_ℓ , and $\ell \geq 4$ for D_ℓ . There

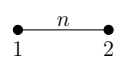
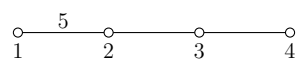
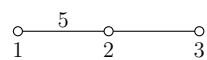
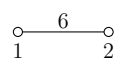
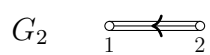
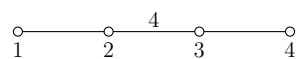
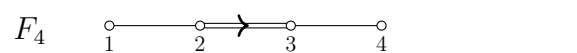
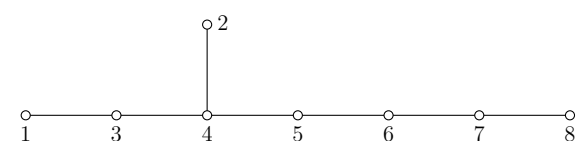
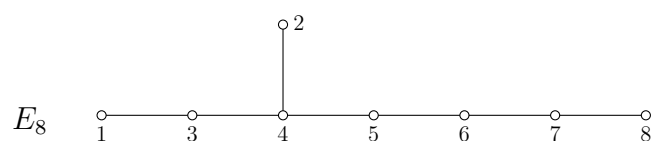
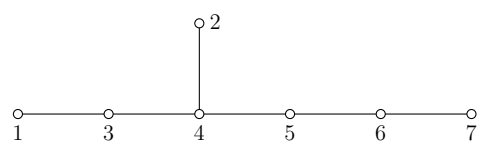
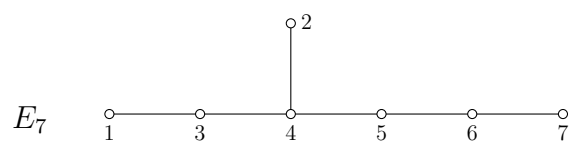
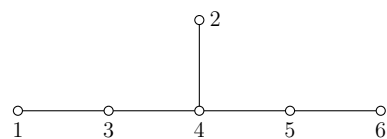
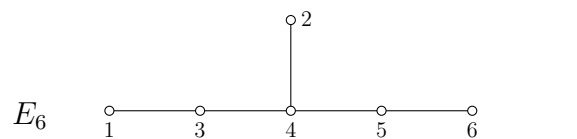
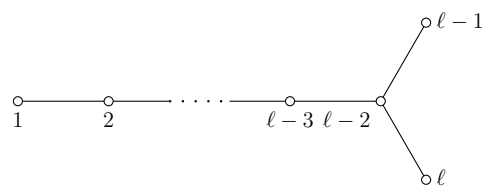
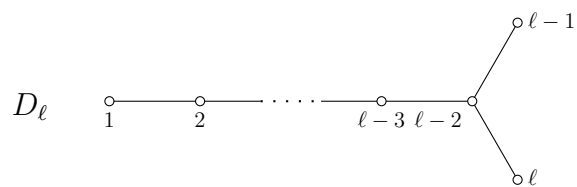
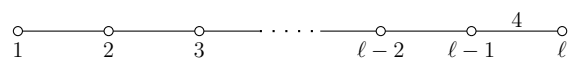
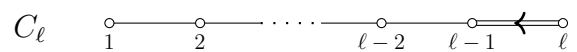
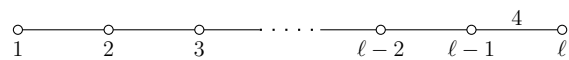
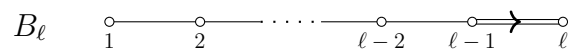
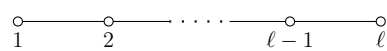
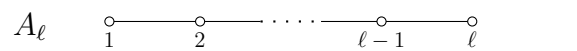
are also diagrams for E_ℓ for $\ell \leq 5$, but they are usually omitted since they are isomorphic to other diagrams.

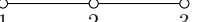
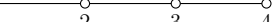
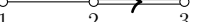
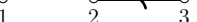
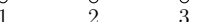
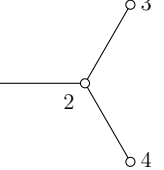
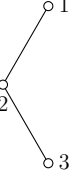
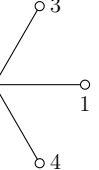
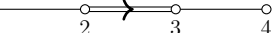
Type	$\ell =$	1	2	3	4
A_ℓ					
B_ℓ					
C_ℓ					
D_ℓ					
D_ℓ alt					
F_ℓ					
G_ℓ					

Next we number the nodes (vertices, roots), as in [Humphreys, §11.4, p. 58].

Type Dynkin Diagram

Coxeter Diagram

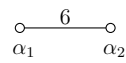
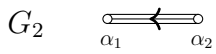
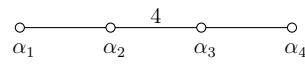
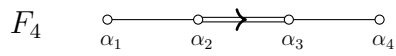
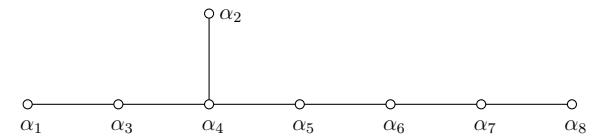
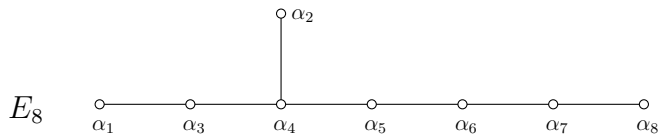
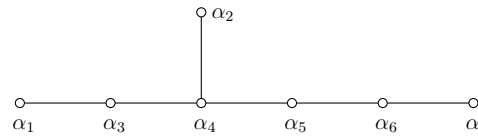
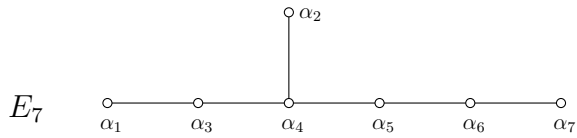
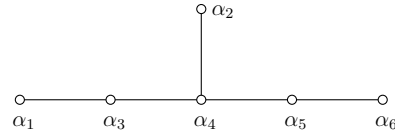
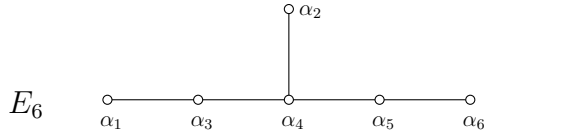
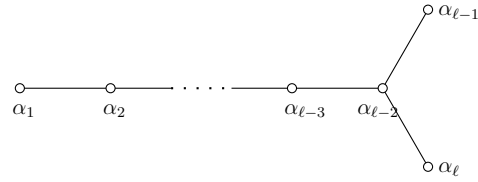
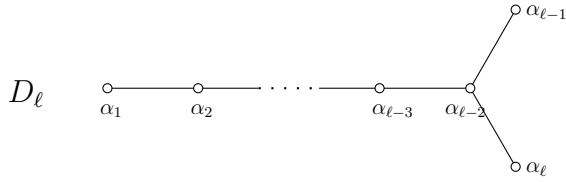
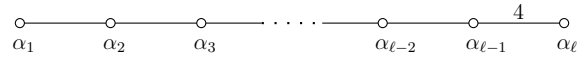
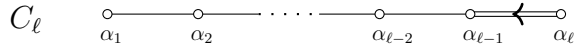
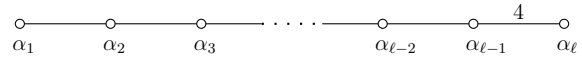
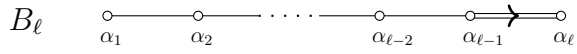
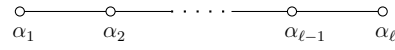
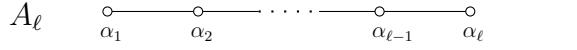
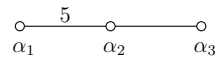
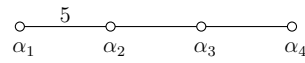
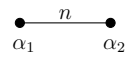


Type	$\ell = 1$	2	3	4
A_ℓ				
B_ℓ				
C_ℓ				
D_ℓ				
D_ℓ alt				
F_ℓ				
G_ℓ				

This time we label the nodes (vertices, roots) with α_i instead of i .

Type Dynkin Diagram

Coxeter Diagram

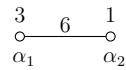
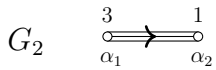
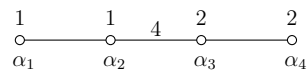
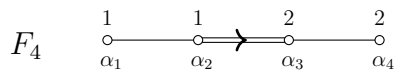
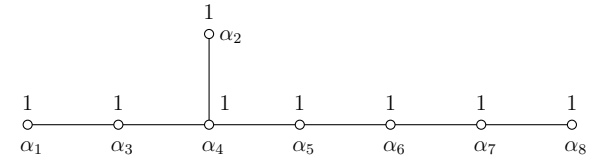
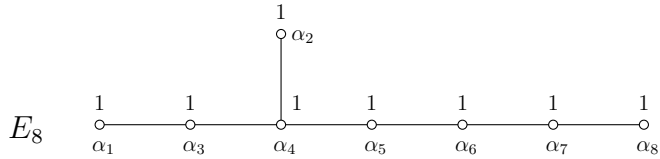
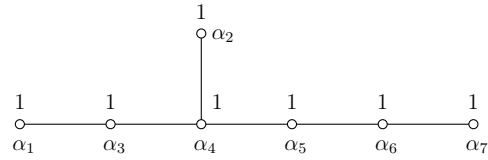
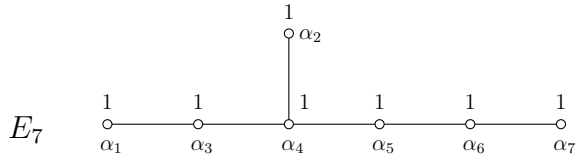
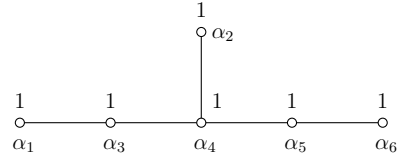
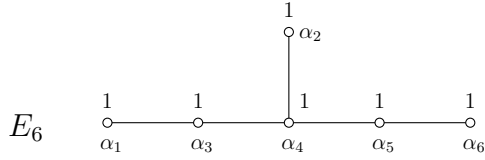
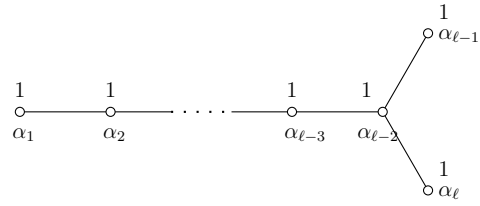
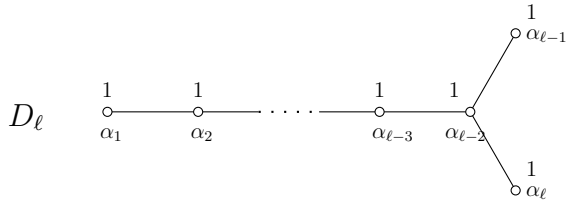
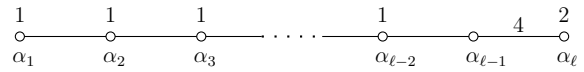
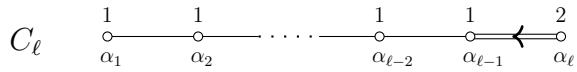
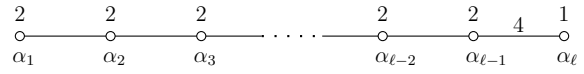
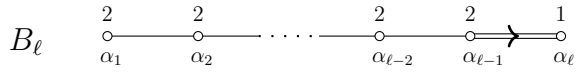
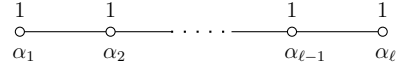
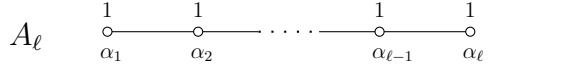
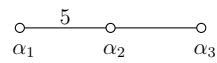
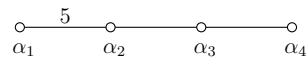
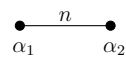

 H_3

 H_4

 $I_2(n)$


Type	$\ell = 1$	2	3	4
A_ℓ	$\circ$ α_1	$\circ \text{---} \circ$ $\alpha_1 \quad \alpha_2$	$\circ \text{---} \circ \text{---} \circ$ $\alpha_1 \quad \alpha_2 \quad \alpha_3$	$\circ \text{---} \circ \text{---} \circ \text{---} \circ$ $\alpha_1 \quad \alpha_2 \quad \alpha_3 \quad \alpha_4$
B_ℓ	$\circ$ α_1	$\circ \text{---} \circ$ $\alpha_1 \quad \alpha_2$	$\circ \text{---} \circ \text{---} \circ$ $\alpha_1 \quad \alpha_2 \quad \alpha_3$	$\circ \text{---} \circ \text{---} \circ \text{---} \circ$ $\alpha_1 \quad \alpha_2 \quad \alpha_3 \quad \alpha_4$
C_ℓ	$\circ$ α_1	$\circ \text{---} \circ$ $\alpha_1 \quad \alpha_2$	$\circ \text{---} \circ \text{---} \circ$ $\alpha_1 \quad \alpha_2 \quad \alpha_3$	$\circ \text{---} \circ \text{---} \circ \text{---} \circ$ $\alpha_1 \quad \alpha_2 \quad \alpha_3 \quad \alpha_4$
D_ℓ	$\circ$ α_1	$\alpha_1 \quad \alpha_2$ $\circ \quad \circ$	$\circ \text{---} \circ \text{---} \circ$ $\alpha_1 \quad \alpha_2 \quad \alpha_3$	$\circ \text{---} \circ$ $\alpha_1 \quad \alpha_2$ $\alpha_3 \quad \alpha_4$
$D_\ell \text{ alt}$	$\circ$ α_1	$\circ$ α_2 $\circ$ α_1	$\circ$ α_1 α_2 α_3	$\circ$ α_3 α_2 α_1 α_4
F_ℓ				$\circ \text{---} \circ \text{---} \circ \text{---} \circ$ $\alpha_1 \quad \alpha_2 \quad \alpha_3 \quad \alpha_4$
G_ℓ		$\circ \text{---} \circ$ $\alpha_1 \quad \alpha_2$		

Next we label the nodes (vertices, roots) and the weights, as in [Jacobson, §IV.5, pp. 134-135]. (We have numbered the nodes on the E_ℓ as in [Humphreys, §11.4, p. 58], rather than [Jacobson, §IV.5, pp. 134-135].)

Type Dynkin Diagram

Coxeter Diagram


 H_3

 H_4

 $I_2(n)$


Type	$\ell = 1$	2	3	4
A_ℓ	$\begin{array}{c} 1 \\ \circ \\ \alpha_1 \end{array}$	$\begin{array}{cc} 1 & 1 \\ \circ & \text{---} \circ \\ \alpha_1 & \alpha_2 \end{array}$	$\begin{array}{ccc} 1 & 1 & 1 \\ \circ & \text{---} \circ & \text{---} \circ \\ \alpha_1 & \alpha_2 & \alpha_3 \end{array}$	$\begin{array}{cccc} 1 & 1 & 1 & 1 \\ \circ & \text{---} \circ & \text{---} \circ & \text{---} \circ \\ \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \end{array}$
B_ℓ	$\begin{array}{c} \circ \\ \alpha_1 \end{array}$	$\begin{array}{cc} 2 & 1 \\ \circ & \text{---} \circ \\ \alpha_1 & \alpha_2 \end{array}$	$\begin{array}{ccc} 2 & 2 & 1 \\ \circ & \text{---} \circ & \text{---} \circ \\ \alpha_1 & \alpha_2 & \alpha_3 \end{array}$	$\begin{array}{cccc} 2 & 2 & 2 & 1 \\ \circ & \text{---} \circ & \text{---} \circ & \text{---} \circ \\ \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \end{array}$
C_ℓ	$\begin{array}{c} \circ \\ \alpha_1 \end{array}$	$\begin{array}{cc} 1 & 2 \\ \circ & \text{---} \circ \\ \alpha_1 & \alpha_2 \end{array}$	$\begin{array}{ccc} 1 & 1 & 2 \\ \circ & \text{---} \circ & \text{---} \circ \\ \alpha_1 & \alpha_2 & \alpha_3 \end{array}$	$\begin{array}{cccc} 1 & 1 & 1 & 2 \\ \circ & \text{---} \circ & \text{---} \circ & \text{---} \circ \\ \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \end{array}$
D_ℓ	$\begin{array}{c} 1 \\ \circ \\ \alpha_1 \end{array}$	$\begin{array}{cc} 1 & 1 \\ \circ & \circ \\ \alpha_1 & \alpha_2 \end{array}$	$\begin{array}{ccc} 1 & 1 & 1 \\ \circ & \text{---} \circ & \text{---} \circ \\ \alpha_1 & \alpha_2 & \alpha_3 \end{array}$	$\begin{array}{c} 1 \\ \circ \alpha_3 \\ \diagup \\ 1 \circ \alpha_2 \text{---} 1 \circ \alpha_1 \\ \diagdown \\ \circ \alpha_4 \end{array}$
$D_\ell \text{ alt}$	$\begin{array}{c} 1 \\ \circ \\ \alpha_1 \end{array}$	$\begin{array}{c} \circ \alpha_2 \\ \circ \alpha_1 \end{array}$	$\begin{array}{c} 1 \circ \alpha_1 \\ \diagup \\ 1 \circ \alpha_2 \\ \diagdown \\ 1 \circ \alpha_3 \end{array}$	$\begin{array}{c} 1 \circ \alpha_3 \\ \diagup \\ 1 \circ \alpha_2 \text{---} 1 \circ \alpha_1 \\ \diagdown \\ 1 \circ \alpha_4 \end{array}$
F_ℓ				$\begin{array}{cccc} 1 & 1 & 2 & 2 \\ \circ & \text{---} \circ & \text{---} \circ & \text{---} \circ \\ \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \end{array}$
G_ℓ		$\begin{array}{cc} 3 & 1 \\ \circ & \text{---} \circ \\ \alpha_1 & \alpha_2 \end{array}$		

14.2. The $\text{\LaTeX}$ code. We begin with the default versions of the Dynkin and Coxeter diagrams produced by the `dynkin-diagrams` package, followed by the $\text{\LaTeX}$ code that produces them.

After that, we show the code for the versions we prefer, with scaling and various parameters set.

Type	Dynkin Diagram	Coxeter Diagram
A_ℓ		
B_ℓ		
C_ℓ		
D_ℓ		
E_6		
E_7		
E_8		
F_4		
G_2		
H_2		
H_3		
H_4		
$I_2(n)$		

Here is the $\text{\LaTeX}$ code for the table above.

```

\begin{array}{c@{\quad}l@{\quad}l}
\textbf{Type}&\textbf{Dynkin Diagram}&\textbf{Coxeter Diagram}
\\
A_{\ell}&\dynkin{A}{\ell}&\dynkin[Coxeter]{A}{\ell}
\\
B_{\ell}&\dynkin{B}{\ell}&\dynkin[Coxeter]{B}{\ell}
\\
C_{\ell}&\dynkin{C}{\ell}&\dynkin[Coxeter]{C}{\ell}
\\
D_{\ell}&\dynkin{D}{\ell}&\dynkin[Coxeter]{D}{\ell}
\\
E_6&\dynkin{E}{6}&\dynkin[Coxeter]{E}{6}
\\
E_7&\dynkin{E}{7}&\dynkin[Coxeter]{E}{7}
\\
E_8&\dynkin{E}{8}&\dynkin[Coxeter]{E}{8}
\\
F_4&\dynkin{F}{4}&\dynkin[Coxeter]{F}{4}

```

```

\\
G_2&\dynkin{G}{2}&\dynkin[Coxeter,gonality={6}]{G}{2}
\\
H_2&&\dynkin[Coxeter]{H}{2}
\\
H_3&&\dynkin[Coxeter]{H}{3}
\\
H_4&&\dynkin[Coxeter]{H}{4}
\\
I_2(n)&&\dynkin[Coxeter,gonality={n}]{I}{n}
\end{array}
\]
```

Note that the first diagram below for D_2 is incorrect: `\dynkin{D}{2}` gives $\bullet \dashrightarrow$. The actual D_2 diagram is $\bullet \bullet$, coded as `\dynkins{D1|D1}`.

Type	$\ell = 1$	2	3	4
A_ℓ	$\bullet$	$\bullet \bullet$	$\bullet \bullet \bullet$	$\bullet \bullet \bullet \bullet$
B_ℓ	$\bullet$	$\bullet \dashrightarrow$	$\bullet \bullet \dashrightarrow$	$\bullet \bullet \bullet \dashrightarrow$
C_ℓ	$\bullet$	$\bullet \dashleftarrow$	$\bullet \bullet \dashleftarrow$	$\bullet \bullet \bullet \dashleftarrow$
D_ℓ	$\bullet$	$\bullet \bullet$	$\bullet \bullet \bullet$	$\bullet \bullet \bullet \bullet$
$D_\ell \text{ alt}$	$\bullet$	$\bullet$	$\bullet \bullet$	$\bullet \bullet \bullet$
F_ℓ				$\bullet \bullet \bullet \bullet$
G_ℓ		$\bullet \bullet \bullet$		

2 1

Here is the L^AT_EX code for the table above.

```

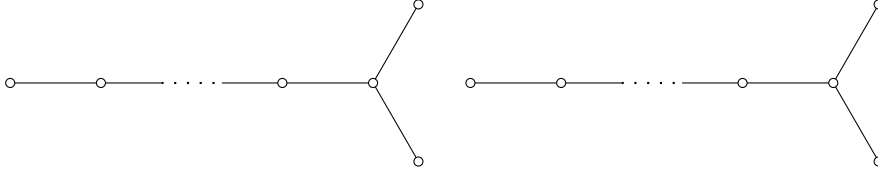
\makeatletter
\newcommand{\extraNodep}[6]%
{%
\dynkinPlaceRootRelativeTo{#1}{#2}{#3}{#4}{#5}
\dynkinDefineSingleEdge{#1}{#2}
\dynkinRootMark{*}{#1}
\advance\dynkin@nodes by 1
\dynkinLabelRoot{#1}{#6} }
\makeatother
```

```

\[
\begin{array}{ccl@{\quad}l@{\quad}l@{\quad}l\textbf{Type}&
| \quad \ell=&\multicolumn{1}{c}{\textbf{1}\quad}&
\multicolumn{1}{c}{\textbf{2}\quad}&
\multicolumn{1}{c}{\textbf{3}\quad}&\multicolumn{1}{c}{\textbf{4}}
\\ \hline
A_{\ell}&&\dynkin{A}{1}&\dynkin{A}{2}&\dynkin{A}{3}&\dynkin{A}{4}
\\
B_{\ell}&&\dynkin{B}{1}&\dynkin{B}{2}&\dynkin{B}{3}&\dynkin{B}{4}
\\
C_{\ell}&&\dynkin{C}{1}&\dynkin{C}{2}&\dynkin{C}{3}&\dynkin{C}{4}
\\
D_{\ell}&&\dynkin{D}{1}&\dynkin{D}{2}&\dynkin{D}{3}&\dynkin{D}{4}
\\
D_{\ell}\text{ alt}&&\dynkin{D}{1}&
\genfrac{}{}{0pt}{0}{\dynkin{D}{1}}{\dynkin{D}{1}}&
\begin{dynkinDiagram}D1
\extraNodep{2}{1}{northeast}{right}{left}{}
\extraNodep{2}{1}{southeast}{right}{left}{}
\end{dynkinDiagram}
&
\begin{dynkinDiagram}[%
]D2
\extraNodep{2}{1}{northeast}{right}{left}{}
\extraNodep{2}{1}{southeast}{right}{left}{}
\end{dynkinDiagram}
\\
F_{\ell}&&&&\dynkin{F}{4}
\\
G_{\ell}&&&\dynkin[labels={2,1}]{G}{2}
\end{array}
\]

```

Next we list examples of code for our scaled versions with various parameters set. First we list both the Dynkin and Coxeter diagrams for D_ℓ .

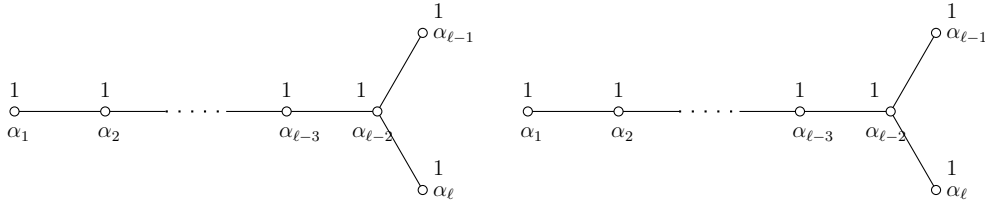


```

 $\backslash$ dynkin[root radius=.03cm,scale=2,mark=o,edge length=0.6cm,
indefinite edge ratio=2,indefinite edge/.style=
{draw=black,fill=white,thick,loosely dotted},
arrow style={black,length=1.5mm,width=3mm,line width=1pt}]{D}{\ell}
\quad
 $\backslash$ dynkin[root radius=.03cm,scale=2,mark=o,edge length=0.6cm,
indefinite edge ratio=2,indefinite edge/.style=
{draw=black,fill=white,thick,loosely dotted},
arrow style={black,length=1.5mm,width=3mm,line width=1pt},Coxeter]{D}{\ell}$

```

Here is the same example with labels and weights, as in [Jacobson, §IV.5, pp. 134-135].



```

 $\backslash$ dynkin[root radius=.03cm,scale=2,mark=o,edge length=0.6cm,
indefinite edge ratio=2,indefinite edge/.style=
{draw=black,fill=white,thick,loosely dotted},
arrow style={black,length=1.5mm,width=3mm,line width=1pt},
label macro/.code={\alpha_{\drlap{\#1}}},
labels={1,2,\ell-3,\ell-2,\ell-1,\ell},labels*={1,1,1,1,1,1}]{D}{\ell}
\quad
 $\backslash$ dynkin[root radius=.03cm,scale=2,mark=o,edge length=0.6cm,
indefinite edge ratio=2,indefinite edge/.style=
{draw=black,fill=white,thick,loosely dotted},
arrow style={black,length=1.5mm,width=3mm,line width=1pt},
label macro/.code={\alpha_{\drlap{\#1}}},
labels={1,2,\ell-3,\ell-2,\ell-1,\ell},labels*={1,1,1,1,1,1},Coxeter]{D}{\ell}$

```

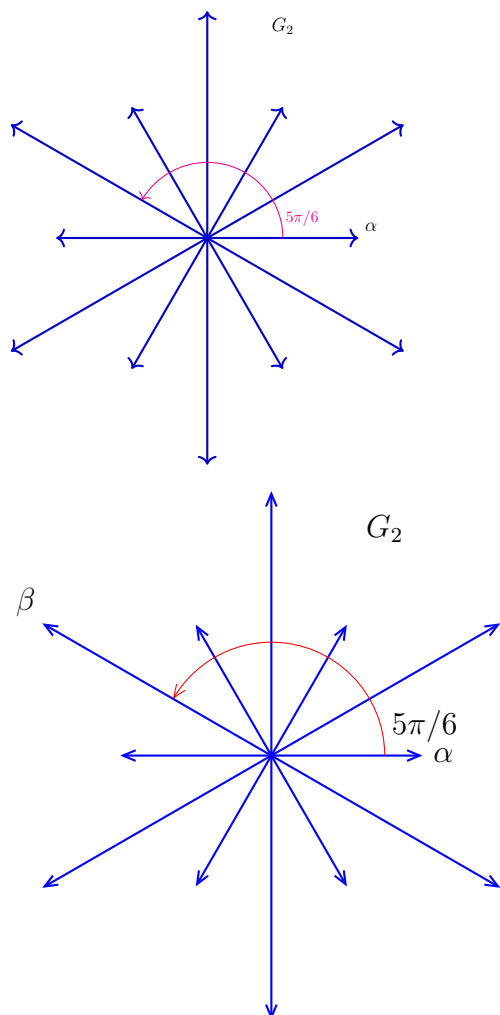
15. Root System Pictures

Here are some root system pictures for low dimensional semisimple Lie algebras, drawn in $\text{\LaTeX}$, with `tikz` and the package `rank-2-roots`. The root system A_2 corresponds to

the Lie algebra $\mathfrak{sl}(3)$, the root system B_2 corresponds to the Lie algebra $\mathfrak{so}(5)$, and the root system C_2 corresponds to the Lie algebra $\mathfrak{sp}(4)$. In fact, $\mathfrak{so}(5, \mathbb{F}) \cong \mathfrak{sp}(4, \mathbb{F})$ when $\mathbb{F} = \mathbb{C}$ but not when $\mathbb{F} = \mathbb{R}$. The root system G_2 is the root system of an exceptional Lie algebra of dimension 14; this Lie algebra is often denoted either G_2 or $\mathfrak{g}_2$.

This section is a work in progress.

First, two pictures of G_2 drawn with `tikz`.



$\LaTeX$ /`tikz` code for the second picture.

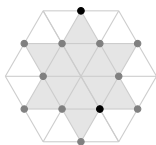
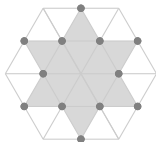
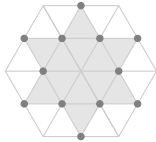
```
%\usetikzlibrary{arrows.meta}
\begin{tikzpicture}[
  % arrow heads for all lines (with narrower arrow head width)
  -{Straight Barb[bend,
    width=\the\dimexpr10\pgflinewidth\relax,
    length=\the\dimexpr12\pgflinewidth\relax]}],
]
```

```

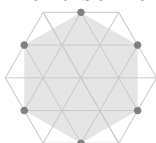
% straight arrows
\foreach \i in {0, 1, ..., 5} {
  \draw[thick, blue] (0, 0) -- (\i*60:2);
  \draw[thick, blue] (0, 0) -- (30 + \i*60:3.5);
}
% arc arrow
\draw[thin, red] (1.5, 0) arc[radius=1.5, start angle=0, end angle=5*30];
% annotations
\node[right] at (2, 0) {$\alpha$};
\node[above left, inner sep=.2em] at (5*30:3.5) {$\beta$};
\node[right] at (15:1.5) {$5\pi/6$};
\node at (1.5, 3) {$G_2$};
\end{tikzpicture}

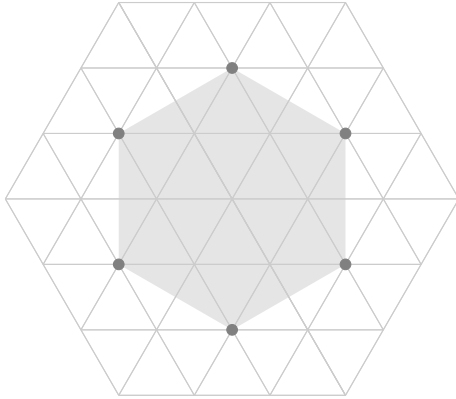
```

Next some versions of G_2 from the `rank-2-roots` L^AT_EX package.



Next some versions of A_2 from the `rank-2-roots` package.

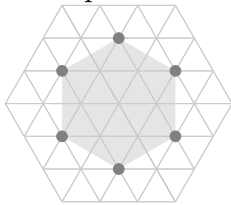




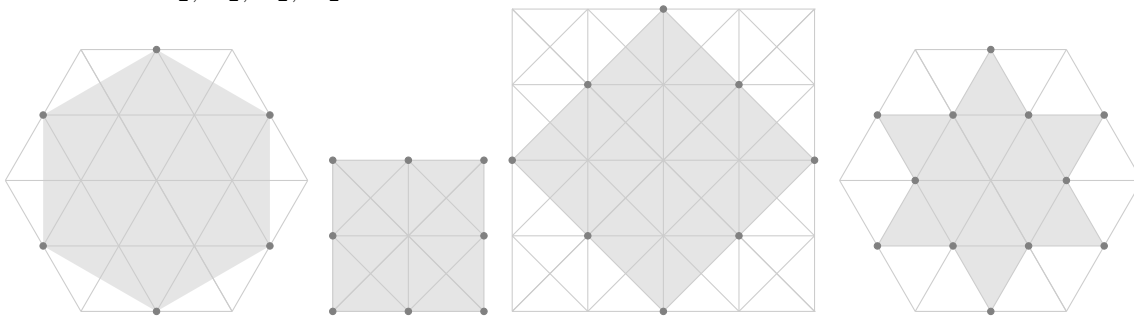
Code for the last picture:

```
%\usepackage{rank-2-roots}
\begin{tikzpicture}
%\AutoSizeWeightLatticefalse
%\begin{rootSystem}[weight length=1cm]{A}
\begin{rootSystem}{A}
\setlength{\weightLength}{1cm}
\setlength{\weightRadius}{2pt}
\weightLattice{3}
\roots
\end{rootSystem}
\end{tikzpicture}
```

Perhaps manual weight size?



Now all of A_2, B_2, C_2, G_2 with rank-2-roots.



Code:

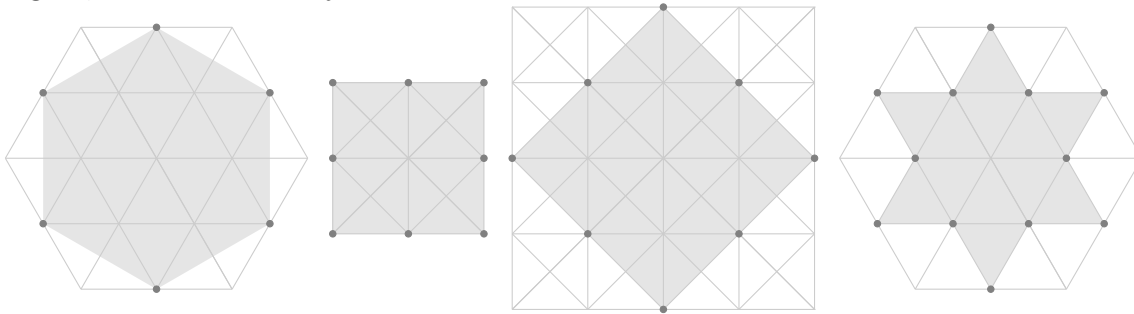
```
\foreach \i in {A,B,C,G}
```

```

{
  \begin{tikzpicture}%[baseline=-.5]
    \begin{rootSystem}{\i}
      \setlength{\weightLength}{1cm}
      \roots
    \end{rootSystem}
  \end{tikzpicture}
}

```

Again, with baseline adjustment.



Bibliography

- [Humphreys] James E. Humphreys, *Introduction to Lie Algebras and Representation Theory*, Graduate Texts in Mathematics, vol. 9, Springer, New York, 1972.
- [Jacobson] Nathan Jacobson, *Lie Algebras*, Dover, New York, 1962.
- [Bourbaki 1-3] Nicolas Bourbaki, *Lie Groups and Lie Algebras, Chapters 1-3*, Elements of Mathematics, Springer, Berlin, 1989.
- [Bourbaki 4-6] ———, *Lie Groups and Lie Algebras, Chapters 4-6*, Elements of Mathematics, Springer, Berlin, 2002.
- [Bourbaki 7-9] ———, *Lie Groups and Lie Algebras, Chapters 7-9*, Elements of Mathematics, Springer, Berlin, 2005.
- [Erdmann-Wildon] Karin Erdmann and Mark J. Wildon, *Introduction to Lie Algebras*, Undergraduate Mathematics Series, Springer, New York, 2006.

- [Fulton-Harris] William Fulton and Joe Harris, *Representation Theory: A First Course*, Graduate Texts in Mathematics, vol. 129, Springer, New York, 2004. *Readings in Mathematics*. MR48#2197
- [Grove-Benson] L. C. Grove and C. T. Benson, *Finite Reflection Groups*, 2nd ed., Graduate Texts in Mathematics, vol. 99, Springer, New York, 1985.
- [Humphreys:Coxeter] James E. Humphreys, *Reflection Groups and Coxeter Groups*, 2nd ed., Cambridge Studies in Advanced Mathematics, vol. 29, Cambridge University Press, Cambridge, 1990.
- [Hall] Brian C. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*, Graduate Texts in Mathematics, vol. 222, Springer, New York, 2003.
- [Knapp] Anthony W. Knap, *Lie Groups: Beyond an Introduction*, 2nd ed., Progress in Mathematics, vol. 140, Birkhäuser, Boston, 2002.
- [Radjavi-Rosenthal] Heydar Radjavi and Peter Rosenthal, *Simultaneous Triangularization*, Universitext, Springer, New York, 2000.
- [Saaty] Thomas L. Saaty (ed.), *Lectures on Modern Mathematics*, Vol. I, John Wiley & Sons, New York, 1963.
- [Serre] Jean-Pierre Serre, *Complex Semisimple Lie Algebras*, Monographs in Mathematics, Springer, Berlin, 1987.
- [Winter] David J. Winter, *Abstract Lie Algebras*, The MIT Press, Cambridge, MA, 1972.
- [Wikipedia] <https://en.wikipedia.org/>.

Index of Definitions

- α -string through β , 136
- Abelian, 46
- about, 142
- abstract Jordan-Chevalley decomposition, 119
- action, 84
- additive commutator, 18
- ADE, 160
- adjoint representation of a Lie algebra, 85
- almost algebraic, 120
- along, 142
- alternating form, 34
- alternating power, 94
- ascending central series, 53
- associative form, 70
- associativity of a form, 35
- base, 149, 165
- bilinear form, 34
- Borel subalgebra, 127
- bracket, 42
- bracket of subspaces, 48
- Cartan integers, 136, 167
- Cartan matrix, 161, 167
- Cartan subalgebra, 126
- Casimir element, 115
- Casimir invariant, 115
- Casimir operator, 116
- center of a Lie algebra, 52
- centralizer, 48
- centralizes, 48
- characteristic polynomial, 19
- characteristic subspace, 64
- characteristic subspace – automorphism, 64
- characteristic subspace – derivation, 64
- closure of a root system, 144
- colon ideal, 189
- commutator, 18
- commutator of subspaces, 48
- complete, 76
- completely invariant form, 70
- completely reducible module, 86
- completely reducible representation, 86
- completely solvable Lie algebra, 53
- component, 153
- composition factors, 89
- composition series, 89
- concrete trace form, 35, 71
- conductor, 189
- connected graph, 153
- contragredient module, 93
- coroot, 166
- Coxeter diagram, 152
- Coxeter graph, 152
- Coxeter graph – abstract, 151
- Coxeter group, 144
- crystallographic condition, 157
- crystallographic group, 157
- crystallographic restriction, 157
- CSA, 126
- decomposable Lie algebra, 120
- derivation, 49
- derived series, 54
- derived subalgebra, 53
- descending central series, 54
- diagonalizable, 19, 25

- direct summand, 86
- direct summands, 86
- dominant weight, 167
- dual Killing form, 135

- effective, 143
- eigenspace, 21, 109, 129
- eigenvalue, 27, 129
- eigenvalues, 109
- eigenvector, 27
- eigenvectors, 109
- elementary automorphism, 51, 127
- Engel subalgebra, 126
- equivariant map, 86
- exponential, 37
- extension of scalars, 37
- exterior power, 94

- faithful, 85
- Fitting null component, 126
- flag, 30
- flag – complete, 29
- flag – full, 29
- flag – partial, 30
- formal character, 170
- fundamental reflection, 149
- fundamental root, 149
- fundamental weights, 166

- general linear Lie algebra, 43
- generalized eigenspace, 21
- Gram matrix, 152

- half-sum of the positive roots, 171
- Heisenberg Lie algebra, 46
- highest weight, 166
- highest weight vector, 104, 112, 166
- hyperoctahedral group, 154
- hyperplane, 142

- ideal, 48
- ideal quotient, 189
- indecomposable module, 86
- indecomposable representation, 86

- inner automorphism, 51, 127
- inner derivation, 49, 60
- integral weight, 167
- integrality condition, 157
- intertwining map, 86
- invariance of a form, 35
- invariant form, 70
- involution of fields, 45
- irreducible Coxeter group, 153
- irreducible module, 86
- irreducible representation, 86
- irreducible root system, 153
- isomorphic, 47
- isomorphism of Lie algebras, 47

- Jordan decomposition, 23
- Jordan-Chevalley decomposition, 23
- Jordan-Chevalley decomposition — abstract, 119
- Jordan-Hölder series, 89

- kernel, 85
- Killing form, 71
- Killing-Cartan form, 71

- lattice, 157
- length of composition series, 89
- Levi factor, 96
- Levi subalgebra, 96
- Lie algebra, 42
- Lie algebra homomorphism, 47
- Lie bracket, 42
- Lie product, 42
- linear Lie algebra, 41
- lower central series, 54

- map of Lie algebras, 47
- maximal vector, 104, 112
- minimal polynomial, 19
- mirror, 142
- module, 84

- nil radical, 58
- nil set of operators, 26

- nilpotent, 20
- nilpotent Lie algebra, 53
- nilpotent radical, 59
- nilpotent set of operators, 26
- non-standard, 26
- nondegenerate form, 34
- nonnegatively spanning, 149
- nonsingular form, 35
- normalizer, 48
- normalizes, 48

- orthogonal complement, 35
- orthogonal Lie algebra, 43
- orthogonal reflection, 142

- parabolic subalgebra, 128
- perfect Lie algebra, 54
- permutable trace, 34
- positive root, 148
- potentially diagonalizable, 19
- primary component, 22

- quadratic Casimir element, 117

- radical, 58
- radical of a form, 34
- reducible root system, 153
- reductive Lie algebra, 60
- reflection about W , 142
- reflection determined by v , 142
- regular element, 126
- regular representation of a Lie algebra, 85
- regular semisimple, 126
- regular with respect to a root system, 148
- representation, 85
- representation of a Lie algebra, 85
- root, 130
- root lattice, 167
- root of a reflection, 142
- root space, 130
- root spaces, 110
- root string, 136
- root system — reduced, 144
- root system — undreduced, 144
- root system of a group — reduced, 143
- root system of a group — unreduced, 143
- root vector, 130
- root vectors, 110
- roots, 109

- Schläfli matrix, 152
- semidirect product, 95
- semidirect sum, 95
- semisimple Lie algebra, 59
- semisimple module, 86
- semisimple representation, 86
- sides of a hyperplane, 147
- sign of Weyl group element, 171
- simple Lie algebra, 59
- simple module, 86
- simple reflection, 149
- simple root, 149
- simple roots, 165
- simple system, 149
- simply laced, 160
- skew-symmetric form, 34
- skew-symmetric Lie algebra relative to $\langle, \rangle$, 43
- solvable Lie algebra, 52
- solvable radical, 58
- special automorphism, 51
- special linear Lie algebra, 43
- special orthogonal Lie algebra, 43
- special unitary Lie algebra, 45
- split, 125
- split linear Lie algebra, 41
- split matrix, 20
- split module, 86
- split operator, 20
- split polynomial, 20
- split set of operators, 26
- split solvable, 53
- split subalgebra, 125
- squarefree polynomial, 19
- standard cyclic module, 168
- strictly triangular Lie algebra, 43
- strictly triangularizable, 25
- strictly upper triangular Lie algebra, 43

strongly dominant weight, 167
subalgebra, 48
subgraph, 153
summand — direct, 86
summands — direct, 86
supersolvable group, 52
symmetric form, 34
symmetric power, 94
symplectic Lie algebra, 45

tensor power, 94
through, 142
toral subalgebra, 125
transporter, 49, 189
triangular Lie algebra, 43
triangularizable, 25
trivial representation, 85

unipotent, 25
unitary Lie algebra, 45
universal Casimir element, 116
universal enveloping algebra, 86
upper central series, 53
upper triangular Lie algebra, 43

Verma module, 168

weight, 106, 129, 166
weight lattice, 167
weight space, 22, 109, 129, 166
weight vector, 106
weight vectors, 109
weights, 109
Weyl group, 171

Index of Notation and Symbols

Items are ordered according to their appearance in the notes.

Starred symbols may denote different things in other parts of the notes.

Symbol	Name/Description	Page(s)
$\mathbf{GL}(n, \mathbb{R})$	general linear group of $n \times n$ invertible real matrices	7
$\mathbf{GL}(n, \mathbb{C})$	general linear group of $n \times n$ invertible complex matrices	7
$\mathbb{F}$	the base field (characteristic 0)	11, 12
$M_n(\mathbb{F})$	algebra of $n \times n$ matrices over $\mathbb{F}$	11
$\mathfrak{gl}(n, \mathbb{F})$	general linear Lie algebra of $n \times n$ matrices over $\mathbb{F}$	11
$\mathfrak{gl}(n)$	general linear Lie algebra of $n \times n$ matrices	11, 43
$\text{End}(V)$	algebra of linear operators on V	11
$\mathfrak{gl}(V)$	general linear Lie algebra of linear operators on V	11
ad	adjoint mapping defined by $\text{ad } x(y) = [x, y]$	11, 60
$\mathfrak{gl}(L)$	general linear Lie algebra of linear operators on L	11
$\overline{\mathbb{F}}$	algebraic closure of the base field $\mathbb{F}$	12, 19
$\mathbb{E}$	extension of the base field $\mathbb{F}$	12
$\mathfrak{gl}$	general linear Lie algebra	13
$\mathfrak{sl}$	special linear Lie algebra	13
$[\cdot, \cdot]$	Lie bracket or additive commutator	18
$\text{mp}(A)$	minimal polynomial of the operator or matrix A	19
$\text{cp}(A)$	characteristic polynomial of the operator or matrix A	19
$\mathbb{F}[x]$	polynomial ring over $\mathbb{F}$ in the indeterminate x	19
tr	trace of a matrix or operator	21
$A _W$	restriction of the operator A to the subspace W	21
V_λ	λ -eigenspace or λ -weight space in V	21
$V_{[\lambda]}$	generalized λ -eigenspace or generalized λ -weight space in V	21
$X _W$	set of restrictions of the operators A in X to the subspace W	26
char	characteristic of a field	34
$^\perp$	orthogonal complement	35
V^*	dual space of V , i.e., all linear maps from V to $\mathbb{F}$	35
τ	concrete trace form	35, 71

Symbol	Name/Description	Page(s)
e_{rs}	either the r, s matrix unit (1 in the r, s -entry and 0 elsewhere) or the usual basis of $\text{End } V$ relative to a given basis $v_1, \dots, v_k$ of V (taking v_t to $\delta_{rt}v_s$)	36
$\mathfrak{sl}(n)$	special linear Lie algebra of $n \times n$ matrices with trace 0	36, 43
$V_{\mathbb{E}}$	$\mathbb{E}$ -vector space $\mathbb{E} \otimes_{\mathbb{F}} V$ with scalars extended from $\mathbb{F}$ to $\mathbb{E}$	36
$A_{\mathbb{E}}$	$\mathbb{E}$ -linear operator $\text{id}_{\mathbb{E}} \otimes A$ extended from $A : V \rightarrow V$	36
e^A	exponential of the matrix or operator A ; same as $\exp A$	37
$\exp A$	exponential of the matrix or operator A	37
CRT	Chinese remainder theorem	38
$\text{GL}(n, \mathbb{F})$	general linear group of $n \times n$ invertible matrices over $\mathbb{F}$	42
Sym_n	symmetric group on n letters	42
A_{ℓ}	Dynkin diagram type; also used to refer to the corresponding Coxeter/Weyl group and simple Lie algebra $\mathfrak{sl}(\ell + 1)$	43
$\mathfrak{t}(n)$	Lie algebra of $n \times n$ upper triangular matrices	43
$\mathfrak{n}(n)$	Lie algebra of $n \times n$ strictly upper triangular matrices	43
$\mathfrak{o}(n)$	orthogonal Lie algebra of $n \times n$ skew-symmetric matrices (automatically trace 0)	43, 44
$\mathfrak{so}(n)$	orthogonal Lie algebra of $n \times n$ skew-symmetric matrices; equals $\mathfrak{o}(n)$	43
$\mathfrak{gl}(V, \langle, \rangle)$	general linear Lie algebra of linear operators on V that are skew-symmetric relative to the form $\langle, \rangle$	43
$\mathfrak{gl}(n, M)$	general linear Lie algebra of $n \times n$ matrices that transpose-skew-commute with the matrix M , i.e., $A^T M = -MA$	44
D_{ℓ}	Dynkin diagram type; also used to refer to the corresponding Coxeter/Weyl group and simple Lie algebra $\mathfrak{o}(2\ell)$	44
B_{ℓ}	Dynkin diagram type; also used to refer to the corresponding Weyl group and simple Lie algebra $\mathfrak{o}(2\ell + 1)$	44
C_{ℓ}	Dynkin diagram type; also used to refer to the corresponding Weyl group and simple Lie algebra $\mathfrak{sp}(2\ell)$	45
$\mathfrak{su}(2)$	special unitary Lie algebra of unitary 2×2 skew-Hermitian matrices of trace 0	45
e_{ij}	either the i, j matrix unit (1 in the i, j -entry and 0 elsewhere) or the usual basis of $\text{End } V$ relative to a given basis $v_1, \dots, v_k$ of V (taking v_k to $\delta_{ik}v_j$)	46, 62, 67
$L_{a,b,c}$	the a,b,c Lie algebra for scalars a, b, c , defined by $[x, y] = cz$, $[y, z] = ax$, $[z, x] = by$	47, 176
$I \triangleleft L$	I is an ideal of L	48

Symbol	Name/Description	Page(s)
$[S, T]$	subspace generated by the brackets $[s, t]$ with $s \in S, t \in T$	48
$C(S)$	The centralizer of S , sometimes written $C_L(S)$	48
$N(S)$	The normalizer of S , sometimes written $N_L(S)$	48
$(A : B)$	transporter of B into A ; also called the ideal quotient or colon ideal	49, 189, 191
$\text{Der } L$	Lie algebra of derivations on L	49
$\text{IDer } L$	Lie algebra of inner derivations on L	50
e^δ	exponential of the derivation δ ; a Lie algebra automorphism; same as $\exp \delta$	50
$\exp \delta$	exponential of the derivation δ ; same as e^δ	50
δ_s	semisimple part in the Jordan-Chevalley decomposition of the derivation δ	51
δ_n	nilpotent part in the Jordan-Chevalley decomposition of the derivation δ	51
$Z(L)$	center of the Lie algebra L	52
Z_i	i^{th} term of the ascending central series of a Lie algebra	53
$\mathcal{C}_i L$	i^{th} term of the ascending central series of L	53
L'	** derived subalgebra of L ; also denoted L^1	53
L^1	derived subalgebra of L ; also denoted $[L, L], L', L^{(1)}, \mathcal{D}L$	53
$L^{(1)}$	derived subalgebra of L ; also denoted L^1	53
$L^{(i)}$	** i^{th} term of the derived series of L	54
$\mathcal{D}^i L$	i^{th} term of the derived series of L ; same as $L^{(i)}$	54
$\mathcal{D}L$	derived subalgebra of L ; also denoted L^1	54
L^i	i^{th} term of the descending central series of L	54
$\mathcal{C}^i L$	i^{th} term of the descending central series of L ; same as L^i	54
Rad	the solvable radical (largest solvable ideal) of a Lie algebra	58
Nil	the nil radical (largest nilpotent ideal) of a Lie algebra	58
$R^{(k)}$	k^{th} term in the derived series of the Lie algebra R	59
ad_L	adjoint action the Lie algebra L	60
ad_S	adjoint action the subalgebra S	60
diag	diagonal matrix (with following elements the diagonal entries)	69
κ	the Killing form on a Lie algebra	71
$\kappa _S$	the Killing form on L restricted to the subalgebra S ; really $\kappa_L _{S \times S}$	72
κ_S	the Killing form on the subalgebra S	72
$\kappa _I$	the Killing form on L restricted to the ideal I ; really $\kappa_L _{I \times I}$	72
κ_I	the Killing form on the ideal I regarded as a Lie algebra	72
$\mathfrak{d}(n)$	Lie algebra of $n \times n$ diagonal matrices	75

Symbol	Name/Description	Page(s)
$U(L)$	universal enveloping algebra of the Lie algebra L	86
$T^k(V)$	k^{th} tensor $V^{\otimes k}$ power of V	94
$V^{\otimes k}$	tensor product $V \otimes V \otimes \cdots \otimes V$ of k copies of V	94
$S^k(V)$	k^{th} symmetric power of V	94
$\text{Sym}^k(V)$	k^{th} symmetric power of V	94
$\Lambda^k(V)$	k^{th} exterior power of V	94
$\text{Alt}^k(V)$	k^{th} exterior power of V	94
$V^{\otimes n}$	tensor product $V \otimes V \otimes \cdots \otimes V$ of n copies of V	107
$S^n(V)$	n^{th} symmetric power of V	107
$\text{Sym}^n(V)$	n^{th} symmetric power of V	107
c	** the (universal, quadratic) Casimir element	115
c_ϕ	the Casimir element for the representation ϕ	116
H	** a split toral subalgebra, usually a maximal toral subalgebra or Cartan subalgebra	126
V_α	α -eigenspace of α -weight space	129
L_α	α -root space of the Lie algebra L relative to a split toral subalgebra or Cartan subalgebra	129
Φ	** set of roots of a semisimple Lie algebra <i>or</i> a root system	130, 143
$C(H)$	The centralizer of H , sometimes written $C_L(H)$	130
$N(H)$	The normalizer of H , sometimes written $N_L(H)$	130
t_α	** element of H with $\alpha(h) = \kappa(t_\alpha, h)$ for all $h \in H$	135
h_α	** $2t_\alpha/\kappa(t_\alpha, t_\alpha)$; basis of $[L_\alpha, L_{-\alpha}]$	135
S_α	$L_\alpha \oplus [L_\alpha, L_{-\alpha}] \oplus L_{-\alpha}$; a copy of $\mathfrak{sl}(2)$	135
$\ v\ $	norm (length) of the vector v	142
$\mathbf{O}(V)$	orthogonal group of the inner product space V	142
(v, w)	** inner product of v and w	142
$\sigma_v(w)$	reflection of w about the hyperplane perpendicular to v	142
σ_v	reflection about the hyperplane perpendicular to v	142
H_-	negative side of the hyperplane H	147
H_+	positive side of the hyperplane H	147
Φ_t^+	set of t -positive roots	148
Φ_t^-	set of t -negative roots	148
Δ	** base of a root system	149
$m(v, w)$	the order of $\sigma_v \circ \sigma_w$	150
Γ	** Coxeter graph	152