

Introduction to Lie Algebras

J.I. Hall

11 November 2015

Contents

Preface	v
1 Introduction	1
1.1 Algebras	1
1.2 Types of algebras	2
1.3 Jordan algebras	6
1.4 Lie algebras and representation	7
1.5 Problems	12
2 Examples of Lie algebras	13
2.1 Abelian algebras	13
2.2 Generators and relations	13
2.3 Matrix algebras	14
2.4 Derivations	16
2.4.1 Derivations of polynomial algebras	17
2.4.2 Derivations of nonassociative algebras	19
2.4.3 Vector fields	19
2.5 Other constructions	20
2.5.1 Extensions	20
2.5.2 Embeddings	21
2.5.3 Nilpotent groups	21
2.6 Problems	23
3 Lie groups	25
3.1 Representation theory as spectral theory	26
3.2 Lie groups and Hilbert's Fifth Problem	26
3.3 Some matrix calculus	27
3.4 One-parameter subgroups	29
3.5 Equivalence of representation	35
3.6 Problems	38

4	Basics of Lie Algebras	39
4.1	Basic structure theory	39
4.2	Basic representation theory	42
4.3	Further structure and representation	42
4.4	Problems	46
5	Nilpotent representations	47
5.1	Engel's Theorem and Cartan subalgebras	47
5.2	Weight spaces and Cartan decomposition	50
5.3	Problems	54
6	Killing forms and semisimple Lie algebras	55
6.1	Killing forms	55
6.2	Semisimple algebras I	58
6.3	Problems	62
7	Representations of $\mathfrak{sl}_2(\mathbb{K})$	63
7.1	Weight modules	63
7.2	Verma modules	68
7.3	The Casimir operator	75
7.4	Problems	78

Preface

These are notes for the MSU Fall Semester 2015 course MTH914: *Lie Algebras*. The primary aim of the course is the introduction and discussion of the complex finite dimensional (semi-)simple Lie algebras and their representations.

Chapter 1

Introduction

1.1 Algebras

Class 1: 2 Sept 15

Let $\mathbb{K}$ be a field. A $\mathbb{K}$ -algebra $(\mathbb{K}A, \mu)$ is a (left) $\mathbb{K}$ -space A equipped with a bilinear multiplication. That is, there is a $\mathbb{K}$ -space homomorphism *multiplication* $\mu: A \otimes_{\mathbb{K}} A \rightarrow A$. We often write ab in place of $\mu(a \otimes b)$. Also we may write A or (A, μ) in place of $(\mathbb{K}A, \mu)$ when the remaining pieces should be evident from the context.

If A is a $\mathbb{K}$ -algebra, then its *opposite algebra* A^{op} has the same underlying vector space but its multiplication μ^{op} is given by $\mu^{\text{op}}(x \otimes y) = \mu(y \otimes x)$.

(1.1). LEMMA. *The map $\mu: A \otimes_{\mathbb{K}} A \rightarrow A$ is a $\mathbb{K}$ -algebra multiplication if and only if the adjoint map* lem-alg

$$\text{ad}: x \mapsto \text{ad}_x \quad \text{given by} \quad \text{ad}_x a = xa$$

is a $\mathbb{K}$ -vector space endomorphism of A into $\text{End}_{\mathbb{K}}(A)$. □

If $\mathcal{V} = \{v_i \mid i \in I\}$ is a $\mathbb{K}$ -basis of A , then the algebra is completely described by the associated *multiplication coefficients* or *structure constants* $c_{ij}^k \in \mathbb{K}$ given by

$$v_i v_j = \sum_{k \in I} c_{ij}^k v_k ,$$

for all i, j .

We may naturally *extend scalars* from $\mathbb{K}$ to any extension field $\mathbb{E}$. Indeed $\mathbb{E} \otimes_{\mathbb{K}} A$ has a natural $\mathbb{E}$ -algebra structure with the same multiplication coefficients for the basis $\mathcal{V}$.

Going the other direction is a little more subtle. If the $\mathbb{E}$ -algebra B has a basis $\mathcal{V}$ for which all the c_{ij}^k belong to $\mathbb{K}$, then the $\mathbb{K}$ -span of the basis is a $\mathbb{K}$ -algebra A for which $B = \mathbb{E} \otimes_{\mathbb{K}} A$. In that case we say that A is a $\mathbb{K}$ -form of the

algebra A . In many cases the $\mathbb{E}$ -algebra B has several pairwise nonisomorphic $\mathbb{K}$ -forms.

Various generalizations of the above are available and often helpful. The extension field $\mathbb{E}$ of $\mathbb{K}$ is a itself special sort of $\mathbb{K}$ -algebra. If C is an arbitrary $\mathbb{K}$ -algebra, then $C \otimes_{\mathbb{K}} A$ is a $\mathbb{K}$ -algebra, with opposite algebra $A \otimes_{\mathbb{K}} C$. The relevant multiplication is $\mu = \mu_C \otimes \mu_A$:

$$\mu((c_1 \otimes a_1) \otimes (c_2 \otimes a_2)) = \mu_C(c_1 \otimes c_2) \otimes \mu_A(a_1 \otimes a_2).$$

We might also wish to consider R -algebras for other rings R with identity. For the tensor product to work reasonably, R should be commutative. A middle ground would require R to be an integral domain, although even in that case we must decide whether or not we wish algebras to be free as R -module.

Of primary interest to us is the case $R = \mathbb{Z}$. A $\mathbb{Z}$ -algebra is a free abelian group (that is, *lattice*) $L = \bigoplus_{i \in I} \mathbb{Z}v_i$ provided with a bilinear multiplication $\mu_{\mathbb{Z}}$ which is therefore completely determined by the integral multiplication coefficients c_{ij}^k . From this we can construct $\mathbb{K}$ -algebras $L_{\mathbb{K}} = \mathbb{K} \otimes_{\mathbb{Z}} L$ for any field $\mathbb{K}$, indeed for any $\mathbb{K}$ -algebra. For instance $C \otimes_{\mathbb{Z}} \text{Mat}_n(\mathbb{Z})$ is the $\mathbb{K}$ -algebra $\text{Mat}_n(C)$ of all $n \times n$ matrices with entries from the $\mathbb{K}$ -algebra C .

Suppose for the basis $\mathcal{V}$ of the $\mathbb{K}$ -algebra A all the c_{ij}^k are integers—that is, belong to the subring of $\mathbb{K}$ generated by 1. Then the $\mathbb{Z}$ -algebra $L = \bigoplus_{i \in I} \mathbb{Z}v_i$ with these multiplication coefficients can be viewed as a $\mathbb{Z}$ -form of A (although we only have its quotient by $\text{char}(\mathbb{K})$ as a subring of A). The original $\mathbb{K}$ -algebra A is then isomorphic to $L_{\mathbb{K}}$.

1.2 Types of algebras

As $\dim_{\mathbb{K}}(A \otimes_{\mathbb{K}} A) \geq \dim_{\mathbb{K}}(A)$, every $\mathbb{K}$ -space admits $\mathbb{K}$ -algebras. We focus on those with some sort of interesting additional structure. Examples are associative algebras, Jordan algebras, alternative algebras, composition algebras, Hopf algebras, and Lie algebras—these last being the primary focus of our study. (All the others will be discussed at least briefly.)

In most cases these algebra types naturally form subcategories of the additive category ${}_{\mathbb{K}}\mathbf{Alg}$ of $\mathbb{K}$ -algebras, the maps φ of $\text{Hom}_{{}_{\mathbb{K}}\mathbf{Alg}}(A, B)$ being those linear transformations $\varphi \in \text{Hom}_{\mathbb{K}}(A, B)$ with $\varphi(xy) = \varphi(x)\varphi(y)$ for all $x, y \in A$. As the category ${}_{\mathbb{K}}\mathbf{Alg}$ is additive, each morphism has a kernel and image, which are defined as usual and enjoy the usual properties.

A subcategory will often be defined initially as belonging to a particular *variety* of $\mathbb{K}$ -algebras. For instance, the *associative* $\mathbb{K}$ -algebras are precisely those $\mathbb{K}$ -algebras satisfying the identical relation

$$(xy)z = x(yz).$$

Alternatively, the associative $\mathbb{K}$ -algebras are those whose multiplication map μ satisfies

$$\mu(\mu(x \otimes y) \otimes z) = \mu(x \otimes \mu(y \otimes z)).$$

As the defining identical relation is equivalent to its reverse $(zy)x = z(yx)$, the opposite of an associative algebra is also associative.

Similarly, the subcategory of *alternative $\mathbb{K}$ -algebras* is the variety of $\mathbb{K}$ -algebras given by the weak associative laws

$$x(xy) = (xx)y \quad \text{and} \quad x(yy) = (xy)y.$$

The opposite of an alternative algebra is also alternative.

Varietal algebras like these have nice local properties:

- (i) A $\mathbb{K}$ -algebra is associative if and only if all its 3-generator subalgebras are associative.
- (ii) A $\mathbb{K}$ -algebra is alternative if and only if all its 2-generator subalgebras are alternative.

The associative identity is linear in that each variable appears at most once in each term, while the alternative identity is not, since a appears twice in each term. The linearity of an identity implies that it only need be checked on a basis of the algebra to ensure that it is valid throughout the algebra. That is, there are appropriate identities among the various c_{ij}^k that are equivalent to the algebra being associative. (Exercise: find them.) This implies the (admittedly unsurprising) fact that extending the scalars of an associative algebra produces an associative algebra. It is also true that extending the scalars of an alternative algebra produces another alternative algebra, but that needs some discussion since the basic identity is not linear. (Exercise.)

The basic model for an associative algebra is $\text{End}_{\mathbb{K}}(V)$ for some $\mathbb{K}$ -space V . Indeed, most associative algebras (including all with an identity) are isomorphic to subalgebras of various $\text{End}_{\mathbb{K}}(V)$. (See Proposition (1.3).) For finite dimensional V we often think in matrix terms by choosing a basis for V and then using that basis to define an isomorphism of $\text{End}_{\mathbb{K}}(V)$ with $\text{Mat}_n(\mathbb{K})$ for $n = \dim_{\mathbb{K}}(V)$.

Of course, every associative algebra is alternative, but we now construct the most famous models for alternative but nonassociative algebras. If we start with $\mathbb{K} = \mathbb{R}$, then we have the familiar construction of the complex numbers as 2×2 matrices: for $a, b \in \mathbb{K}$ we set

$$(a, b) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

with multiplication given by

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} c & d \\ -d & c \end{pmatrix} = \begin{pmatrix} ac - bd & ad + bc \\ -bc - ad & -bd + ac \end{pmatrix}$$

and conjugation given by

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}^{-} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

As $\mathbb{R}$ is commutative and conjugation is trivial on $\mathbb{R}$, these can be rewritten:

For $a, b \in \mathbb{K}$ and $a \mapsto \bar{a}$ an antiautomorphism of $\mathbb{K}$, we set

$$(a, b)_{\mathbb{K}} = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix},$$

with

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \begin{pmatrix} c & d \\ -\bar{d} & \bar{c} \end{pmatrix} = \begin{pmatrix} ac - \bar{d}b & da + b\bar{c} \\ -c\bar{b} - \bar{a}d & -\bar{b}d + \bar{c}a \end{pmatrix}$$

and

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}^{-1} = \begin{pmatrix} \bar{a} & -b \\ \bar{b} & a \end{pmatrix}.$$

This then gives us the complex numbers $\mathbb{C}$ as the collection of all pairs $(a, b)_{\mathbb{R}}$ of real numbers. Feeding the complex numbers back into the machine produces Hamilton's *quaternions* $\mathbb{H}$ as all pairs $(a, b)_{\mathbb{C}}$ with the multiplication and the conjugation antiautomorphism described. As $\mathbb{C}$ is commutative the quaternions are associative, but they are no longer commutative.

Finally with $\mathbb{K} = \mathbb{H}$, the resulting $\mathbb{O}$ of all pairs $(a, b)_{\mathbb{H}}$ is the *octonions* of Cayley and Graves. The octonions are indeed alternative but not associative, although this requires checking. Again conjugation is an antiautomorphism.

In each case, the 2×2 "scalar matrices" are only those with $b = 0$ and $a = \bar{a} \in \mathbb{R}$, so we have constructed $\mathbb{R}$ -algebras with respective dimensions $\dim_{\mathbb{R}}(\mathbb{C}) = 2$, $\dim_{\mathbb{R}}(\mathbb{H}) = 4$, $\dim_{\mathbb{R}}(\mathbb{O}) = 8$.

Class 2: 4 Sept 15

A *quadratic form* on the $\mathbb{K}$ -space V is a map $q: V \rightarrow \mathbb{K}$ for which the associated map $b: V \times V \rightarrow \mathbb{K}$ given by polarization

$$b(x, y) = q(x + y) - q(x) - q(y)$$

is a nondegenerate bilinear form. (See Section 2.3 for a brief discussion of bilinear forms.)

The $\mathbb{R}$ -algebras $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, and $\mathbb{O}$ furnish examples of composition $\mathbb{R}$ -algebras. A *composition algebra* is a $\mathbb{K}$ -algebra A with multiplicative identity, admitting a nondegenerate quadratic form $\delta: A \rightarrow \mathbb{K}$ that is multiplicative:

$$\delta(x)\delta(y) = \delta(xy),$$

for all $x, y \in A$. The codimension 1 subspace $1^{\perp}$ consists of the *pure imaginary* elements of A , and (in characteristic not 2) the conjugation map $a1 + \bar{b} = a1 - b$, for $b \in 1^{\perp}$, is an antiautomorphism of A whose fixed point subspace is $\mathbb{K}1$.

In the above $\mathbb{R}$ -algebras the form δ is given by $\delta(x)1 = x\bar{x}$:

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \begin{pmatrix} \bar{a} & -b \\ \bar{b} & a \end{pmatrix} = a\bar{a} + b\bar{b} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

In $\mathbb{O}$ specifically, for $a, b, c, d, e, f, g, h \in \mathbb{R}$, we find

$$\begin{aligned} \delta(((a, b)_{\mathbb{R}}, (c, d)_{\mathbb{R}})_{\mathbb{C}}, ((e, f)_{\mathbb{R}}, (g, h)_{\mathbb{R}})_{\mathbb{C}})_{\mathbb{H}} &= \delta(a, b, c, d, e, f, g, h) = \\ &= a^2 + b^2 + c^2 + d^2 + e^2 + f^2 + g^2 + h^2. \end{aligned}$$

Thus in $\mathbb{O}$ (and so its subalgebras $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{H}$) all nonzero vectors have nonzero norm.

An immediate consequence of the composition law is that an invertible element of A must have nonzero norm. As $\delta(x)1 = x\bar{x}$ in composition algebras, the converse is also true. Therefore if 0 is the only element of the composition algebra A with norm 0, then all nonzero elements are invertible and A is a *division algebra*. Prime examples are the division composition $\mathbb{R}$ -algebras $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, and $\mathbb{O}$. The following remarkable theorem of Hurwitz shows that this situation is typical

(1.2). THEOREM. (HURWITZ' THEOREM) *If A is a composition algebra over $\mathbb{K}$, then $\dim_{\mathbb{K}}(A)$ is 1, 2, 4, or 8.* thm-hurwitz $\square\square$

If the composition $\mathbb{K}$ -algebra A is not a division algebra, then it is called *split*. It turns out that a split composition algebra over $\mathbb{K}$ is uniquely determined up to isomorphism by its dimension. In dimension 1, the algebra is $\mathbb{K}$ itself, always a division algebra. In dimension 4, a split composition $\mathbb{K}$ -algebra is always $\text{Mat}_2(\mathbb{K})$ with $\delta = \det$, and the diagonal matrices provide a split subalgebra of dimension 2.

Composition algebras of dimension 8 are called *octonion algebras*. The original is the real division algebra $\mathbb{O}$ presented above and due to Graves (1843, unpublished) and Cayley (1845) [SpV00, p. 23].

A split octonion algebra $\mathbb{O}^{\text{sp}}(\mathbb{K})$ over any field $\mathbb{K}$ is provided by *Zorn's vector matrices* [Zor31]

$$m = \begin{pmatrix} a & \vec{b} \\ \vec{c} & d \end{pmatrix}$$

with $a, d \in \mathbb{K}$ and $\vec{b}, \vec{c} \in \mathbb{K}^3$. Multiplication is given by

$$\begin{pmatrix} a & \vec{b} \\ \vec{c} & d \end{pmatrix} \begin{pmatrix} x & \vec{y} \\ \vec{z} & w \end{pmatrix} = \begin{pmatrix} ax + \vec{b} \cdot \vec{z} & a\vec{y} + w\vec{b} \\ x\vec{c} + d\vec{z} & \vec{c} \cdot \vec{y} + dw \end{pmatrix} + \begin{pmatrix} 0 & \vec{c} \times \vec{z} \\ -\vec{b} \times \vec{y} & 0 \end{pmatrix}$$

using the standard dot (inner) and cross (outer, exterior, vector) products of 3-vectors. The associated norm is

$$\delta(m) = ad - \vec{b} \cdot \vec{c}.$$

For any $\vec{v}$ with $\vec{v} \cdot \vec{v} = k \neq 0$ the subalgebra of all

$$m = \begin{pmatrix} a & b\vec{v} \\ ck^{-1}\vec{v} & d \end{pmatrix}$$

is a copy of the split quaternion algebra $\text{Mat}_2(F)$ with norm the usual determinant.

Zorn (and others) gave a slightly different version of the vector matrices, replacing our entry $\vec{c}$ with its negative. This gives the more symmetrical norm form $\delta(m) = ad + \vec{b} \cdot \vec{c}$ but makes the connection with standard matrix multiplication and determinants less clear.

Extending coefficients in a composition algebra produces a composition algebra (although this is more than an exercise). For every composition $\mathbb{K}$ -algebra

O , there is an extension $\mathbb{E}$ of degree at most 2 over $\mathbb{K}$ with $\mathbb{E} \otimes_{\mathbb{K}} O$ a split composition $\mathbb{E}$ -algebra. In particular every composition algebra over algebraically closed $\mathbb{E}$ is split and so unique up to isomorphism. The split algebra over $\mathbb{C}$ (for instance given by Zorn's vector matrices) has two isomorphism classes of $\mathbb{R}$ -forms—the class of the split algebra $\mathbb{O}^{\text{sp}}(\mathbb{R})$ and that of the Cayley-Graves division algebra $\mathbb{O}$.

1.3 Jordan algebras

As mentioned above, the basic models for associative algebras are the endomorphism algebras $\text{End}_{\mathbb{K}}(V)$ for some $\mathbb{K}$ -space V and the related matrix algebras $\text{Mat}_n(\mathbb{K})$. While Jordan and Lie algebras both have abstract varietal definitions (given below for Jordan algebras and in the next section for Lie algebras), they are first seen in canonical models coming from $\text{End}_{\mathbb{K}}(V)$.

We start with the observation that any pure tensor from $V \otimes V$ is the sum of its symmetric and skew-symmetric parts:

$$v \otimes w = \frac{1}{2}(v \otimes w + w \otimes v) + \frac{1}{2}(v \otimes w - w \otimes v).$$

In 1933 P. Jordan [JvNW34] initiated the study of the $\mathbb{K}$ -algebra $A^+ = (A, \mu^+) = (A, \circ)$ that is the associative $\mathbb{K}$ -algebra A equipped with the *Jordan product*

$$\mu^+(x \otimes y) = x \circ y = \frac{1}{2}(xy + yx).$$

This requires, of course, that the characteristic of the field $\mathbb{K}$ not be 2. We could also consider the algebra without the factor of $\frac{1}{2}$, but we keep it for various reasons—in particular $x \circ x = \frac{1}{2}(xx + xx) = xx = x^2$ and $1 \circ x = \frac{1}{2}(1x + x1) = x$.

The model for all Jordan algebras is then $\text{End}_{\mathbb{K}}^+(V)$, the vector space of all $\mathbb{K}$ -endomorphisms of V equipped with the Jordan product.

Clearly the algebra $\text{End}_{\mathbb{K}}^+(V)$ is commutative. Not so obvious is the fact that we also have the identity

$$(x \circ x) \circ (y \circ x) = ((x \circ x) \circ y) \circ x,$$

for all $x, y \in \text{End}_{\mathbb{K}}^+(V)$. (Exercise.)

We are led to the general, varietal definition: the $\mathbb{K}$ -algebra A is a *Jordan algebra* if it is commutative and satisfies the identical relation

$$x^2(yx) = (x^2y)x.$$

The canonical models are $\text{End}_{\mathbb{K}}^+(V)$ and so also $\text{Mat}_n^+(\mathbb{K})$ (in finite dimension).

Any subspace of $\text{End}_{\mathbb{K}}^+(V)$ that is closed under the Jordan product is certainly a Jordan subalgebra. Especially if τ is an automorphism of $\text{End}_{\mathbb{K}}(V)$, then its fixed-point-space is certainly closed under the Jordan product and so is a subalgebra. More subtly, if τ is an antiautomorphism of $\text{End}_{\mathbb{K}}(V)$, then it

induces an automorphism of $\text{End}_{\mathbb{K}}^+(V)$ whose fixed points are again a Jordan subalgebra.

For instance, in the $\mathbb{K}$ -algebra $\text{Mat}_n(\mathbb{K})$ the transpose map is an antiautomorphism, so the symmetric matrices from $\text{Mat}_n(\mathbb{K})$ form a Jordan subalgebra of $\text{Mat}_n^+(\mathbb{K})$. More generally, if A is a $\mathbb{K}$ -algebra with an antiautomorphism $a \mapsto \bar{a}$ fixing $\mathbb{K}$, then we can try the same trick with the $\mathbb{K}$ -algebra $\text{Mat}_n(A)$. The transpose-conjugate map

$$\bar{\tau}: (a_{ij}) \mapsto (\bar{a}_{ji})$$

is then an antiautomorphism of $\text{Mat}_n(A)$ (Exercise.), and so the associated fixed space of Hermitian matrices

$$\text{H}_n(A) = \{ M \in \text{Mat}_n(A) \mid M = M^{\bar{\tau}} \}$$

is closed under the Jordan product

$$M \circ N = \frac{1}{2}(MN + NM).$$

If A is associative then we have a Jordan algebra. Indeed this with $A = \mathbb{C}$ and $\mathbb{K} = \mathbb{R}$ was the original motivation for the physicist Jordan: in quantum mechanics the observables for the Hilbert space $\mathbb{C}^n$ are characterized by the hermitian matrices $\text{H}_n(\mathbb{C})$, a set which is not closed under the standard matrix product but is a real Jordan algebra under the Jordan product.

Class 3: 9 Sept '15

When A is not associative, there is no reason to assume that this gives $\text{H}_n(A)$ the structure of an (abstract) Jordan algebra. But if we choose A to be an octonion algebra over $\mathbb{K}$ and let $n \leq 3$, then this is in fact the case. (Recall that the alternative law is a weak version of the associative law, so this is not completely unreasonable.)

For the octonion $\mathbb{K}$ -algebra O , the Jordan algebra $\text{H}_3(O)$ is called an *Albert algebra*. Each matrix of $\text{H}_3(O)$ has the shape

$$\begin{pmatrix} a & \alpha & \beta \\ \bar{\alpha} & b & \gamma \\ \bar{\beta} & \bar{\gamma} & c \end{pmatrix}$$

with $a, b, c \in \mathbb{K}$ (the fixed field of conjugation in O) and $\alpha, \beta, \gamma \in O$. Thus the $\mathbb{K}$ -dimension of the Albert algebra $\text{H}_3(O)$ is $3 + 3 \times 8 = 27$.

1.4 Lie algebras and representation

In the previous section we only discussed the symmetric part of the tensor decomposition displayed at the beginning of the section. But even at the time of Jordan, the corresponding skew part had been studied for years, starting with the Norwegian Sophus Lie [Lie??]. If A is an associative algebra, then we define a skew algebra $A^- = (A, \mu^-) = (A, [\cdot, \cdot])$ by furnishing A with the multiplication

$$\mu^-(x \otimes y) = [x, y] = xy - yx.$$

(Note that the scaling factor $\frac{1}{2}$ does not appear.) The algebras A^- and in particular $\text{End}_{\mathbb{K}}^-(V)$ and $\text{Mat}_n^-(\mathbb{K})$ are the canonical models for Lie algebras over $\mathbb{K}$.

In a given category, a representation of an object M is loosely a morphism of M into one of the canonical examples from the category. So a linear representation of a group M is a homomorphism from M to some $\text{GL}_{\mathbb{K}}(V)$. With this in mind, we will say that a *linear representation* of an associative algebra A , a Jordan algebra J , and a Lie algebra L (all over $\mathbb{K}$), respectively, is a $\mathbb{K}$ -algebra homomorphism φ belonging to, respectively, some

$$\text{Hom}_{\mathbb{K}\text{Alg}}(A, \text{End}_{\mathbb{K}}(V)), \quad \text{Hom}_{\mathbb{K}\text{Alg}}(J, \text{End}_{\mathbb{K}}^+(V)), \quad \text{Hom}_{\mathbb{K}\text{Alg}}(L, \text{End}_{\mathbb{K}}^-(V)),$$

which in the finite dimensional case can be viewed as

$$\text{Hom}_{\mathbb{K}\text{Alg}}(A, \text{Mat}_n(\mathbb{K})), \quad \text{Hom}_{\mathbb{K}\text{Alg}}(J, \text{Mat}_n^+(\mathbb{K})), \quad \text{Hom}_{\mathbb{K}\text{Alg}}(L, \text{Mat}_n^-(\mathbb{K})).$$

The corresponding image of φ is then a *linear associative algebra*, *linear Jordan algebra*, or *linear Lie algebra*, respectively. The representation is *faithful* if its kernel is 0. The underlying space V or $\mathbb{K}^n$ is then an A -module which *carries* the representation and upon which the algebra *acts*.

It turns out that in each of these categories, many of the important examples are linear. For instance

prop-assoc-linear **(1.3). PROPOSITION.** *Every associative algebra with a multiplicative identity element is isomorphic to a linear associative algebra.*

PROOF. Let A be an associative algebra. For each $x \in A$, consider the map $\text{ad}: A \rightarrow \text{End}_{\mathbb{K}}(A)$ of Lemma (1.1), given by $x \mapsto \text{ad}_x$ where $\text{ad}_x a = xa$ as before. That lemma states that ad is a vector space endomorphism.

Thus we need to check that multiplication is respected. But the associative identity

$$(xy)a = x(ya)$$

can be restated as

$$\text{ad}_{xy} a = \text{ad}_x \text{ad}_y a,$$

for all $x, y, a \in A$. Hence $\text{ad}_{xy} = \text{ad}_x \text{ad}_y$ as desired.

The kernel of ad consists of those x with $xa = 0$ for all $a \in A$. In particular, the kernel is trivial if A contains an identity element. $\square$

It is clear from the proof that the multiplicative identity plays only a small role—the result should and does hold in greater generality. But for us the main message is that the adjoint map is a representation of every associative algebra. The proposition should be compared with Cayley's Theorem which proves that every group is (isomorphic to) a faithful permutation group by looking at the regular representation, which is the permutation version of adjoint action.

What about Jordan and Lie representation? Of course we still have not defined general Lie algebras, but we certainly want to include all the subalgebras of $\text{End}_{\mathbb{K}}^-(V)$ and $\text{Mat}_n^-(\mathbb{K})$.

As above, the multiplication map μ of an arbitrary Lie algebra $A = (A, [\cdot, \cdot])$ will be written as a bracket, in deference to the commutator product in an associative algebra:

$$\mu(x \otimes y) = [x, y].$$

In the linear Lie algebras $\text{End}_{\mathbb{K}}^-(V)$ and $\text{Mat}_n^-(\mathbb{K})$ we always have

$$[x, x] = xx - xx = 0,$$

so we require that an abstract Lie algebra satisfy the *null identical relation*

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

This identity is not linear, but we may “linearize” it by setting $x = y + z$. We then find

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

giving as an immediate consequence the linear *skew identical relation*

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

If $\text{char}\mathbb{K} \neq 2$, these two identities are equivalent. (This is typical of linearized identities: they are equivalent to the original except where neutralized by the characteristic.)

Our experience with groups and associative algebras tells us that having adjoint representations available is of great benefit, so we make an initial hopeful definition:

A Lie algebra is an algebra $(\mathbb{K}L, [\cdot, \cdot])$ in which all squares $[x, x]$ are 0 and for which the $\mathbb{K}$ -endomorphism $\text{ad}: L \longrightarrow \text{End}_{\mathbb{K}}^-(L)$ is a representation of L .

Are $\text{End}_{\mathbb{K}}^-(V)$ and $\text{Mat}_n^-(\mathbb{K})$ Lie algebras in this sense? Indeed they are:

$$\begin{aligned} \text{ad}_x \text{ad}_y a &= \text{ad}_x(ya - ay) \\ &= x(ya - ay) - (ya - ay)x \\ &= xya - xay - yax + ayx \end{aligned}$$

hence

$$\begin{aligned} [\text{ad}_x, \text{ad}_y]a &= (\text{ad}_x \text{ad}_y - \text{ad}_y \text{ad}_x)a \\ &= (xya - xay - yax + ayx) - (yxa - yax - xay + axy) \\ &= (xya - axy) - (yxa - ayx) \\ &= [xy, a] - [yx, a] \\ &= [xy - yx, a] \\ &= \text{ad}_{[x, y]} a. \end{aligned}$$

That is, $[\text{ad}_x, \text{ad}_y] = \text{ad}_{[x,y]}$, as desired.

Let us now unravel the consequences of the identity $\text{ad}_{[x,y]} = [\text{ad}_x, \text{ad}_y]$ for the algebra $(L, [\cdot, \cdot])$:

$$\begin{aligned} \text{ad}_{[x,y]} z &= [\text{ad}_x, \text{ad}_y] z \\ [[x, y], z] &= (\text{ad}_x \text{ad}_y - \text{ad}_y \text{ad}_x) z \\ [[x, y], z] &= (\text{ad}_x \text{ad}_y) z - (\text{ad}_y \text{ad}_x) z \\ [[x, y], z] &= [x, [y, z]] - [y, [x, z]] \\ [[x, y], z] &= -[[y, z], x] - [[z, x], y]. \end{aligned}$$

That is,

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

We arrive at the standard definition of a Lie algebra:

A Lie algebra is an algebra $(\mathbb{K}L, [\cdot, \cdot])$ that satisfies the two identical relations:

- (i) $[x, x] = 0$;
- (ii) (Jacobi Identity) $[[x, y], z] + [[y, z], x] + [[z, x], y] = 0$.

Negating the Jacobi Identity gives us the equivalent identity

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

In particular, the opposite of a Lie algebra is again a Lie algebra. (Exercise: the map $x \mapsto -x$ is an isomorphism of the Lie algebra L with its opposite algebra.)

The Jacobi Identity and the skew law $[y, z] = -[z, y]$ are both linear, and these serve to define Lie algebras if the characteristic is not 2. This is good enough to prove that tensor product extensions of Lie algebras are still Lie algebras as long as the characteristic is not 2

In all characteristics the null law $[x, x] = 0$ admits a weaker form of linearity. Assume that we already know $[y, y] = 0$, $[z, z] = 0$, and $[y, z] = -[z, y]$. Then for all constants a, b we have

$$\begin{aligned} [ay + bz, ay + bz] &= [ay, ay] + [ay, bz] + [bz, ay] + [bz, bz] \\ &= a^2[y, y] + ab([y, z] + [z, y]) + b^2[z, z] \\ &= 0 + 0 + 0 = 0. \end{aligned}$$

This, together with the linearity of the Jacobi Identity, gives

(1.4). PROPOSITION. *Let L be Lie $\mathbb{K}$ -algebra and $\mathbb{E}$ an extension field over $\mathbb{K}$. Then $\mathbb{E} \otimes_{\mathbb{K}} L$ is a Lie $\mathbb{E}$ -algebra. $\square$*

Our discussion of representation and our ultimate definition of Lie algebras immediately give

thm-ad-repn **(1.5).** THEOREM. *For any Lie $\mathbb{K}$ -algebra L , the map $\text{ad}: L \longrightarrow \text{End}_{\mathbb{K}}^-(L)$ is a representation of L . The kernel of this representation is the center of L*

$$Z(L) = \{ z \in L \mid [z, a] = 0, \text{ for all } a \in L \}. \quad \square$$

As was the case in Proposition (1.3) the small additional requirement that the center of A be trivial gives an easy proof that A has a faithful representation which has finite dimension provided A does. Far deeper is:

(1.6). THEOREM.

thm-ado

- (a) (PBW THEOREM) *Every Lie algebra has a faithful representation as a linear Lie algebra.*
- (b) (ADO-IWASAWA THEOREM) *Every finite dimensional Lie algebra has a faithful representation as a finite dimensional linear Lie algebra.* $\square\square$

Both these theorems are difficult to prove, although we will return to the easier PBW Theorem later. Notice that the Ado-Iwasawa Theorem is not an immediate consequence of PBW. Indeed the representation produced by the PBW Theorem is almost always a representation on an infinite dimensional space.

For Jordan algebras, the efforts of this section are largely a failure. In particular the adjoint action of a Jordan algebra A on itself does not give a representation in $\text{End}_{\mathbb{K}}^+(A)$. (Exercise.)

Jordan algebras that are (isomorphic to) linear Jordan algebras are usually called *special Jordan algebras*, while those that are not linear are the *exceptional Jordan algebras*.¹ A.A. Albert [Alb34] proved that the Albert algebras—the dimension 27 Jordan $\mathbb{K}$ -algebras described in Section 1.3—are exceptional rather than special. Indeed Cohn [Coh54] proved that Albert algebras are not even quotients of special algebras. Results of Birkhoff imply that the category of images of special Jordan algebras is varietal and does not contain the Albert algebras, but it is unknown what additional identical relations suffice to define this category.

¹So, taking a page out of the Montessori book, there are exactly two types of Jordan algebras: those that are special and those that are exceptional.

1.5 Problems

prob-lin-alt (1.7). PROBLEM.

- (a) Give two linear identities that characterize alternative $\mathbb{K}$ -algebras when $\text{char}\mathbb{K} \neq 2$.
- (b) Let A be an alternative $\mathbb{K}$ -algebra and $\mathbb{E}$ an extension field over $\mathbb{K}$. Prove that $\mathbb{E} \otimes_{\mathbb{K}} A$ is an alternative $\mathbb{E}$ -algebra.

prob-jord-not-rep (1.8). PROBLEM. Let A be an associative $\mathbb{K}$ -algebra for $\mathbb{K}$ a field of characteristic not equal to 2.

- (a) Prove that in general the adjoint action of a Jordan algebra does not give a representation. Consider specifically the Jordan algebra $A^+ = (A, \circ)$ and its adjoint map $\text{ad}: A^+ \rightarrow \text{End}_{\mathbb{K}}^+(A)$ where you can compare $\text{ad}_{a \circ a}$ and $\text{ad}_a \circ \text{ad}_a$.
- (b) Consider the two families of maps from A to itself:

$$L_a: x \mapsto a \circ x = \frac{1}{2}(ax + xa)$$

and

$$U_a: x \mapsto axa.$$

Prove that the $\mathbb{K}$ -subspace V of A is invariant under all L_a , for $a \in V$, if and only if it is invariant under all U_a , for $a \in V$.

HINT: The two parts of this problem are not unrelated.

REMARK. Observe that saying V is invariant under the L_a is just the statement that V is a Jordan subalgebra of $\text{End}_{\mathbb{K}}^+(A)$, the map L_a being the adjoint. Therefore the problems tells us that requiring U_a -invariance is another way of locating Jordan subalgebras, for instance the important and motivating spaces of hermitian matrices $H_n(\mathbb{C})$ in $\text{Mat}_n(\mathbb{C})$.

The crucial thing about U_a is that division by 2 is not needed. Therefore the maps U_a and their properties can be, and are, used to extend the study of Jordan algebras to include characteristic 2. The appropriate structures are called quadratic Jordan algebras, although some care must be taken as the “multiplication” $a \star x = U_a(x)$ is not bilinear. It is linear in its second variable but quadratic in its first variable; for instance $(\alpha a) \star x = \alpha^2(a \star x)$ for $\alpha \in \mathbb{K}$.

Chapter 2

Examples of Lie algebras

We give many examples of Lie algebras $(\mathbb{K}L, [\cdot, \cdot])$. These also suggest the many contexts in which Lie algebras are to be found.

2.1 Abelian algebras

Any $\mathbb{K}$ -vector space V is a Lie $\mathbb{K}$ -algebra when provided with the trivial product $[v, w] = 0$ for all $v, w \in V$. These are the *abelian Lie algebras*.

Class 4: 11 Sept '15
cancelled
Class 5: 14 Sept '15

2.2 Generators and relations

As with groups and most other algebraic systems, one effective way of producing examples is by providing a generating set and a collection of relations among the generators. For a $\mathbb{K}$ -algebra that would often be through supplying a basis $\mathcal{V} = \{v_i \mid i \in I\}$ together with appropriate equations restricting the various associated c_{ij}^k .

For a Lie algebra, the Jacobi Identity is linear and leads to ([Exercise.](#)) the equations:

$$\sum_k c_{ij}^k c_{kl}^m + c_{jl}^k c_{kl}^m + c_{li}^k c_{kl}^m = 0,$$

for all $i, j, l, m \in I$.

The law $[x, x] = 0$ gives the equations

$$c_{ii}^k = 0.$$

Since the null law is not linear, we also must include the consequences of its linearized skew law $[x, y] = -[y, x]$; so we also require

$$c_{ij}^k = -c_{ji}^k.$$

An algebra whose multiplication coefficients satisfy these three sets of equations is a Lie algebra. (Exercise.)

When presenting a Lie algebra it is usual to leave the non-Jacobi equations implicit, assuming without remark that the bracket multiplication is null and skew-symmetric.

For instance, we have the $\mathbb{K}$ -algebra $L = \mathbb{K}h \oplus \mathbb{K}e \oplus \mathbb{K}f$ where we state

$$[e, f] = h, \quad [h, e] = 2e, \quad [h, f] = -2f,$$

but in the future will not record the additional, necessary, but implied relations, which in this case are

$$[h, h] = [e, e] = [f, f] = 0, \quad [f, e] = -h, \quad [e, h] = -2e, \quad [f, h] = 2f.$$

Of course at this point in order to be sure that L really is a Lie algebra, we must verify the Jacobi Identity equations for all quadruples $(i, j, l, m) \in \{h, e, f\}^4$. (Exercise.)

2.3 Matrix algebras

Many Lie algebras occur naturally as matrix algebras. We have already mentioned $\text{Mat}_n^-(\mathbb{K})$. This is often written $\mathfrak{gl}_n(\mathbb{K})$, the *general linear algebra*, in part because it is the Lie algebra of the Lie group $\text{GL}_n(\mathbb{K})$; see Theorem (3.7)(a) below. The Gothic (or Fraktur) font is also a standard for Lie algebras.

A standard matrix calculation shows that $\text{tr}(MN) = \text{tr}(NM)$, so the subset of matrices of trace 0 is a dimension $n^2 - 1$ subalgebra $\mathfrak{sl}_n(\mathbb{K})$ of the algebra $\mathfrak{gl}_n(\mathbb{K})$, which itself has dimension n^2 . Indeed the *special linear algebra* $\mathfrak{sl}_n(\mathbb{K})$ is the *derived subalgebra* $[\mathfrak{gl}_n(\mathbb{K}), \mathfrak{gl}_n(\mathbb{K})]$ spanned by all $[M, N]$ for $M, N \in \mathfrak{gl}_n(\mathbb{K})$; see Section 4.1 below.

The subalgebras $\mathfrak{n}_n^+(\mathbb{K})$ and $\mathfrak{n}_n^-(\mathbb{K})$ are, respectively, composed of all strictly upper triangular and all strictly lower triangular matrices. Both have dimension $\binom{n}{2}$. Next let $\mathfrak{d}_n(\mathbb{K})$ and $\mathfrak{h}_n(\mathbb{K})$ be the abelian subalgebras of, respectively, all diagonal matrices (dimension n) and all diagonal matrices of trace 0 (dimension $n - 1$). Then

$$\mathfrak{gl}_n(\mathbb{K}) = \mathfrak{n}_n^+(\mathbb{K}) \oplus \mathfrak{d}_n(\mathbb{K}) \oplus \mathfrak{n}_n^-(\mathbb{K})$$

and

$$\mathfrak{sl}_n(\mathbb{K}) = \mathfrak{n}_n^+(\mathbb{K}) \oplus \mathfrak{h}_n(\mathbb{K}) \oplus \mathfrak{n}_n^-(\mathbb{K}).$$

This second decompositions and ones resembling it will be important later.

Within the Lie algebra $\mathfrak{sl}_2(\mathbb{K})$, consider the three elements

$$h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

so that $\mathfrak{h}_2(\mathbb{K}) = \mathbb{K}h$, $\mathfrak{n}_2^+(\mathbb{K}) = \mathbb{K}e$, and $\mathfrak{n}_2^-(\mathbb{K}) = \mathbb{K}f$, and

$$\mathfrak{sl}_2(\mathbb{K}) = \mathbb{K}h \oplus \mathbb{K}e \oplus \mathbb{K}f.$$

We then have (Exercise.)

$$[e, f] = h, \quad [h, e] = 2e, \quad [h, f] = -2f,$$

and the algebra presented at the end of the previous section is indeed a Lie algebra, namely a copy of $\mathfrak{sl}_2(\mathbb{K})$.

Let σ be an automorphism of $\mathbb{K}$ with fixed field $\mathbb{F}$. For the $\mathbb{K}$ -space V , the map $b: V \times V \rightarrow K$ is a σ -sesquilinear form provided it is biadditive and

$$b(pv, qw) = pb(v, w)q^\sigma$$

for all $v, w \in V$ and $p, q \in \mathbb{K}$. The form is additionally σ -hermitian provided

$$b(v, w) = b(w, v)^\sigma$$

always. The special case $\sigma = 1$ is that of *bilinear forms* and *symmetric bilinear forms*, respectively. The form b is *nondegenerate* provided 0 is the only vector v of V with $b(v, w) = 0$ for all $w \in V$.

Woowoo!!

perps and dimensions

For σ -hermitian b , the $\mathbb{F}$ -space of endomorphisms

$$\mathfrak{L}(V, b) = \{ x \in \text{End}_{\mathbb{K}}(V) \mid b(xv, w) = -b(v, xw) \text{ for all } v, w \in V \}$$

is then an Lie $\mathbb{F}$ -subalgebra of $\text{End}_{\mathbb{K}}(V)^\perp$. (Exercise.)

With $V = \mathbb{K}^n$ and $\text{End}_{\mathbb{K}}(V) = \text{Mat}_n(\mathbb{K}) = \mathfrak{gl}_n(K)$, we have some special cases of $\mathfrak{L} = \mathfrak{L}(V, b)$. $G = (b(e_i, e_j))_{i,j}$ be the *Gram matrix* of b on V (with respect to the usual basis). The condition above then becomes

$$\mathfrak{L}(V, b) = \{ M \in \text{Mat}_n(\mathbb{K}) \mid MG = -GM^{\sigma^\top} \}.$$

For simplicity's sake we assume that $\mathbb{K}$ has characteristic 0.

- (i) **Orthogonal algebras.** If $\sigma = 1$ and b is the usual nondegenerate form with an orthonormal basis, then **fix o versus so** $\mathfrak{so}_n(\mathbb{K}) = \mathfrak{L}$. As matrices,

Woowoo!!

$$\mathfrak{so}_n(\mathbb{K}) = \{ M \in \text{Mat}_n(\mathbb{K}) \mid M = -M^\top \}.$$

p and q ?

Woowoo!!

- (ii) **Symplectic algebras.** If $\sigma = 1$ and b is the usual nondegenerate symplectic form with symplectic basis $\mathcal{S} = \{ v_i, w_i \mid 1 \leq i \leq n \}$ subject to $b(v_i, v_j) = b(w_i, w_j) = 0$ and $b(v_i, w_j) = \delta_{i,j} = -b(w_j, v_i)$, then $\mathfrak{sp}_{2n}(\mathbb{K}) = \mathfrak{L}$. As matrices,

$$\mathfrak{sp}_{2n}(\mathbb{K}) = \{ M \in \text{Mat}_{2n}(\mathbb{K}) \mid MJ = -JM^\top \},$$

where J is the $2n \times 2n$ matrix with n blocks $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ down the diagonal.

- (iii) **Unitary algebras.** If σ has order 2 and b is the usual nondegenerate form with an orthonormal basis, then $\mathfrak{u}_n(\mathbb{K}) = \mathfrak{L}$. As matrices,

$$\mathfrak{u}_n(\mathbb{K}) = \{ M \in \text{Mat}_n(\mathbb{K}) \mid M = -M^{\sigma^\top} \}.$$

Furthermore $\mathfrak{su}_n(\mathbb{K}) = \mathfrak{u}_n(\mathbb{K}) \cap \mathfrak{sl}_n(\mathbb{K})$. Note that these are $\mathbb{F}$ -algebras rather than $\mathbb{K}$ -algebras.

The notation is not uniform. Especially, when $\mathbb{K} = \mathbb{R}$ the field is sometimes omitted, hence one may find

$$\mathfrak{gl}_n(\mathbb{R}) = \mathfrak{gl}(n, \mathbb{R}) = \mathfrak{gl}(n) = \mathfrak{gl}_n, \quad \mathfrak{sl}_n(\mathbb{R}) = \mathfrak{sl}(n, \mathbb{R}) = \mathfrak{sl}(n) = \mathfrak{sl}_n;$$

and

$$\mathfrak{so}_n(\mathbb{R}) = \mathfrak{so}(n, \mathbb{R}) = \mathfrak{so}(n) = \mathfrak{so}_n;$$

and similarly, with σ complex conjugation,

$$\mathfrak{u}_n(\mathbb{C}) = \mathfrak{u}(n, \mathbb{C}) = \mathfrak{u}(n) = \mathfrak{u}_n, \quad \mathfrak{su}_n(\mathbb{C}) = \mathfrak{su}(n, \mathbb{C}) = \mathfrak{su}(n) = \mathfrak{su}_n.$$

More confusingly, in the case of symplectic algebras the actual definition can vary as well as the notation; see [Tu11, p. 160].

2.4 Derivations

A *derivation* D on the $\mathbb{K}$ -algebra A is a linear transformation $D \in \text{End}_{\mathbb{K}}(A)$ with

$$D(fg) = fD(g) + D(f)g,$$

for all $f, g \in A$. This should be recognized as the *Leibniz product rule*. Clearly the set $\text{Der}_{\mathbb{K}}(A)$ is a $\mathbb{K}$ -subspace of $\text{End}_{\mathbb{K}}(A)$, but in fact this provides an amazing machine for constructing Lie algebras:

Class 6: 16 Sept '15

(2.1). THEOREM. $\text{Der}_{\mathbb{K}}(A) \leq \text{End}_{\mathbb{K}}^-(A)$. *That is, the derivation space is a Lie $\mathbb{K}$ -algebra under the bracket product.*

PROOF. Let $D, E \in \text{Der}_{\mathbb{K}}(A)$. Then, for all $f, g \in A$,

$$\begin{aligned} [D, E](fg) &= (DE - ED)(fg) = DE(fg) - ED(fg) \\ &= D(fEg + (Ef)g) - E(fDg + (Df)g) \\ &= D(fEg) + D((Ef)g) - E(fDg) - E((Df)g) \\ &= fDEg + DfEg + EfDg + (DEf)g \\ &\quad - fEDg - EfDg - DfEg - (EDf)g \\ &= fDEg - fEDg + (DEf)g - (EDf)g \\ &= f([D, E]g) + ([D, E]f)g. \quad \square \end{aligned}$$

The definition of derivations then tells us that the injection of $\text{Der}_{\mathbb{K}}(A)$ into $\text{End}_{\mathbb{K}}^-(A)$ gives a representation of the Lie *derivation algebra* $\text{Der}_{\mathbb{K}}(A)$ on the $\mathbb{K}$ -space A .

(2.2). COROLLARY. *The image of A under the adjoint representation is a subalgebra of $\text{Der}_{\mathbb{K}}(A)$ and $\text{End}_{\mathbb{K}}^-(A)$.*

PROOF. The image of A under ad is a $\mathbb{K}$ -subspace of $\text{End}_{\mathbb{K}}(A)$ by our very first Lemma (1.1). It remains to check that each ad_a is a derivation of A .

We start from the Jacobi Identity:

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

hence

$$-[[y, z], a] = [[a, y], z] + [[z, a], y].$$

That is,

$$[a, [y, z]] = [[a, y], z] + [y, [a, z]],$$

or

$$\text{ad}_a[y, z] = [\text{ad}_a y, z] + [y, \text{ad}_a z]. \quad \square$$

The map ad_a is then an *inner derivation* of A , and the Lie subalgebra $\text{InnDer}_{\mathbb{K}}(A) = \{\text{ad}_a \mid a \in A\}$ is the *inner derivation algebra*.

We have an easy but useful observation:

(2.3). PROPOSITION. *Every linear transformation of $\text{End}_{\mathbb{K}}(A)$ is a derivation of the abelian Lie algebra A .* prop-ab-der

PROOF. For $D \in \text{End}_{\mathbb{K}}(A)$ and $a, b \in A$

$$D[a, b] = 0 = 0 + 0 = [Da, b] + [a, Db]. \quad \square$$

2.4.1 Derivations of polynomial algebras

(2.4). PROPOSITION. prop-poly-der

- (a) $\text{Der}_{\mathbb{K}}(\mathbb{K}) = 0$.
- (b) *If the $\mathbb{K}$ -algebra A has an identity element 1, then for each $D \in \text{Der}_{\mathbb{K}}(A)$ and each $c \in \mathbb{K}$ we have $D(c) = 0$.*
- (c) $\text{Der}_{\mathbb{K}}(\mathbb{K}[t]) = \{p(t)\frac{d}{dt} \mid p(t) \in \mathbb{K}[t]\}$, a Lie algebra of infinite $\mathbb{K}$ -dimension with basis $\{t^i\frac{d}{dt} \mid i \in \mathbb{N}\}$.

PROOF. Part (b) clearly implies (a).

- (b) Let $c = c1 \in \mathbb{K}1$. Then for all $x \in A$ and all $D \in \text{Der}_{\mathbb{K}}(A)$ we have

$$D(cx) = cD(x)$$

as D is a $\mathbb{K}$ -linear transformation. But D is also a derivation, so

$$D(cx) = cD(x) + D(c)x.$$

We conclude that $D(c)x = 0$ for all $x \in A$, and so $D(c) = 0$.

(c) Let $D \in \text{Der}_{\mathbb{K}}(A)$. By (b) we have $D(\mathbb{K}1) = 0$. As the algebra A is generated by 1 and t , the knowledge of $D(t)$ together with the product rule should give us everything. Set $p(t) = D(t)$.

We claim that $D(t^i) = p(t)it^{i-1}$ for all $i \in \mathbb{N}$. We prove this by induction on i , the result being clear for $i = 0, 1$. Assume the claim for $i - 1$. Then

$$\begin{aligned} D(t^i) &= D(t^{i-1}t) = t^{i-1}D(t) + D(t^{i-1})t \\ &= t^{i-1}p(t) + p(t)(i-1)t^{i-2}t = p(t)it^{i-1}, \end{aligned}$$

as claimed.

As D is a linear transformation, if $a(t) = \sum_{i=0}^m a_i t^i$, then

$$D(a(t)) = \sum_{i=0}^m a_i D(t^i) = \sum_{i=0}^m a_i p(t)it^{i-1} = p(t) \sum_{i=0}^m i a_i t^{i-1} = p(t) \frac{d}{dt} a(t),$$

completing the proposition. $\square$

In $\text{Der}_{\mathbb{K}}(\mathbb{K}[t])$ there is the subalgebra $A = \mathbb{K}h \oplus \mathbb{K}e \oplus \mathbb{K}f$ with $e = \frac{d}{dt}$, $h = -2t \frac{d}{dt}$, $f = -t^2 \frac{d}{dt}$, and relations (Exercise.)

$$[e, f] = h, \quad [h, e] = 2e, \quad [h, f] = -2f;$$

so we have $\mathfrak{sl}_2(\mathbb{K})$ again.

We next consider $\mathbb{K}[x, y]$. A similar argument to that of the proposition proves

$$\text{Der}_{\mathbb{K}}(\mathbb{K}[x, y]) = \left\{ p(x, y) \frac{\partial}{\partial x} + q(x, y) \frac{\partial}{\partial y} \mid p(x, y), q(x, y) \in \mathbb{K}[x, y] \right\}.$$

(See Problem (2.8).) We examine two special situations—a subalgebra and a quotient algebra.

- (i) Consider the Lie subalgebra that leaves each homogeneous piece of $\mathbb{K}[x, y]$ invariant. This subalgebra has basis

$$h_x = x \frac{\partial}{\partial x}, \quad e = x \frac{\partial}{\partial y}, \quad f = y \frac{\partial}{\partial x}, \quad h_y = y \frac{\partial}{\partial y}.$$

Set $h = h_x - h_y = x \frac{\partial}{\partial x} - y \frac{\partial}{\partial y}$. Then

$$[e, f] = h, \quad [h, e] = 2e, \quad [h, f] = -2f,$$

giving $\mathfrak{sl}_2(\mathbb{K})$ yet again. The 4-dimensional algebra $\mathbb{K}h_x \oplus \mathbb{K}h_y \oplus \mathbb{K}e \oplus \mathbb{K}f$ is isomorphic to $\mathfrak{gl}_2(\mathbb{K})$ with the correspondences

$$h_x = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad h_y = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Each homogeneous piece of $\mathbb{K}[x, y]$ carries a representation of $\mathfrak{gl}_2(\mathbb{K})$ and $\mathfrak{sl}_2(\mathbb{K})$ via restriction from the action of $\text{Der}_{\mathbb{K}}(\mathbb{K}[x, y])$. The degree m homogeneous component $\mathbb{K}[x, y]_m$ is then a cyclic $\mathbb{K}e$ - hence $\mathfrak{sl}_2(\mathbb{K})$ -module $M_0(m+1)$ of dimension $m+1$ with generator y^m . This will be important in Chapter 7.

- (ii) The algebra $\mathbb{K}[x, y]$ has as quotient the algebra $\mathbb{K}[x, x^{-1}]$ of all *Laurent polynomials* in x . A small extension of the arguments from Proposition (2.4)(c) ([Exercise.](#)) proves that $\text{Der}_{\mathbb{K}}(\mathbb{K}[x, x^{-1}])$ has $\mathbb{K}$ -basis consisting of the distinct elements

$$L_m = -x^{m+1} \frac{d}{dx} \quad \text{for } m \in \mathbb{Z}.$$

We write the generators in this form, since they then have the nice presentation

$$[L_m, L_n] = (m - n)L_{m+n}.$$

All the multiplication coefficients are integers. The $\mathbb{Z}$ -algebra with this presentation has infinite dimension. It is called the *Witt algebra* over $\mathbb{Z}$, just as its tensor with $\mathbb{K}$, $\text{Der}_{\mathbb{K}}(\mathbb{K}[x, x^{-1}])$, is the Witt algebra over $\mathbb{K}$.

2.4.2 Derivations of nonassociative algebras

We may also consider derivations of the nonassociative algebras we have encountered, specifically the octonion $\mathbb{K}$ -algebra O and (in characteristic not 2) its related Albert algebra—the exceptional Jordan $\mathbb{K}$ -algebra $H_3(O)$. The derivation algebra $\text{Der}_{\mathbb{K}}(O)$ has dimension 14 and is said to have *type* $\mathfrak{g}_2$ while $\text{Der}_{\mathbb{K}}(O)$ has dimension 52 and is said to have *type* $\mathfrak{f}_4$. Especially when $\mathbb{K}$ is algebraically closed and of characteristic 0 we have the uniquely determined algebras $\mathfrak{g}_2(\mathbb{K})$ and $\mathfrak{f}_4(\mathbb{K})$, respectively.

Class 7: 18 Sept '15

2.4.3 Vector fields

We shall see in the next chapter that the tangent space to a Lie group at the identity is a Lie algebra. As the group acts regularly on itself by translation, this space is isomorphic to the Lie algebra of invariant vector fields on the group.

Indeed often a *vector field* on the smooth manifold M is defined to be a derivation of the algebra $C^\infty(M)$ of all smooth functions; for instance, see [Hel01, p. 9]. Thus the space of all vector fields is the corresponding derivation algebra and so automatically has a Lie algebra structure.

For instance, the Lie group of rotations of the circle S^1 is the group $\text{SO}_2(\mathbb{R})$ of all matrices

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix},$$

which becomes $e^{i\theta}$ when we extend coefficients to the complex numbers. The corresponding spaces of invariant vector fields have dimension 1.

The space $C^\infty(S^1)$ of all smooth functions on the circle consists of those functions that can be expanded as convergent *Fourier series*

$$\sum_{m \in \mathbb{Z}} a_m \sin(m\theta) + b_m \cos(m\theta),$$

which after extension to $\mathbb{C}$ becomes the simpler

$$\sum_{m \in \mathbb{Z}} c_m e^{im\theta}.$$

This space has as a dense subalgebra the space of all *Fourier polynomials*, whose canonical basis is $\{e^{im\theta} \mid m \in \mathbb{Z}\}$.

The group of all complex orientation preserving diffeomorphisms of the circle (an “infinite dimensional Lie group”) is an open subset of $C^\infty_\mathbb{C}(S^1)$ and has as corresponding space of smooth vector fields (not just those that are invariant) all $f \frac{d}{d\theta}$ for f smooth. The dense Fourier polynomial subalgebra with basis $L_m = ie^{im\theta} \frac{d}{d\theta}$ then has

$$[L_m, L_n] = (m - n)L_{m+n},$$

giving the complex Witt algebra again.

2.5 Other constructions

2.5.1 Extensions

As we have seen and expect, subalgebras and quotients are ways of constructing new algebras out of old algebras. We can also extend old algebras to get new ones. As with groups, central extensions are important since the information we have about a given situation may come to us in projective rather than affine form.

The *Virasoro algebra* is a central extension of the complex Witt algebra. If W is the Witt $\mathbb{Z}$ -algebra, then

$$\text{Vir}_\mathbb{C} = (\mathbb{C} \otimes_\mathbb{Z} W) \oplus \mathbb{C}c$$

with $[w, c] = 0$ for all $w \in W$ and

$$[L_m, L_n] = (m - n)L_{m+n} + \delta_{m, -n} \frac{m(m^2 - 1)}{12} c.$$

The multiplication coefficients are half-integers.

The Virasoro algebra is important in applications to physics and other situations. As seen after Proposition (2.4), the Witt and Virasoro algebras both contain the subalgebra $\mathbb{C}L_{-1} \oplus \mathbb{C}L_0 \oplus \mathbb{C}L_1$ isomorphic to $\mathfrak{sl}_2(\mathbb{C})$. As we shall find starting in Section 6.2, large parts of the finite dimensional Lie algebra theory depend upon the construction of Lie subalgebras $\mathfrak{sl}_2(\mathbb{K})$. Similarly, the

infinite dimensional Lie algebras that come up in physics and elsewhere are often handled using Witt and Virasoro subalgebras, which are in a sense the infinite dimensional substitutes for the finite dimensional $\mathfrak{sl}_2(\mathbb{K})$.

Given a complex simple Lie algebra like $\mathfrak{sl}_2(\mathbb{C})$, the corresponding *affine Lie algebra* comes from a two step process. First extend scalars to the Laurent polynomials and second take an appropriate central extension. So:

$$\widehat{\mathfrak{sl}}_2(\mathbb{C}) = (\mathbb{C}[t, t^{-1}] \otimes_{\mathbb{C}} \mathfrak{sl}_2(\mathbb{C})) \oplus \mathbb{C}c$$

where the precise cocycle on the complex Lie algebra $\mathbb{C}[t, t^{-1}] \otimes_{\mathbb{C}} \mathfrak{sl}_2(\mathbb{C})$ that gives the extension is defined in terms of the Killing form on the algebra $\mathfrak{sl}_2(\mathbb{C})$. (See Chapter 6 below.)

In such situations it is more usual to write the Lie algebra $\mathbb{C}[t, t^{-1}] \otimes_{\mathbb{C}} \mathfrak{sl}_2(\mathbb{C})$ instead as $\mathfrak{sl}_2(\mathbb{C}) \otimes_{\mathbb{C}} \mathbb{C}[t, t^{-1}]$, viewing its elements as “Laurent polynomials” with coefficients from the algebra $\mathfrak{sl}_2(\mathbb{C})$.

It is also possible to form split extensions of Lie algebras, with derivations playing the role that automorphisms play in group extensions. (See Section 4.3.) The canonical derivation $\frac{d}{dt}$ on the Laurent polynomials induces a derivation of the affine algebra which is then used to extend the affine algebra so that it has codimension 1 in the corresponding *Kac-Moody Lie algebra*.

2.5.2 Embeddings

We saw above that derivations of octonion and Jordan algebras give new Lie algebras. Tits, Kantor, and Koecher [Tit??] used these same nonassociative algebras to construct Lie algebras that are still more complicated. In particular, the space

$$\mathrm{Der}_{\mathbb{C}}(\mathbb{O}^{\mathrm{sp}}(\mathbb{C})) \oplus (\mathbb{O}^{\mathrm{sp}}(\mathbb{C})_0 \otimes_{\mathbb{C}} \mathrm{H}_{\mathbb{C}}(\mathbb{O}^{\mathrm{sp}}(\mathbb{C}))_0) \oplus \mathrm{Der}_{\mathbb{C}}(\mathrm{H}_{\mathbb{C}}(\mathbb{O}^{\mathrm{sp}}(\mathbb{C})))$$

of dimension $14 + (8-1) \times (27-1) + 52 = 248$ can be provided with a Lie algebra product (extending that of the two derivation algebra pieces) that makes it into the Lie algebra $\mathfrak{e}_8(\mathbb{C})$. Here $\mathbb{O}^{\mathrm{sp}}(\mathbb{C})_0$ is $1^{\perp}$ in $\mathbb{O}^{\mathrm{sp}}(\mathbb{C})$ and $\mathrm{H}_{\mathbb{C}}(\mathbb{O}^{\mathrm{sp}}(\mathbb{C}))_0$ is a similarly defined subspace of codimension 1 in $\mathrm{H}_{\mathbb{C}}(\mathbb{O}^{\mathrm{sp}}(\mathbb{C}))$. The Lie algebra $\mathfrak{e}_8(\mathbb{C})$ furthermore has the important subalgebras $\mathfrak{e}_6(\mathbb{C})$ of dimension 78 and $\mathfrak{e}_7(\mathbb{C})$ of dimension 133.

2.5.3 Nilpotent groups

Let G be a nilpotent group with *lower central series*

$$G = L^1(G) \supseteq L^2(G) \supseteq \cdots \supseteq L^{n+1}(G) = 1$$

where $L^{k+1}(G)$ is defined as $[G, L^k(G)]$. For each $1 \leq k \leq n$ set

$$L_k = L^k(G)/L^{k+1}(G),$$

an abelian group as is the sum

$$L = \bigoplus_{k=1}^n L_k .$$

As G is nilpotent, always

$$[L^i(G), L^j(G)] \leq L^{i+j}(G) .$$

This provides the relations that turn the group $L = L_G$ into a *Lie ring*—we do not require it to be free as $\mathbb{Z}$ -module—within which we have

$$[L_i, L_j] \leq L_{i+j} .$$

Certain questions about nilpotent groups are much more amenable to study in the context of Lie rings and algebras [Hig58]. A particular important instance is the *Restricted Burnside Problem*, which states that an m -generated finite nilpotent group of exponent e has order less than or equal to some function $f(m, e)$, dependent only on m and e . E. Zelmanov received a Fields Medal in 1994 for the positive solution of the Restricted Burnside Problem. His proof [Zel97] makes heavy use of Lie methods.

2.6 Problems

(2.5). PROBLEM. Classify up to isomorphism all Lie $\mathbb{K}$ -algebras of dimension 2. (Of course, the abelian algebra gives the only isomorphism class in dimension 1.) prob-la-dim2-bis

Woowoo!!

(2.6). PROBLEM. Prove that over an arbitrary field $\mathbb{K}$ of characteristic not 2, the Lie algebra $\mathfrak{sl}_2(\mathbb{K})$ is isomorphic to $\mathfrak{so}_{2,1}(\mathbb{K})$. prob-sl2-so21

(2.7). PROBLEM. Find all subalgebras of $\mathfrak{sl}_2(\mathbb{K})$ that contain the subalgebra $H = \mathbb{K}h$. prob-sl2-sub
HINT: Small characteristic can produce anomalous results.

(2.8). PROBLEM. Calculate $\text{Der}_{\mathbb{K}}(\mathbb{K}[x_1, \dots, x_n])$. prob-poly-der-bis

(2.9). PROBLEM. Consider the matrix subgroup $\text{UT}_n(\mathbb{K})$ of $\text{GL}_n(\mathbb{K})$, consisting of the upper unitriangular matrices—those which have 1's on the diagonal, anything above the diagonal, and 0's below the diagonal. prob-la-upper-tri

- (a) Prove that $G = \text{UT}_n(\mathbb{K})$ is a nilpotent group.
- (b) Starting with this group G , construct the Lie algebra $L = L_G$ as in Section 2.5.3. Prove that L is isomorphic to the Lie algebra $\mathfrak{n}_n^+(\mathbb{K})$.

(2.10). PROBLEM. Consider the subgroup $X_n(\mathbb{K})$ of upper unitriangular matrices that have 1's on the diagonal, anything in the nondiagonal part of the first row and last column, and 0's elsewhere. prob-heisenberg-extraspec

- (a) By the previous problem $X = X_n(\mathbb{K})$ is nilpotent. Prove that for $n \geq 2$ it has nilpotence class exactly 2 and that its center is equal to its derived group and consists only of those matrices with 1's down the diagonal and the only other nonzero entries found in the upper-righthand corner.
- (b) Starting with this group X , construct the Lie algebra $L = L_X$ as in Section 2.5.3. Prove that L is isomorphic to the Lie algebra on the space

$$M = \mathbb{K}z \oplus \bigoplus_{i=1}^{n-1} (\mathbb{K}x_i \oplus \mathbb{K}y_i)$$

with relations given by

$$[x_i, y_i] = -[y_i, x_i] = z,$$

for all i , and all other brackets among generators equal to 0.

REMARK. This Lie algebra is the Heisenberg algebra of dimension $2n - 1$ over $\mathbb{K}$.

Chapter 3

Lie groups

(N. Jacobson [Jac79, p. 1]:) *The theory of Lie algebras is an outgrowth of the Lie theory of continuous groups.*

(R. Carter [Car05, p. xiii]:) *Lie algebras were originally introduced by S. Lie as algebraic structures used for the study of Lie groups.*

It would be wrong for us to talk at length about Lie algebras without devoting at least some time to the way in which they arise in the theory of Lie groups. We do that in an abbreviated form in this chapter.

For us, Lie's work and the work that it motivated contain two basic observations:

- (i) *If G is a Lie group, then the tangent space to the identity is a Lie algebra $\Lambda(G)$.*
- (ii) *The representation theory of the Lie group G and of the Lie algebra $\Lambda(G)$ are essentially the same.*

The second observation displays real progress, since a Lie algebra is a linear object whereas the Lie group is not. This is the same advantage obtained in the passage from a nilpotent group to its associated Lie ring in Section 2.5.3.

This chapter is included in order to place Lie algebras in one of their most important contexts, historically and practically. Its material will not be used in the rest of the notes or course. Therefore for ease of presentation we assume uniformly throughout that the vector spaces, groups, and algebras we examine are defined over the real numbers. Given our later focus on algebraically closed fields of characteristic 0, it might make more sense to restrict to the complex case; but that would require more sophisticated calculus/analysis than we care to use.

3.1 Representation theory as spectral theory

The two observations beg the question, “What is so good about representation theory?” After all, many of our important Lie groups and algebras are already defined in terms of matrices. Why worry about more representations?

Lie and those who followed him were interested in using Lie theory to solve problems, and it is often easier to solve a problem in pieces rather than all at once. An important example is the analysis of the action of a linear transformation in terms of its eigenspaces. Such decompositions are collected together under the heading of spectral theory, and they are served by various canonical form results.

The representation theory of groups (and other algebras) can be thought of as a general form of spectral or canonical form theory. If the initial, say physical, statement of a problem has some inherent symmetry, then that symmetry should also be evident in the space of solutions. Lie noted that this action could be exploited to decompose the solution space and so perhaps find nice descriptions for the solutions. At the heart of matrix canonical form results is the feeling that matrices containing lots of zeros are the easiest to deal with.

Lie was interested in particular in solving differential equations. Dresner [Dre99, p. 16] shows how, starting from the differential equation

$$\frac{d}{dx}y = -\frac{y(y^2 - x)}{x},$$

once one has noticed that the solution set is invariant under the change of variables

$$x = x_1 \longrightarrow x_s = e^s x \quad y = y_1 \longrightarrow y_s = e^{s/2} y,$$

for all $s \in \mathbb{R}$, it is relative easy to construct an integrating factor

$$\varphi(x, y) = \left(xy^3 - \frac{x^2 y}{2} \right)^{-1}$$

and so reach the closed form solution set

$$y = x(2x + c)^{1/2}.$$

The displayed symmetry group $\{e^s \mid s \in \mathbb{R}\} \simeq (\mathbb{R}, +)$ is continuous and even smooth in its variable. This type of symmetry is evident in many physical situations, and this led Lie (and others) to the study of smooth groups and their representations. We shall see in Section 3.4 that the most basic Lie group $(\mathbb{R}, +)$ is also one of the most important.

3.2 Lie groups and Hilbert’s Fifth Problem

A *Lie group* is a smooth manifold G that is also a group. These two conditions are linked by the requirements that the group multiplication $m: G \times G \longrightarrow G$

given by $m(x, y) = xy$ and the group inverse map $i: G \rightarrow G$ given by $i(x) = x^{-1}$ are smooth maps on the manifold. Here (recalling that we are speaking of real manifolds) by *smooth* we mean C^∞ . (For a complex manifold, smooth means holomorphic.)

Examples are provided by the *closed subgroups* of $GL_n(\mathbb{R})$: those subgroups containing the limit of every sequence of group matrices for which that limit exists and is invertible. This already might be a surprise, since closure is a topological property, determined only by examining C^0 continuity issues. The C^0 condition is very weak when compared to the smooth C^∞ assumptions of the manifold definition.

If G is a Lie group, then certainly

- (i) G is a topological group (that is, the maps m and i are continuous) and
- (ii) G is locally a finite dimensional Euclidean space.

One reading of *Hilbert's Fifth Problem* is that, in fact, the Lie groups are exactly the locally Euclidean topological groups. Once made precise, this version of the Fifth Problem was proven by Montgomery and Zippin [MoZi55] and Gleason [Gle52] in 1952. (See [Tao14] for more.)

Cartan first proved that closed subgroups of $GL_n(\mathbb{R})$ are Lie groups. As such, it is reasonable to focus on such examples when initially discussing Lie groups. This is the approach taken by several modern introductions to Lie groups [Eld15, Hal15, How83, vNe29, Ros02, Sti08, Tap05] and is largely what we do here. In particular, those not comfortable with manifolds need not worry—just focus on closed subgroups of $GL_n(\mathbb{R})$.

Essentially everything we prove (or state) goes over to the general case, although some of the definitions and proofs would require more subtlety. In particular, in place of the concrete functions $\exp$ and $\log$ provided by convergent power series of matrices, one appeals to the uniqueness of solutions for appropriate ordinary differential equations and to the Inverse Function Theorem; see [CSM95, pp. 69–74].

3.3 Some matrix calculus

For the matrix $M = (m_{ij})_{ij} \in \text{Mat}_{k,l}(\mathbb{R})$, set $|M| = \sqrt{\sum_{i,j} m_{ij}^2}$. This is the standard Euclidean norm on $\mathbb{R}^{kl}$; especially for $k = l = 1$ we have the usual $|(m)| = |m|$. We can then define limits of matrix functions, using this norm to determine “closeness.” In turn, this gives meaning to statements that a function from one matrix space to another is continuous, for instance in our discussion above of multiplication and inversion in Lie groups.

For smoothness we need derivatives as well. The usual derivative of $f(x)$ at $x = a$ is given by

$$\lim_{t \rightarrow 0} \frac{f(t+a) - f(a)}{t} = f'(a) = \left. \frac{df}{dx} \right|_{x=a}.$$

If we rewrite this as

$$\lim_{t \rightarrow 0} \frac{f(t+a) - (f(a) + f'(a)t)}{t} = 0,$$

we are observing that near a (near $t = 0$), the line $f(a) + f'(a)t$ is a good approximation to the function $f(t+a)$. This motivates the following definition of the *derivative of a matrix function*; see [Spi65, p. 16].

The linear transformation $D: \text{Mat}_{k,l}(\mathbb{R}) \rightarrow \text{Mat}_{m,n}(\mathbb{R})$ is the derivative at A of the matrix function $F: \text{Mat}_{k,l}(\mathbb{R}) \rightarrow \text{Mat}_{m,n}(\mathbb{R})$ provided

$$\lim_{T \rightarrow 0} \frac{|F(T+A) - F(A) - D(T)|}{|T|} = 0.$$

As derivatives are locally determined, to calculate the derivative of F at A we only need to know F on some neighborhood of A in $\text{Mat}_{k,l}(\mathbb{R})$.

This definition is the appropriate one for checking properties, but our applications later in this chapter will only be concerned with the special case $k = l = 1$ and $m = n$. That is, we will consider matrix functions $F: I \rightarrow \text{Mat}_n(\mathbb{R})$ for some open interval I in $\mathbb{R}$ that contains the point a . There we will use the equivalent but more familiar formulation

$$F'(a) = \lim_{t \rightarrow 0} \frac{F(t+a) - F(a)}{t} \in \text{Mat}_n(\mathbb{R}).$$

Once we have checked that matrix limits and derivatives behave as hoped and expected¹ (see, for instance, [Eld15], [Hal15]), we have

prop-mat-calc-omnibus **(3.1).** PROPOSITION.

- (a) If the power series $A(t) = \sum_{k=0}^{\infty} A_k t^k$ converges for all $|t| < r$, then its derivative $A'(t) = \sum_{k=0}^{\infty} k A_k t^{k-1}$ also converges for all $|t| < r$.
- (b) $\exp(A) = \sum_{k=0}^{\infty} \frac{1}{k!} A^k$ converges for all $A \in \text{Mat}_n(\mathbb{R})$. For $A, B \in \text{Mat}_n(\mathbb{R})$ with $[A, B] = 0$ we have $\exp(A+B) = \exp(A)\exp(B)$. Especially $I = \exp(A)\exp(-A)$.
- (c) For all $A \in \text{Mat}_n(\mathbb{R})$ the unique solution of the matrix ordinary differential equation

$$f'(t) = f(t)A, \quad f(0) = I$$

is $f(t) = \exp(tA)$.

- (d) $\log(1+X) = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{1}{k} X^k$ converges for all X with $|X| < 1$. For $|X| < 1$, we have $\exp(\log(1+X)) = 1+X$. $\square$

¹Exercise: Check the matrix versions of Leibniz' $\frac{d}{dt}(p(t)q(t)) = p(t)q'(t) + p'(t)q(t)$ and of the chain rule.

It is important that we can only guarantee $\exp(A+B) = \exp(A)\exp(B)$ when the matrices A and B commute. When they do, the corresponding power series multiplication goes through exactly as in the standard case. But if they do not commute, then things like BAB and AB^2 on the lefthand side can be different, so collecting of like terms is greatly restricted.

Also note that we are defining the logarithm via its Taylor series, rather than the usual calculus definitions that use an integral or that legislate it to be the inverse function for the exponential. Thus for us it is only defined (convergent) near the identity. This will be good enough. (See the proof of Proposition (3.2) below.)

The next proposition is an extension of the familiar result/definition from calculus

$$\exp(a) = \lim_{k \rightarrow \infty} \left(1 + \frac{a}{k}\right)^k,$$

which is the special case $n = 1$ and $g(t) = 1 + at$ of the proposition.

(3.2). PROPOSITION. *Let $g: (-r, r) \rightarrow \mathrm{GL}_n(\mathbb{R})$ be differentiable at 0 with $g(0) = I$ and $g'(0) = A$. Then $\lim_{k \rightarrow \infty} g\left(\frac{1}{k}\right)^k = \exp(A)$.* prop-defn-exp

PROOF. Set $q(t) = \log(g(t))$ (for t small enough so that $|g(t) - I| < 1$). By the chain rule, $q'(t) = g'(t)g(t)^{-1}$ (again for small t), so $q(0) = 0$ and $q'(0) = A$. Therefore by the definition of the matrix derivative of q at 0 (with $k = t^{-1}$)

$$\begin{aligned} 0 &= \lim_{t \rightarrow 0} \frac{\log g(t) - tA}{t} = \lim_{k \rightarrow \infty} k (\log g(k^{-1}) - k^{-1}A) \\ &= -A + \lim_{k \rightarrow \infty} k \log g(k^{-1}). \end{aligned}$$

That is, $A = \lim_{k \rightarrow \infty} k \log g(k^{-1})$. As exponentiation is everywhere continuous,

$$\begin{aligned} \exp(A) &= \exp\left(\lim_{k \rightarrow \infty} k \log g(k^{-1})\right) = \lim_{k \rightarrow \infty} \exp(k \log g(k^{-1})) \\ &= \lim_{k \rightarrow \infty} g\left(\frac{1}{k}\right)^k, \end{aligned}$$

as desired. □

3.4 One-parameter subgroups

If G is a Lie group, then a *one-parameter subgroup* of G is a continuous homomorphism $\varphi: (\mathbb{R}, +) \rightarrow G$. This links the weakest C^0 continuity property of G (and $\mathbb{R}$) with group theoretic structure. We shall see that this forces very strong continuity—not just C^∞ but C^ω (analytic). For every $A \in \mathrm{Mat}_n(\mathbb{R})$, the analytic map $\varphi_A: \mathbb{R} \rightarrow \mathrm{Mat}_n(\mathbb{R})$ given by $\varphi_A(t) = \exp(tA)$ is a one-parameter subgroup of $\mathrm{GL}_n(\mathbb{R})$ by Proposition (3.1). Surprisingly, the converse is true. This can be viewed as an important special case of Hilbert's Fifth Problem.

(3.3). THEOREM. *Let $\varphi: (\mathbb{R}, +) \rightarrow G$ be a one-parameter subgroup of the closed subgroup G of $\mathrm{GL}_n(\mathbb{R})$. Then there is a unique matrix $A \in \mathrm{Mat}_n(\mathbb{R})$ with $\varphi(t) = \exp(tA)$ for all $t \in \mathbb{R}$. In particular φ is C^∞ and indeed analytic. We have $A = \varphi'(0) = \frac{d}{dt}\varphi|_{t=0}$.* thm-one-param

PROOF. Our proof follows [Eld15]. It has two parts. We first prove that φ is differentiable and then prove that it is an exponential.

Set $F(t) = \int_0^t \varphi(u) du$. As φ is continuous, F is differentiable with $F(0) = 0$ and $F'(t) = \varphi(t)$, hence $F'(0) = I$. We use the fact that φ is a homomorphism and make the change of variable $v = u - t$ to find

$$\begin{aligned} F(t+s) &= \int_0^{t+s} \varphi(u) du \\ &= \int_0^t \varphi(u) du + \int_t^{t+s} \varphi(u) du \\ &= \int_0^t \varphi(u) du + \int_t^{t+s} \varphi(t)\varphi(u-t) du \\ &= \int_0^t \varphi(u) du + \varphi(t) \int_t^{t+s} \varphi(u-t) du \\ &= \int_0^t \varphi(u) du + \varphi(t) \int_0^s \varphi(v) dv \\ &= F(t) + \varphi(t)F(s). \end{aligned}$$

Next note that

$$I = F'(0) = \lim_{s \rightarrow 0} \frac{F(s) - F(0)}{s - 0} = \lim_{s \rightarrow 0} \frac{F(s)}{s},$$

hence

$$1 = \det I = \det \left(\lim_{s \rightarrow 0} \frac{F(s)}{s} \right) = \lim_{s \rightarrow 0} (s^{-n} \det F(s)),$$

as $\det$ is continuous. Especially, for some small s_0 we must have $\det F(s_0) \neq 0$ and so $F(s_0)$ is invertible. But then the above tells us that

$$\varphi(t) = (F(t+s_0) - F(t))F(s_0)^{-1}$$

is differentiable, as desired for the first part of our argument.

We now have φ differentiable with $\varphi(0) = I$. As φ is a homomorphism

$$\begin{aligned} \varphi'(t) &= \lim_{h \rightarrow 0} \frac{\varphi(t+h) - \varphi(t)}{h} = \lim_{h \rightarrow 0} \frac{\varphi(t)\varphi(h) - \varphi(t)}{h} \\ &= \lim_{h \rightarrow 0} \varphi(t) \frac{\varphi(h) - I}{h} = \varphi(t) \lim_{h \rightarrow 0} \frac{\varphi(h) - \varphi(0)}{h} \\ &= \varphi(t)\varphi'(0) \end{aligned}$$

That is, for $\varphi'(0) = A$ the function $\varphi(t)$ solves the ordinary differential equation

$$\varphi'(t) = \varphi(t)A \quad \text{and} \quad \varphi(0) = I.$$

By the omnibus Proposition (3.1)(c) we have $\varphi(t) = \exp(tA)$, as claimed. $\square$

(3.4). COROLLARY. $\det(\exp(tA)) = e^{t \operatorname{tr}(A)}$. cor-det-exp

PROOF. The map $t \mapsto \det \exp(tA)$ is a one-parameter subgroup of $\operatorname{GL}_1(\mathbb{R})$. (Exercise.) Therefore there is a nonzero $a \in \mathbb{R}$ with $\det \exp(tA) = e^{ta}$ for $a = \frac{d}{dt} \det \exp(tA)|_{t=0}$.

We have $\exp(tA) = I + tA + t^2B(t)$ (for an appropriate convergent power series $B(t)$), hence with $A = (a_{ij})_{ij}$ the standard expansion of the determinant gives

$$\det \exp(tA) = 1 + t(a_{11} + \cdots + a_{nn}) + t^2c(t) = 1 + t \operatorname{tr}(A) + t^2c(t).$$

Therefore $a = \frac{d}{dt} \det \exp(tA)|_{t=0} = \operatorname{tr}(A)$. $\square$

Let G be a closed subgroup of $\operatorname{GL}_n(\mathbb{R})$. A *curve* in G is a differentiable map $c: J \rightarrow G$, for some open interval J in $\mathbb{R}$. In particular, a one-parameter subgroup is a special type of curve.

There are several ways of defining the *tangent space* at the identity element I of the group G . We offer two—a relatively weak C^1 (differentiable) version and a very strong C^ω (analytic) condition. Set

- (i) $T_I(G) = \{ c'(0) \mid \text{curve } c: (-r, r) \rightarrow G, \text{ some } r \in \mathbb{R}^+, c(0) = I \};$
- (ii) $\Lambda(G) = \{ A \mid \exp(tA) \leq G \}.$

Clearly $\Lambda(G) \subseteq T_I(G)$, but we will prove in Theorem (3.6) below that we have equality. Again, this is in the spirit of Hilbert's Fifth Problem.

We first show that the tangent space is indeed a subspace.

(3.5). LEMMA. $T_I(G)$ is a subspace of $\operatorname{Mat}_n(\mathbb{R})$. lem-tan-subspace

PROOF. Let $A, B \in T_I(G)$ and $a, b \in \mathbb{R}$. We must show that $aA + bB \in T_I(G)$. Let differentiable

$$g: (-q, q) \rightarrow G, \quad g(0) = I, \quad g'(0) = A$$

and

$$h: (-s, s) \rightarrow G, \quad h(0) = I, \quad h'(0) = B$$

testify to $A, B \in T_I(G)$.

First consider $c(t) = h(bt)$ on $(-r, r)$ with $r = |b^{-1}s|$ ($= \infty$ for $b = 0$). Then

$$c(0) = h(0) = I \quad \text{and} \quad c'(0) = bh'(0) = bB,$$

so $T_I(G)$ is closed under scalar multiplication.

It remains to prove $A+B \in T_I(G)$. For $r = \frac{1}{2} \min(q, s)$, the curve $c: (-r, r) \rightarrow G$ given by

$$c(t) = \frac{1}{2}(g(2t) + h(2t)),$$

has

$$c(0) = \frac{1}{2}(g(0) + h(0)) = \frac{1}{2}(I + I) = I.$$

and

$$c'(0) = \frac{1}{2}(2g'(0) + 2h'(0)) = \frac{1}{2}(2A + 2B) = A + B.$$

Thus $A + B \in T_I(G)$ as desired. $\square$

thm-equal-tan **(3.6).** THEOREM. $\Lambda(G) = T_I(G)$.

PROOF. We have already pointed out that $\Lambda(G) \subseteq T_I(G)$. Now, for fixed but arbitrary $t \in \mathbb{R}$ and for each $B \in T_I(G)$, we must prove that the matrix $\exp(tB)$ is in G , as then $t \mapsto \exp(tB)$ will be a one-parameter subgroup of G , exhibiting $B \in \Lambda(G)$ and providing the reverse containment $\Lambda(G) \supseteq T_I(G)$. By the previous lemma $T_I(G)$ is a $\mathbb{R}$ -space, so it is enough to prove that $\exp(A) \in G$ for all $A \in T_I(G)$.

For some $r \in \mathbb{R}^+$, let the curve $g: (-r, r) \rightarrow G$ have $g(0) = I$ and $g'(0) = A$. Then for all integral k greater than some N we have $g(\frac{1}{k}) \in G$. As G is a group, in turn $g(\frac{1}{k})^k \in G$. By Proposition (3.2), this gives $\lim_{k \rightarrow \infty} g(\frac{1}{k})^k = \exp(A)$, which is always invertible. As G is a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$, we conclude $\exp(A) \in G$ as desired. $\square$

It is now appropriate for us to define the *tangent space* at the identity element I of the the group G , closed in $\mathrm{GL}_n(\mathbb{R})$, to be the $\mathbb{R}$ -space $\Lambda(G) = T_I(G)$.

Of course $\mathrm{GL}_n(\mathbb{R})$ is closed in itself. Additionally $\mathrm{SL}_n(\mathbb{R})$ is closed in $\mathrm{GL}_n(\mathbb{R})$ as it consists of all matrices X with $\det(X) - 1 = 0$.

thm-la-gl-sl **(3.7).** THEOREM.

- (a) $\Lambda(\mathrm{GL}_n(\mathbb{R})) = \mathfrak{gl}_n(\mathbb{R})$ and $\langle \exp(tA) \mid A \in \mathfrak{gl}_n(\mathbb{R}) \rangle = \mathrm{GL}_n(\mathbb{R})^+$, the subgroup of index 2 in $\mathrm{GL}_n(\mathbb{R})$ of all matrices with positive determinant.
- (b) $\Lambda(\mathrm{SL}_n(\mathbb{R})) = \mathfrak{sl}_n(\mathbb{R})$ and $\langle \exp(tA) \mid A \in \mathfrak{sl}_n(\mathbb{R}) \rangle = \mathrm{SL}_n(\mathbb{R})$.

PROOF. The equality $\Lambda(\mathrm{GL}_n(\mathbb{R})) = \mathrm{Mat}_n(\mathbb{R}) = \mathfrak{gl}_n(\mathbb{R})$ is clear from Proposition (3.1)(b).

We next consider $A \in \Lambda(\mathrm{SL}_n(\mathbb{R}))$. By Corollary (3.4), for the one-parameter subgroup $\exp(tA)$ of $\mathrm{SL}_n(\mathbb{R})$ we have

$$1 = \det(\exp(tA)) = e^{t \operatorname{tr}(A)}.$$

That is, $\operatorname{tr}(A) = 0$ and $A \in \mathfrak{sl}_n(\mathbb{R})$. Conversely, for $A \in \mathfrak{sl}_n(\mathbb{R})$, by the same corollary

$$1 = e^{t \operatorname{tr}(A)} = \det(\exp(tA)).$$

This is true for arbitrary t , so $t \mapsto \exp(tA)$ is a one-parameter subgroup of $\mathrm{SL}_n(\mathbb{R})$. Thus $A \in \Lambda(\mathrm{SL}_n(\mathbb{R}))$, hence $\Lambda(\mathrm{SL}_n(\mathbb{R})) = \mathfrak{sl}_n(\mathbb{R})$.

For each elementary matrix unit $e_{ij} \in \mathrm{Mat}_n(\mathbb{R})$ with $i \neq j$, we have $e_{ij} \in \mathfrak{sl}_n(\mathbb{R})$ and $e_{ij}^2 = 0$. Thus $\exp(te_{ij}) = I + te_{ij}$, an elementary transvection subgroup. By Gaussian elimination,

$$\begin{aligned} \langle \exp(tA) \mid A \in \mathfrak{sl}_n(\mathbb{R}) \rangle &\leq \mathrm{SL}_n(\mathbb{R}) = \langle I + te_{ij} \mid i \neq j, t \in \mathbb{R} \rangle \\ &\leq \langle \exp(tA) \mid A \in \mathfrak{sl}_n(\mathbb{R}) \rangle. \end{aligned}$$

Therefore $\langle \exp(tA) \mid A \in \mathfrak{sl}_n(\mathbb{R}) \rangle = \mathrm{SL}_n(\mathbb{R})$.

If $D = \mathrm{diag}(d_{11}, \dots, d_{ii}, \dots, d_{nn})$ is a diagonal matrix, then $\exp(D)$ is also diagonal with entries $e^{d_{ii}}$. Every diagonal matrix with positive entries on the diagonal can be found this way, and these together with $\mathrm{SL}_n(\mathbb{R})$ generate $\mathrm{GL}_n(\mathbb{R})^+$. By Corollary (3.4), every matrix exponential has positive determinant; so $\langle \exp(tA) \mid A \in \mathfrak{gl}_n(\mathbb{R}) \rangle = \mathrm{GL}_n(\mathbb{R})^+$. $\square$

(3.8). COROLLARY. *Although $\mathrm{GL}_n(\mathbb{R})^+$ has index 2 in $\mathrm{GL}_n(\mathbb{R})$, the two groups have the same tangent space at the identity* cor-diff-gp

$$\Lambda(\mathrm{GL}_n(\mathbb{R})^+) = \Lambda(\mathrm{GL}_n(\mathbb{R})) = \mathfrak{gl}_n(\mathbb{R}). \quad \square$$

Class 10: 25 Sept '15

In the remain results of this subsection, we let G be a closed subgroup of $\mathrm{Mat}_n(\mathbb{R})$ and set $L = \Lambda(G) = T_1(G)$.

(3.9). LEMMA. *If $g \in G$ and $A \in L$, then $gAg^{-1} \in L$.* lem-conj-auto

PROOF. As $g \in G$ and $A \in L$, the group G contains $\exp(tA)$ and

$$\begin{aligned} g(\exp(tA))g^{-1} &= g \left(\sum_{k=0}^{\infty} \frac{t^k A^k}{k!} \right) g^{-1} = \sum_{k=0}^{\infty} g \left(\frac{t^k A^k}{k!} \right) g^{-1} \\ &= \sum_{k=0}^{\infty} \frac{t^k (gA^k g^{-1})}{k!} = \sum_{k=0}^{\infty} \frac{t^k (gAg^{-1})^k}{k!} \\ &= \exp(t(gAg^{-1})). \end{aligned}$$

Therefore $gAg^{-1} \in L$. $\square$

Thus we have the *adjoint representation* of the group G on its Lie algebra L :

$$\mathrm{Ad}: G \longrightarrow \mathrm{GL}_{\mathbb{R}}(L) \quad \text{given by} \quad \mathrm{Ad}_g(A) = gAg^{-1}.$$

It should come as no surprise that in general a Lie group acts on its Lie algebra, the corresponding representation being always called adjoint.

(3.10). LEMMA. *For $A, B \in L$,* lem-Ad-calc

$$\mathrm{Ad}_{\exp(tB)}(A) = A + t(BA - AB) + t^2 D(t)$$

with $D(t) = \sum_{k,l \in \mathbb{N}} d_{kl} B^k A B^l t^{k+l}$ for $d_{kl} \in \mathbb{R}$.

PROOF.

$$\begin{aligned}\mathrm{Ad}_{\exp(tB)}(A) &= (I + tB + t^2B_1(t))A(I - tB + t^2B_2(t)) \\ &= A + t(BA - AB) + t^2D(t). \quad \square\end{aligned}$$

As written, the adjoint representation appears to involve matrix calculation of degree $\dim_{\mathbb{R}}(L)$. On the other hand already $L \leq \mathrm{Mat}_n(\mathbb{R})$; so the next result, among other things, makes the calculation more manageable.

thm-dAd-ad **(3.11).** THEOREM. For $B \in L$, $\mathrm{Ad}_{\exp(B)} = \exp(\mathrm{ad}_B)$.

PROOF. Clearly $t \mapsto \mathrm{Ad}_{\exp(tB)}$ is a one-parameter subgroup of $\mathrm{GL}_{\mathbb{R}}(L)$, so there is an $X \in \mathrm{End}_{\mathbb{R}}^-(L)$ with $\mathrm{Ad}_{\exp(tB)} = \exp(tX)$. By the lemma

$$\mathrm{Ad}_{\exp(tB)}(A) = (I + t\mathrm{ad}_B + t^2E(t))(A),$$

for $E(t) = \sum_{k,l \in \mathbb{N}} d_{kl}B^k(B^l - \mathrm{ad}_{B^l})t^{k+l}$. Thus

$$X = \frac{d}{dt} \mathrm{Ad}_{\exp(tB)}|_{t=0} = \frac{d}{dt} (I + t\mathrm{ad}_B + t^2E(t))|_{t=0} = \mathrm{ad}_B. \quad \square$$

thm-tan-la **(3.12).** THEOREM. L is a Lie subalgebra of $\mathrm{Mat}_n^-(\mathbb{R}) = \mathfrak{gl}_n(\mathbb{R})$.

PROOF. Let $A, B \in L$. By Lemma (3.10), for all $t \in \mathbb{R}$,

$$F(t) = A + t(BA - AB) + t^2D(t)$$

is in the $\mathbb{R}$ -space L . Therefore, for each nonzero $t \in \mathbb{R}$,

$$t^{-1}(F(t) - A) = (BA - AB) + tD(t)$$

is also in L .

The Lie algebra L is a subspace of $\mathrm{Mat}_n(\mathbb{R})$ and especially is closed, hence

$$\lim_{t \rightarrow 0} (BA - AB) + tD(t) = [B, A]$$

is in L . We conclude $[A, B] = -[B, A] \in L$. $\square$

Thus we have the matrix version of Lie's first observation from the beginning of the chapter:

- (i) If G is a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$, then the tangent space to the identity is a Lie algebra $\Lambda(G)$.

In this case we say that $\Lambda(G)$ is the Lie algebra of G .

3.5 Equivalence of representation

In this section we discuss Lie's second basic observation:

- (ii) *The representation theory of the Lie group G and of the Lie algebra $\Lambda(G)$ are essentially the same.*

Even in the case of closed subgroups of $\mathrm{GL}_n(\mathbb{R})$, the results are more difficult than those of the previous subsections. We offer them without proof, but see [CSM95, pp. 75-81] and [Kir08, §3.8] for nice discussions of the general results and their proofs. In the closed group case, each of [Hal15, Ros02, Sti08] proves the first two theorems of this section. Serre's notes [Ser06] contain a proof of Lie's Third Theorem, which makes use of Ado's Theorem (1.6)(b).

Theorem (3.11) could be summarized by the commutative diagram

$$\begin{array}{ccc} \Lambda(G) & \xrightarrow{\mathrm{ad}} & \Lambda(\mathrm{GL}_{\mathbb{K}}(L)) \\ \exp \downarrow & & \downarrow \exp \\ G & \xrightarrow{\mathrm{Ad}} & \mathrm{GL}_{\mathbb{K}}(L) \end{array} \quad .$$

The next theorem provides an important extension of this.

(3.13). THEOREM. *If $f: G \rightarrow H$ is a Lie group homomorphism, then there is a unique Lie algebra homomorphism $df: \Lambda(G) \rightarrow \Lambda(H)$ with $f \exp = \exp df$. That is, we have the following commutative diagram:*

$$\begin{array}{ccc} \Lambda(G) & \xrightarrow{!df} & \Lambda(H) \\ \exp \downarrow & & \downarrow \exp \\ G & \xrightarrow{f} & H \end{array} \quad \square \square$$

This is the easiest theorem of the present section. Especially the candidate for the *differential* df of f is evident, and we give it in the matrix case. If $A \in \Lambda(G)$, then $\varphi(t) = \exp(tA)$ is a one-parameter subgroup of G . If we compose it with f , then $f\varphi(t) = f(\exp(tA))$ is a one-parameter subgroup of H . Therefore there is a $B \in \Lambda(H)$ with $f\varphi(t) = \exp(tB)$. We set $df(A) = B$. This is clearly unique. The remaining verification (in the matrix case) that this gives a Lie algebra homomorphism is achieved through calculations similar to those of the previous two sections; see [Eld15].

A functor F from the category $\mathbf{A}$ to the category $\mathbf{B}$ is an *equivalence* if it is faithful, full, and dense [Jac89]:

- (i) F is *faithful* if the maps $F: \mathrm{Hom}_{\mathbf{A}}(X, Y) \rightarrow \mathrm{Hom}_{\mathbf{B}}(F(X), F(Y))$ are always injections.

- (ii) F is *full* if the maps $F: \text{Hom}_A(X, Y) \longrightarrow \text{Hom}_B(F(X), F(Y))$ are always surjections.
- (iii) F is *dense* if for every object Z of B there is an object X of A with $F(X)$ isomorphic to Z in B .

One should think of category equivalence as saying that the two categories are essentially the same, although the names of the isomorphism classes may have been changed. (For instance, the category of all finite sets is equivalent to the category of all finite subsets of the integers.) In particular, equivalent categories have the same representation theory (subject to some changing of names).

Theorem (3.13) could be restated to say that Λ with $\Lambda(f) = df$ is a faithful functor from the category of Lie groups ${}_{\mathbb{R}}\text{LieGp}$ to the category of Lie algebras ${}_{\mathbb{R}}\text{LieAlg}$. The next two results say that, given appropriate restrictions, Λ is also full and dense. Thus we get a category equivalence.

thm-lie-second **(3.14). THEOREM. (LIE'S SECOND THEOREM)** *If G and H are Lie groups with G simply connected, then for each Lie algebra homomorphism $d: \Lambda(G) \longrightarrow \Lambda(H)$ there is a Lie group homomorphism $f: G \longrightarrow H$ with $d = df$.* $\square\square$

We must restrict to *simply connected* G . This is a stronger requirement than *path connectivity*, which requires that, for every group element, there is a curve containing the identity and that element. Path connectivity makes sense, since our discussion of the tangent space can only reach those elements of G joined to the identity by some curve. Indeed the Lie algebra of any Lie group is equal to that of the connected component of the identity. As we saw in Corollary (3.8), the two groups $\text{GL}_n(\mathbb{R})^+$ and $\text{GL}_n(\mathbb{R})$ have the same Lie algebra $\mathfrak{gl}_n(\mathbb{R})$. That is because any continuous path from the identity I of positive determinant 1 to a matrix of negative determinant would have to pass through a matrix of determinant 0; the path would have to leave the group $\text{GL}_n(\mathbb{R})$.

A simply connected group must be path connected but also satisfy an additional requirement, which we do not give precisely. It states that all paths from the identity to a given element in that component are fundamentally the same. For example, when the group $\text{SL}_2(\mathbb{R})$ acts on $\mathfrak{sl}_2(\mathbb{R})$ via the adjoint representation, the scalar $-I$ acts trivially. This action is that of the quotient group $\text{SL}_2(\mathbb{R})/\{\pm I\} \simeq \text{SO}_{2,1}(\mathbb{R})$, an orthogonal Lie group acting on $\mathbb{R}^3$ preserving the symmetric form whose Gram matrix is diagonal with entries $(1, 1, -1)$. Therefore the nonisomorphic Lie groups $\text{SL}_2(\mathbb{R})$ and $\text{SO}_{2,1}(\mathbb{R})$ have isomorphic Lie algebras $\mathfrak{sl}_2(\mathbb{R}) \simeq \mathfrak{so}_{2,1}(\mathbb{R})$. The difficulty is that in (the path connected group) $\text{SL}_2(\mathbb{R})$ there are paths $c_g^+(t)$ from I to the group element g and paths $c_g^-(t)$ from I to $-g$. In the quotient $\text{SO}_{2,1}(\mathbb{R})$ the images $\bar{c}_g^+(t)$ and $\bar{c}_g^-(t)$ are fundamentally different paths from the identity element $\bar{I} = \{\pm I\}$ to the image $\bar{g} = \{\pm g\}$. This testifies to that fact that the Lie group $\text{SO}_{2,1}(\mathbb{R})$ is not simply connected. The quotient map described above provides a lift for the algebra isomorphism $\mathfrak{sl}_2(\mathbb{R}) \simeq \mathfrak{so}_{2,1}(\mathbb{R})$ to a Lie homomorphism from $\text{SL}_2(\mathbb{R})$ to $\text{SO}_{2,1}(\mathbb{R})$; but there is no corresponding Lie homomorphism from $\text{SO}_{2,1}(\mathbb{R})$ to $\text{SL}_2(\mathbb{R})$.

Woowoo!!

Ben suggests just using $\mathbb{R} \longrightarrow \mathbb{S}^1$. loop groups?

thm-lie-third **(3.15).** THEOREM. (LIE'S THIRD THEOREM) *For all finite dimensional Lie algebras L , there is a Lie group G with $\Lambda(G)$ isomorphic to L .* $\square\square$

As the Lie group G is a manifold, its Lie algebra must be finite dimensional.

(3.16). THEOREM. *The functor Λ gives a category equivalence of the category of simply connected Lie groups $\mathbb{R}\text{LieGp}^{sc}$ and the category of finite dimensional Lie algebras $\mathbb{R}\text{LieAlg}^{fd}$.* thm-cat-equiv

PROOF. We have already observed that Theorem (3.13) says that Λ is faithful. By Lie's Second Theorem (3.14) it is full on $\mathbb{R}\text{LieGp}^{sc}$, and by Lie's Third Theorem (3.14) it is dense to $\mathbb{R}\text{LieAlg}^{fd}$. $\square$

In particular, we now know that the (appropriately restricted) Lie group G and Lie algebra $\Lambda(G)$ have essentially the same representation theory.

3.6 Problems

prob-cbh (3.17). PROBLEM.

- (a) In $\mathrm{GL}_n(\mathbb{R})$ prove that $\exp(tA)\exp(tB) = \exp(t(A+B))$, for all $t \in \mathbb{R}$, if and only if $[A, B] = 0$.

HINT: The function $\exp(t(A+B)) - \exp(tA)\exp(tB)$ is smooth on $\mathbb{R}$.

- (b) Let e_{12} and e_{23} be matrix units in $\mathrm{Mat}_3(\mathbb{R})$. Do the calculations in $\mathrm{SL}_3(\mathbb{R})$ and $\mathfrak{sl}_3(\mathbb{R})$ that exhibit $A+B \in \mathfrak{sl}_3(\mathbb{R})$ but $\exp(A)\exp(B) \neq \exp(A+B)$.

REMARK. For small enough values of t , the smooth curve $\exp(tA)\exp(tB)$ has norm less than 1, so $\log(\exp(tA)\exp(tB))$ exists. Its precise calculation in terms of A and B is the content of the Campbell-Baker-Hausdorff Theorem, which begins

$$\log(\exp(tA)\exp(tB)) = t(A+B) + \frac{1}{2}t^2[A, B] + t^3(\cdots).$$

As such, it also provides a proof of (a). Even at this level it is more sophisticated than what we have done up to now. In particular it involves composing $\log$ and $\exp$ in the order $\log \exp$ as opposed to the simpler $\exp \log$, which we used in our proof of Proposition (3.2).

prob-lie-heisenberg-extraspec (3.18). PROBLEM. Consider the group $X = X_n(\mathbb{R})$ of Problem (2.10). Prove that its Lie algebra is a Heisenberg algebra isomorphic to L_X .

prob-la-deriv (3.19). PROBLEM. Let G be a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$. Prove that if $c : I \rightarrow \Lambda(G)$ is a curve, differentiable on the open interval I , then $c'(t) \in \Lambda(G)$ for all $t \in I$.
HINT: Examine the proof of Theorem (3.12).

Chapter 4

Basics of Lie Algebras

Class 11: 28 Sept '15

The previous chapters were, in a sense, introduction and justification. The actual work starts here. We repeat our basic definition: a *Lie algebra* is a $\mathbb{K}$ -algebra $(\mathbb{K}A, [\cdot, \cdot])$ that satisfies the two identical relations:

- (i) $[x, x] = 0$;
- (ii) (Jacobi Identity) $[[x, y], z] + [[y, z], x] + [[z, x], y] = 0$.

Our overall goals are to classify and understand Lie algebras and their representations under suitable additional hypotheses. We will focus on finite dimensional Lie algebras over algebraically closed fields of characteristic 0, but various parts of what we say are valid in a more general context. In particular, in this chapter we make no restriction on dimension or field, except where expressly noted.

4.1 Basic structure theory

Let L be a Lie $\mathbb{K}$ -algebra. A *subalgebra* of L is a $\mathbb{K}$ -subspace M that is closed under the bracket multiplication. In this case we write $M \leq L$ and $L \geq M$.

A *Lie homomorphism* is a $\mathbb{K}$ -linear transformation $\varphi: L \rightarrow M$ with $\varphi([x, y]) = [\varphi(x), \varphi(y)]$ for all $x, y \in L$. The *kernel* of φ is then the kernel of φ as a linear transformation. In view of the First Isomorphism Theorem below, the image of φ is sometimes referred to as the *quotient algebra* of L by the kernel.

The kernel I is a subalgebra, indeed it is an *ideal* of L —a subspace of L with $[x, a] \in I$ for all $x \in L$ and $a \in I$. We do not need to distinguish right ideals from left ideals, since $[a, x] = -[x, a] \in I$; all right and left ideals are immediately 2-sided ideals.

We have the standard *Isomorphism Theorems*:

(4.1). THEOREM. *Let L be a Lie $\mathbb{K}$ -algebra.*

thm-iso-thms

- (a) (FIRST ISOMORPHISM THEOREM) *If $\varphi: L \rightarrow M$ is a Lie homomorphism with kernel I , then the image algebra $\varphi(L)$ is canonically isomorphic via*

$\varphi(a) \mapsto a + I$ to the quotient algebra L/I provided with the Lie bracket $[a + I, b + I] = [a, b] + I$.

- (b) (SECOND ISOMORPHISM THEOREM) Let I be a subalgebra and J an ideal of L . Then $I+J$ is a subalgebra, $I \cap J$ is an ideal of I and $(I+J)/J \simeq I/(I \cap J)$.
- (c) (THIRD ISOMORPHISM THEOREM) If I is an ideal of L contained in the ideal K , then L/K is isomorphic to $(L/I)/(K/I)$. In particular, there is a bijection between the set of ideals of L/I and the set of ideals of L that contain I . $\square$

By definition, the subspace I is an ideal precisely when it is invariant under all inner derivations ad_x . The ideal I is additionally *characteristic* in L if it is invariant under all derivations of L , not just the inner derivations.

lem-char-ideal **(4.2).** LEMMA. Let J be an ideal of L and I a characteristic ideal of J .

- (a) I is an ideal of L .
- (b) If J is a characteristic ideal of L , then I is a characteristic ideal of L . $\square$

As in the Second Isomorphism Theorem, old ideals can be used to construct new ones.

lem-ideal-sum **(4.3).** LEMMA. Let L be a Lie $\mathbb{K}$ -algebra.

- (a) If A and B are ideals of L , then $A + B$ is an ideal of L .
- (b) If A and B are characteristic ideals of L , then $A + B$ is a characteristic ideal of L . $\square$

For subspaces A and B of L we let the *commutator* $[A, B]$ be the subspace of L spanned by $[a, b]$ for all $a \in A$ and $b \in B$.

lem-ideal-prod **(4.4).** LEMMA. Let L be a Lie $\mathbb{K}$ -algebra.

- (a) If A and B are ideals of L , then $[A, B]$ is an ideal of L .
- (b) If A and B are characteristic ideals of L , then $[A, B]$ is a characteristic ideal of L .

PROOF. For each derivation D , we have $D([a, b]) = [D(a), b] + [a, D(b)]$. $\square$

simple
abelian algebra; largest abelian quotient $L/[L, L]$.

derived series; solvable algebras; length
 Z is kernel of ad
lower central series; nilpotent algebras; class
 $L^1 = L, L^{n+1} = [L^n, L] = [L, L^n]$
 $L^{(0)} = L, L^{(n+1)} = [L^{(n)}, L^{(n)}]$
upper central series?
Example: $\mathfrak{b} = \mathfrak{d} \oplus \mathfrak{n}$

prop-series **(4.5).** PROPOSITION. *Let L be a Lie $\mathbb{K}$ -algebra.*

- (a) $[L^m, L^n] \leq L^{m+n}$.
- (b) $L^{(m)} \leq L^{2^m}$.
- (c) *If L is nilpotent, then L is solvable.*

PROOF. Part (a) follows from induction on n , with $n = 1$ given by the definition of L^{m+1} and the induction step coming from the Jacobi Identity:

$$\begin{aligned}
 [L^m, L^{n+1}] &= [L^m, [L^n, L]] \\
 &\leq [L^n, [L, L^m]] + [L, [L^m, L^n]] \\
 &\leq [L^n, L^{m+1}] + [L, L^{m+n}] \\
 &\leq L^{n+m+1} + L^{1+m+n} \\
 &\leq L^{m+n+1}.
 \end{aligned}$$

Now (b) follows from (a) and (c) follows from (b). □

(4.6). LEMMA. *Let L be a Lie $\mathbb{K}$ -algebra.*

lem-sol-str

- (a) *Subalgebras and quotient algebras of solvable L are solvable.*
- (b) *The sum of solvable ideals in L is a solvable ideal of L*
- (c) *If $\dim_{\mathbb{K}}(L)$ is finite, then L has a unique maximal solvable ideal.*
- (d) *If the ideal I and the quotient L/I are solvable, then L is solvable.* □

(4.7). LEMMA. *Let L be a Lie $\mathbb{K}$ -algebra.*

lem-nil-str

- (a) *Subalgebras and quotient algebras of nilpotent L are solvable.*
- (b) *The sum of nilpotent ideals in L is a nilpotent ideal of L*
- (c) *If $\dim_{\mathbb{K}}(L)$ is finite, then L has a unique maximal nilpotent ideal.* □

It is noteworthy that the last part of the previous lemma does not have a counterpart here; the extension of a nilpotent Lie algebra by a nilpotent Lie algebra need not be nilpotent. (Otherwise, all solvable Lie algebras would also be nilpotent.) In Proposition (4.12) below we will introduce an additional necessary and sufficient condition for such extensions to be nilpotent.

If L has a unique maximal nilpotent ideal, then it is the *nilpotent radical* of L . Similarly if L has a unique maximal solvable ideal, then it is the *radical* or *solvable radical* of L . The nilpotent radical is of course contained in the solvable radical. On the other hand, the last term in the derived series of a solvable ideal is an abelian ideal and so is nilpotent. Therefore the solvable radical is 0 if and only if the nilpotent radical is 0.

A Lie algebra is *semisimple* if its (solvable) radical is 0. By Lemma (4.6)(d) the quotient of L by its radical is then always semisimple.

Woowoo!!

$$\mathfrak{b} = \mathfrak{d} \oplus \mathfrak{n}$$

4.2 Basic representation theory

Class 13: 2 October '15

repn; degree; dimension
 modules; extrinsic and intrinsic; iso and equiv
 repn on trivial Lie algebra ??

“universal” algebra:
 submodule; quotient;
 simple; irreducible; trivial; indecomposable
 cr; semisimple module
 composition series
 Jordan-Hölder
 module duals

adjoint module and consequence for Lie structure

thm-ks (4.8). THEOREM. (KRULL-SCHMIDT THEOREM) *If $V = \bigoplus_{i \in I} V_i = \bigoplus_{j \in J} V_j$ are decompositions of the L -module V into indecomposable summands, then there is a bijection $\sigma: I \rightarrow J$ with V_i and $V^{\sigma(i)}$ isomorphic for all $i \in I$.* $\square$

thm-comp-red (4.9). THEOREM. *Let V be a module for the Lie algebra L . Then the following are equivalent:*

- (1) *for every submodule W of V , there is a submodule W' with $V = W \oplus W'$;*
- (2) *V is a sum of irreducible submodules;*
- (3) *V is a direct sum of irreducible submodules.* $\square$

thm-cr-ks (4.10). THEOREM. *For the finite dimensional and completely reducible module V , let $\mathcal{I}$ be a set of representatives for the isomorphism classes of irreducible submodules of V and let V_i be the sum of all irreducible submodules isomorphic to $i \in \mathcal{I}$. Then $V = \bigoplus_{i \in \mathcal{I}} V_i$.* $\square$

thm-schur (4.11). THEOREM. (SCHUR'S LEMMA) *Let V be a finite dimensional, irreducible L -module over the algebraically closed field $\mathbb{K}$. Then the scalars are the only endomorphisms of V that commute with the action of L .* $\square$

4.3 Further structure and representation

In general an extension of a nilpotent algebra by a nilpotent algebra need not be nilpotent. We do get a nilpotent algebra if we have an additional *Engel condition*, requiring the vanishing of an appropriate iterated commutator. Define $[A; B, n]$ by $[A; B, 1] = [A, B]$ and $[A; B, k+1] = [[A; B, k], B]$.

prop-engel-condition (4.12). PROPOSITION. *Let the Lie algebra L contain an ideal I such that I and L/I are nilpotent. Further assume L has a subalgebra M such that $L = I + M$. Then L is nilpotent if and only if there is a positive m with $[I; M, m] = 0$.*

PROOF. See [Ste70, Lemma 2.1].

If L is nilpotent, then letting m be the class of L gives the required condition.

Now we consider the converse. We first claim that for all positive n and r

$$[I^n; L, r] \leq I^{n+1} + [I^n; M, r].$$

We prove this by induction on r , the result being clear for $r = 1$ as $L = I + M$.

Assume the result for r . Then

$$\begin{aligned} [I^n; L, r+1] &= [[I^n; L, r], L] \\ &\leq [I^{n+1} + [I^n; M, r], I + M] \\ &\leq [I^{n+1}, I + M] + [[I^n; M, r], I] + [[I^n; M, r], M]. \end{aligned}$$

The first two summands are in I^{n+1} (as I^{n+1} and I^n are ideals of L) and the last is equal to $[I^n; M, r+1]$. This gives the claim.

Let k be the maximum of m and the nilpotence class of L/I . We prove $L^{kn} \leq I^n$ by induction on n , with the case $n = 1$ valid by the definition of k . By definition, induction, the claim, and hypothesis

$$L^{kn+k} = [L^{kn}; L, k] \leq [I^n; L, k] \leq I^{n+1} + [I^n; M, k] \leq I^{n+1} + [I; M, k] = I^{n+1},$$

as desired.

For large enough n , nilpotent I has $I^n = 0$. Thus $L^{kn} = 0$, and L is nilpotent. $\square$

The next lemma describes the elementary *internal semidirect product* for Lie algebras. The corresponding *external semidirect product* or *split extension* of Lie algebras is then the construction of the proposition that follows.

(4.13). LEMMA. *Let Lie algebra $L = M \oplus I$ where M is a subalgebra and I is an ideal. Then for $m, n \in M$ and $i, j \in I$ we have* lem-la-int-sdp

$$[m + i, n + j] = [m, n] + [i, j] + [m, j] + [i, n],$$

where $[m, n] \in M$ and $[i, j] + [m, j] + [i, n] = [i, j] + [m, j] - [n, i] \in I$. $\square$

(4.14). PROPOSITION. *Let M and I be Lie $\mathbb{K}$ -algebras, and let $\delta: M \rightarrow \text{Der}_{\mathbb{K}}(I)$ be a Lie homomorphism of M into the derivation algebra of I given by $m \mapsto \delta_m$. Then $M \oplus I$ with bracket multiplication* prop-la-ext-sdp

$$[(m, i), (n, j)] = ([m, n], [i, j] + \delta_m(j) - \delta_n(i))$$

is a Lie $\mathbb{K}$ -algebra in which $0 \oplus I$ is an ideal isomorphic to I and $M \oplus 0$ is a subalgebra isomorphic to M . Furthermore, for each $m \in M$, $\text{ad}_{(m,0)}$ induces $(0, \delta_m)$ on $0 \oplus I$.

PROOF. **Exercise:** the only difficulty is the verification of the Jacobi Identity. In doing that, the corresponding calculation from the lemma can be used as a guide. $\square$

We emphasize two cases.

(4.15). EXAMPLE.

eg-sdp

- (a) *If δ is a derivation of the Lie algebra A , then with $M = \mathbb{K}\delta$ and $I = A$ we make $L = \mathbb{K}\delta \oplus A$ into a Lie algebra as in the proposition. Here A is an ideal of codimension 1 upon which the derivation δ is now induced by the inner derivation ad_δ of the new algebra L .*
- (b) *Let V be a module for the Lie algebra M . As in the proposition $L = M \oplus V$ becomes a Lie algebra after we declare $V(= I)$ to be an abelian Lie algebra: $[V, V] = 0$. (Any endomorphism of an abelian Lie algebra is a derivation by Proposition (2.3).)*

The second example suggests some notation. Let $\varphi: L \longrightarrow \text{End}_{\mathbb{K}}^-(V)$ be a Lie representation of L . For $x \in L$, we may write ad_x^V for $\varphi(x)$. In particular ad_x^L is the usual adjoint action ad_x of x on L in the adjoint representation.

prop-leib-calc **(4.16).** PROPOSITION. *Let δ be a derivation of L . For $x, y \in L$ and $a, b \in \mathbb{K}$:*

$$(\delta - a1 - b1)^n[x, y] = \sum_{i=0}^n \binom{n}{i} [(\delta - a1)^{n-i}(x), (\delta - b1)^i(y)].$$

PROOF. We prove this by induction on n with the case $n = 0$ being trivial

and the case $n = 1$ following from the definition of a derivation.

$$\begin{aligned}
(\delta - a1 - b1)^n[x, y] &= (\delta - a1 - b1)((\delta - a1 - b1)^{n-1}[x, y]) \\
&= (\delta - a1 - b1) \sum_{i=0}^{n-1} \binom{n-1}{i} [(\delta - a1)^{n-1-i}(x), (\delta - b1)^i(y)] \\
&= \sum_{i=0}^{n-1} \binom{n-1}{i} \delta [(\delta - a1)^{n-1-i}(x), (\delta - b1)^i(y)] \\
&\quad + (-a1 - b1) \sum_{i=0}^{n-1} \binom{n-1}{i} [(\delta - a1)^{n-1-i}(x), (\delta - b1)^i(y)] \\
&= \sum_{i=0}^{n-1} \binom{n-1}{i} [\delta(\delta - a1)^{n-1-i}(x), (\delta - b1)^i(y)] \\
&\quad + \sum_{i=0}^{n-1} \binom{n-1}{i} [(\delta - a1)^{n-1-i}(x), \delta(\delta - b1)^i(y)] \\
&\quad + \sum_{i=0}^{n-1} \binom{n-1}{i} [-a(\delta - a1)^{n-1-i}(x), (\delta - b1)^i(y)] \\
&\quad + \sum_{i=0}^{n-1} \binom{n-1}{i} [(\delta - a1)^{n-1-i}(x), -b(\delta - b1)^i(y)] \\
&= \sum_{i=0}^{n-1} \binom{n-1}{i} [(\delta - a1)^{n-i}(x), (\delta - b1)^i(y)] \\
&\quad + \sum_{i=0}^{n-1} \binom{n-1}{i} [(\delta - a1)^{n-1-i}(x), (\delta - b1)^{i+1}(y)] \\
&= \sum_{i=0}^{n-1} \binom{n-1}{i} [(\delta - a1)^{n-i}(x), (\delta - b1)^i(y)] \\
&\quad + \sum_{j=1}^n \binom{n-1}{j-1} [(\delta - a1)^{n-j}(x), (\delta - b1)^j(y)] \\
&= \sum_{k=0}^n \binom{n}{k} [(\delta - a1)^{n-k}(x), (\delta - b1)^k(y)]. \quad \square
\end{aligned}$$

4.4 Problems

- prob-nilpt-ex (4.17). PROBLEM. For $L = \mathfrak{n}$ calculate L^k and $L^{(k)}$.
- prob-exp-der (4.18). PROBLEM. Field indep $\exp(\delta)$ auto; see [Ros02, p. 51].
- prob-nilp-der-aut (4.19). PROBLEM. nilpotent derivations and automorphisms
- prob-jord-chev (4.20). PROBLEM. Jordan-Chevalley decomposition
- prob-la-cent-extn-bis (4.21). PROBLEM. Lie algebra central extension.

Chapter 5

Nilpotent representations

Example: $\mathfrak{b} = \mathfrak{d} \oplus \mathfrak{n}$

Class 14: 5 October '15

Woowoo!!

5.1 Engel's Theorem and Cartan subalgebras

(5.1). PROPOSITION. *Let N be a nilpotent Lie algebra and V a $\mathbb{K}$ -module. For each element x of N and for each $\lambda \in \mathbb{K}$, the generalized eigenspace* prop-submodules

$$V_{x,\lambda} = \{ v \in V \mid (x - \lambda 1)^k v = 0, \text{ some } k = k_{x,\lambda,v} \in \mathbb{N} \}$$

for x on V is an N -submodule of V .

PROOF. Let N have nilpotence class l . For $v \in V_{x,\lambda}$ set $n = l + k_{x,\lambda,v}$. As in Example (4.15)(ii), we calculate within the semidirect product of V by L . By Proposition (4.16) with $y \in L$, $\delta = \text{ad}_x$, $a = 0$, and $b = \lambda$, we have

$$\begin{aligned} (x - \lambda 1)^n(yv) &= (\text{ad}_x - \lambda 1)^n[y, v] \\ &= \sum_{i=0}^n \binom{n}{i} [\text{ad}_x^{n-i}(y), (\text{ad}_x - \lambda 1)^i(v)] = 0, \end{aligned}$$

since $\text{ad}_x^k(y) = 0$ for $k \geq l$ and $(\text{ad}_x - \lambda 1)^i(v) = 0$ for $i \geq k_{x,\lambda,v}$.

This shows that $yV_{x,\lambda} \leq V_{x,\lambda}$, hence the subspace $V_{x,\lambda}$ of V is in fact a submodule. □

An endomorphism is *nil* if some power of it is 0, and a *nil representation* of the Lie algebra N is one in which each element of N acts as a nil endomorphism.

(5.2). PROPOSITION. *If σ is an nil irreducible representation of the nilpotent Lie algebra N , then σ is the trivial 1-dimensional representation.* prop-irred-nil

PROOF. Certainly $L^n V = 0$, where n is the nilpotence class of N . Suppose $L^k V = 0$. If $k = 1$, then $LV = 0$ and irreducible V has dimension 1, as desired.

For $k > 1$ let $x \in L^{k-1}$. As the representation is nil, for nonzero $u \in V$ there is a positive n_u with $x^{n_u}u = 0$. For minimal such n_u , the element $w = x^{n_u-1}u$ is nonzero with $xw = 0$. Thus $W = \{v \in V \mid xv = 0\}$ is nonzero. For all $y \in L$ and $w \in W$

$$x(yw) = y(xw) - [x, y]w = 0 - 0 = 0$$

as $w \in W$ and $[x, y] \in L^k$. Therefore $yw \in W$, which is thus a nonzero submodule. By irreducibility $W = V$, hence $xV = 0$. But this implies $L^{k-1}V = 0$, and we are done. $\square$

cor-nil-comp-series **(5.3). COROLLARY.** *If σ is an finite dimensional nil representation of the nilpotent Lie algebra N , then V has an N -composition series with all factors of dimension 1 and trivial.* $\square$

thm-engel **(5.4). THEOREM.** (ENGEL'S THEOREM) *If the adjoint representation of the finite dimensional Lie algebra N is nil, then N is nilpotent.*

PROOF. We prove this by induction on $\dim_{\mathbb{K}}(N)$ with the result clearly true in dimensions 0 and 1. Assume $N \neq 0$.

Let I be a maximal proper subalgebra of N . As $\text{ad}_x^I = \text{ad}_x^N|_I$ for $x \in I$, the adjoint representation of I is nil. Therefore by induction I is nilpotent.

By Corollary (5.3) there is a 1-dimensional submodule P/I for the nil action of nilpotent I on N/I . Let $x \in P \setminus I$ and $M = \mathbb{K}x$. Then

$$\begin{aligned} [P, P] &= [M + I, M + I] \\ &= [M, M] + [M, I] + [I, M] + [I, I] \\ &= [M, I] + [I, M] + [I, I] \\ &\leq I, \end{aligned}$$

so P is a subalgebra of N in which I is an ideal of codimension 1. By maximality of I , $N = P$.

We now have $N = M \oplus I$ with $M = \mathbb{K}x \simeq N/I$ an abelian algebra and I a nilpotent ideal. Furthermore by hypothesis $\text{ad}_x^m = 0$ for some k , hence $[I; M, m] = [I; x, m] = 0$. By Proposition (4.12), the algebra N is nilpotent. $\square$

Class 15: 7 October '15

If A is a subspace of the Lie algebra L , then the *normalizer* of A in L , $N_L(A)$, is $\{x \in L \mid [x, A] \leq A\}$. The subspace A is then *self-normalizing* if $A = N_L(A)$.

lem-norm-subspace **(5.5). LEMMA.**

- (a) *If A is a subspace of the Lie algebra L , then $N_L(A)$ is a subalgebra.*
- (b) *If A is a self-normalizing subspace of the Lie algebra L , then A is a subalgebra.*

PROOF. For $x, y \in N_L(A)$ and $a \in A$, the Jacobi Identity gives

$$[[x, y], a] = -[[y, a], x] - [[a, x], y] \in A,$$

so the vector space $N_L(A)$ is a subalgebra. The second part then follows from the first. $\square$

lem-10-subalg (5.6). LEMMA. Let L be a Lie algebra, $x \in L$, and

$$L_{x,0} = \{ y \mid \text{ad}_x^k(y) = 0, \text{ some } k = k_{x,0,y} \in \mathbb{N} \}$$

be the generalized eigenspace for x acting on L in the adjoint representation with eigenvalue 0. Then $L_{x,0}$ is a self-normalizing subalgebra of L .

PROOF. Let $a \in N_L(L_{x,0})$. Then $[x, a] \in L_{x,0}$, so $\text{ad}_x^k([x, a]) = 0$ for $k = k_{x,0,[x,a]}$. But then $\text{ad}_x^{k+1}(a) = \text{ad}_x^k([x, a]) = 0$, hence $a \in L_{x,0}$ and the subspace $L_{x,0}$ is self-normalizing. By the previous lemma it is then a subalgebra. $\square$

The element w of the finite dimensional Lie algebra is said to be *regular* in L if the dimension of the subalgebra $L_{w,0}$ is minimal. This dimension is then the *rank* of L . As long as $L \neq 0$ this is positive since $w \in L_{w,0}$.

(5.7). THEOREM. Assume $\mathbb{K}$ has characteristic 0. Let w be a regular element of the finite dimensional Lie algebra L and set $H = L_{w,0}$. Then H is a nilpotent and self-normalizing subalgebra of L . thm-cart-subalg

PROOF. We follow [Eld15].

By the previous lemma, H is a self-normalizing subalgebra. We must prove it to be nilpotent.

For fixed but arbitrary $h \in H$ and $\alpha \in \mathbb{K}$, the element $w + \alpha h$ belongs to H . Consider the linear transformation $\text{ad}_{w+\alpha h}$ of L , which leaves the subspace H invariant and so also acts on the quotient space L/H . Therefore its characteristic polynomial $\chi_\alpha^L(z) \in \mathbb{K}[z]$ is $\varphi_\alpha(z)\gamma_\alpha(z)$ where

$$\varphi_\alpha(z) = z^r + \sum_{i=0}^{r-1} f_i(\alpha)z^i$$

is the characteristic polynomial of $\text{ad}_{w+\alpha h}$ on H and

$$\gamma_\alpha(z) = z^{n-r} + \sum_{j=0}^{n-r-1} g_j(\alpha)z^j$$

is the characteristic polynomial of $\text{ad}_{w+\alpha h}$ on L/H . The standard calculation of the characteristic polynomial as a determinant reveals the polynomials $f_i(x)$ of $\mathbb{K}[x]$ to have degree at most r while the $g_j(x)$ have degree at most $n - r$.

As $H = L_{w,0}$ we have $\gamma_0(0) \neq 0$ hence $g_0(0) \neq 0$. Especially the polynomial $g_0(x)$ of degree at most $n - r$ is not identically 0. As $\mathbb{K}$ has characteristic 0 we have $|\mathbb{K}| > n$, so there are distinct elements $\alpha_1, \dots, \alpha_{r+1}$ of $\mathbb{K}$ with $g_0(\alpha_k) \neq 0$ for $1 \leq k \leq r+1$. In particular $L_{w+\alpha_k h,0} \leq H$ for each k . As w is regular, this forces $L_{w+\alpha_k h,0} = H$, which is to say $\varphi_{\alpha_k}(z) = z^r$ for $1 \leq k \leq r+1$. But then each of the polynomials $f_i(x)$, for $1 \leq i < r$, vanishes at $\alpha_1, \dots, \alpha_{r+1}$. As these polynomials have degree at most r , they must be identically 0.

Therefore $\varphi_\alpha(z) = z^r$ for all values of $\alpha \in \mathbb{K}$, and every $w + \alpha h$ is nil on H . As h was fixed but arbitrary, we find that every element of H is nil on H . By Engel's Theorem (5.4), H is nilpotent as desired. $\square$

Class 16: 9 October '15

A *Cartan subalgebra* of the Lie algebra L is a nilpotent, self-normalizing subalgebra. The theorem tells us that Cartan subalgebras always exist in finite dimension and characteristic 0. More is true: for finite dimensional Lie algebras over algebraically closed fields of characteristic 0, the automorphism group of L is transitive on the Cartan subalgebras (so all arise as in the theorem); see [Jac79, p. 273]. We shall only use this in abusing notation and terminology, where we often will not mention the specific Cartan subalgebra being used since they are all essentially equivalent.

There are many characterizations of Cartan subalgebras. The following is important here.

prop-cart-char **(5.8). PROPOSITION.** *Suppose H is a nilpotent subalgebra of the finite dimensional Lie algebra L . Then H is a Cartan subalgebra if and only if in the action of H on L via the adjoint, H is equal to*

$$L_{H,0} = \{ x \in L \mid \text{ad}_h^k(x) = 0 \text{ for all } h \in H \text{ and some } k = k_{h,0,x} \in \mathbb{N} \},$$

the largest subspace of L upon which H is nil.

PROOF. The nilpotent algebra H is certainly contained in $L_{H,0}$. We show that H is proper in $L_{H,0}$ if and only if H is not self-normalizing. As the Cartan subalgebras are by definition the self-normalizing nilpotent subalgebras, this will give the result.

Let $x \in N_L(H) \setminus H$. Then, for each $h \in H$ we have $[h, x] \in H$. As H is nilpotent, $\text{ad}_h^k[h, x]$ is 0 for sufficiently large $k = k_h$. But then $\text{ad}_h^{k+1}(x) = 0$ and x is in $L_{H,0}$ but not in H .

Suppose $L_{H,0} > H$. By Corollary (5.3), there is a trivial H -submodule P/H of dimension 1 in $L_0(H)/H$. For $x \in P \setminus H$, we have $[x, H] \leq H$. That is, x is in the normalizer of H but not in H . $\square$

5.2 Weight spaces and Cartan decomposition

thm-indec-nilpt **(5.9). THEOREM.** *Assume $\mathbb{K}$ is algebraically closed of characteristic 0. Let V be an indecomposable $\mathbb{K}N$ -module for the nilpotent Lie algebra N with $0 < n = \dim_{\mathbb{K}}(V)$. Then there is a 1-dimensional Lie homomorphism $\lambda: N \rightarrow \mathbb{K}$ with*

$$V = \{ v \in V \mid (x - \lambda(x)1)^{n-1}v = 0 \text{ for all } x \in N \}.$$

PROOF. We may replace N with its image in $\text{End}_{\mathbb{K}}^-(V) \simeq \text{Mat}_n^-(\mathbb{K})$. As $\mathbb{K}$ is algebraically closed, all $x \in N$ have eigenvalues in their action on V . By standard linear algebra (say, Jordan Canonical Form), for each $x \in N$ the module V is the direct sum of its generalized eigenspaces

$$V_{x,\lambda} = \{ v \in V \mid (x - \lambda 1)^k v = 0, \text{ some } k = k_{x,\lambda,v} \in \mathbb{N} \}.$$

Indeed $\max_v(k_{x,\lambda,v}) \leq n-1$, so

$$V_{x,\lambda} = \{v \in V \mid (x - \lambda 1)^{n-1}v = 0\}.$$

By Proposition (5.1), indecomposability, and the above remarks, each $x \in N$ has a unique eigenvalue $\lambda(x)$ on V , and for every x the whole space V is equal to the generalized x -eigenspace $V_{x,\lambda(x)}$:

$$V = V_{x,\lambda(x)} = \{v \in V \mid (x - \lambda(x)1)^{n-1}v = 0\}.$$

In particular $\text{tr}(x) = n\lambda(x)$. As $\mathbb{K}$ has characteristic 0, we find that $\lambda(x) = n^{-1}\text{tr}(x)$ is a linear map $\lambda: N \rightarrow \mathbb{K}$. Furthermore

$$\lambda([x, y]) = n^{-1}\text{tr}(xy - yx) = 0,$$

for all $x, y \in N$; that is, $\lambda|_{[N, N]} = 0$. Therefore the linear transformation $\lambda: N \rightarrow \mathbb{K}$ is a 1-dimensional representation of the abelian Lie algebra $N/[N, N]$ and so of N itself. $\square$

A 1-dimensional representation of a Lie algebra L is called a *weight* of the algebra. For an L -module V and weight λ of L ,

$$V_{L,\lambda} = V_\lambda = \{v \in V \mid (x - \lambda(x)1)^k v = 0 \text{ for all } x \in L \text{ and some } k = k_{x,\lambda,v} \in \mathbb{N}\}$$

is the corresponding *weight space* in V .

For every nonzero Lie algebra, the trivial representation is the *trivial weight* or *zero weight*. We have already encountered a weight space in Proposition (5.8), where the Cartan subalgebra H was characterized among all nilpotent subalgebras of L by being equal to its corresponding weight space $L_{H,0}$. Note that all weights belong to the dual of the $\mathbb{K}$ -space $H/[H, H]$.

A nonzero weight of L is a *root*.

(5.10). THEOREM. *Assume $\mathbb{K}$ is algebraically closed of characteristic 0. For the nilpotent Lie algebra N and the N -module V of finite dimension n , N has only finitely many weights on V ; each weight space* thm-weight-spaces

$$V_{N,\lambda} = V_\lambda = \{v \in V \mid (x - \lambda(x)1)^{n-1}v = 0 \text{ for all } x \in N\}$$

is a submodule; and V is the direct sum of its weight spaces.

PROOF. As V is finite dimensional, we can write V as a direct sum of finitely many nonzero indecomposable submodules. By the previous theorem, each of these summands is contained in one of the weight spaces V_μ for some weight μ of N . Let the submodule $V(\mu)$ be the sum of those indecomposable summands with weight μ . The previous theorem gives

$$V(\mu) \leq \{v \in V \mid (x - \mu(x)1)^{n-1}v = 0 \text{ for all } x \in N\} \leq V_\mu.$$

Now we have

$$V = \bigoplus_{\mu \in J} V(\mu),$$

where J is a finite set of weights for N on V . In particular, every $v \in V$ can be uniquely written $v = \sum_{\mu \in M} v_\mu$ with $v_\mu \in V(\mu)$.

Let λ be an arbitrary weight of N on V , and consider $0 \neq v \in V_\lambda$. We claim:

$$v_\mu \neq 0 \implies \mu = \lambda.$$

As the various nonzero v_μ are linearly independent and each $V(\mu)$ is a submodule, $(x - \lambda(x)1)^m v = 0$ implies $(x - \lambda(x)1)^m v_\mu = 0$ and so $v_\mu \in V_\lambda \cap V(\mu) \leq V_\lambda \cap V_\mu$.

Assume $v_\mu \neq 0$. For fixed but arbitrary $x \in N$, choose $k (= k_{x, \lambda(x), v}) \in \mathbb{N}$ minimal with $(x - \lambda(x)1)^k v_\mu = 0$. Set $u = (x - \lambda(x)1)^{k-1} v_\mu \neq 0$, so that $(x - \lambda(x))u = 0$; that is, $xu = \lambda(x)u$. As $V(\mu)$ is a submodule, $u \in V(\mu) \leq V_\mu$; so there is an $m \in \mathbb{Z}^+$ with $(x - \mu(x)1)^m u = 0$. But

$$(x - \mu(x)1)u = xu - \mu(x)u = \lambda(x)u - \mu(x)u = (\lambda(x) - \mu(x))u,$$

hence

$$0 = (x - \mu(x)1)^m u = (\lambda(x) - \mu(x))^m u.$$

Now $u \neq 0$ forces $\lambda(x) - \mu(x) = 0$. That is, for all $x \in N$ we have $\lambda(x) = \mu(x)$, hence $\lambda = \mu$ as claimed.

For every weight λ , each nonzero $v \in V_\lambda$ must project nontrivially onto at least one of the summands $V(\mu)$ for $\mu \in M$. By the claim, there is only one such summand, namely $V(\lambda)$, and $v \in V(\lambda)$. Thus $\lambda \in J$ and there are only finitely many weights for N on V . Also $V_\lambda \leq V(\lambda) \leq V_\lambda$, hence

$$V(\lambda) = \{v \in V \mid (x - \mu(x)1)^{n-1} v = 0 \text{ for all } x \in N\} = V_\lambda.$$

Finally V is the direct sum of the submodules $V(\mu)$, so it is equally well the direct sum of the weight spaces V_λ , each a submodule. $\square$

Class 17: 12 October '15

Now let L be a finite dimensional Lie algebra over the algebraically closed field $\mathbb{K}$ of characteristic 0. By Theorem (5.7) there is a Cartan subalgebra H of L . By Theorem (5.10) we have

$$L = \bigoplus_{\lambda} L_{H, \lambda} = \bigoplus_{\lambda} L_{\lambda},$$

where λ runs over the finite set of weights of H on L . This is a *Cartan decomposition* of the Lie algebra L —the decomposition of L as the direct sum of its weight spaces for a Cartan subalgebra H .

Here and above we see the common abuse of notation and terminology that refers to the weights and weight spaces of L without specifying the Cartan subalgebra H being used, say, writing L_λ in place of $L_{H, \lambda}$. Usually H will be clear from the context, and in the cases of most interest to us all Cartan algebras are equivalent; see the remarks on page 50.

thm-root-bracket

(5.11). THEOREM. *Let L be a finite dimensional Lie algebra over the algebraically closed field $\mathbb{K}$ of characteristic 0. Let α and β be weights of L for the Cartan subalgebra $H = L_0$. Then*

$$[L_\alpha, L_\beta] \leq L_{\alpha+\beta},$$

where the weight space L_λ for $\lambda \in (H/[H, H])^*$ is taken to be 0 when λ is not a weight.

PROOF. Let $x \in H$, $y \in L_\alpha$, and $z \in L_\beta$. Then, for $n = 2 \dim_{\mathbb{K}}(H)$, by Proposition (4.16) and Theorem (5.10)

$$(\text{ad}_x - \alpha 1 - \beta 1)^n [y, z] = \sum_{i=0}^n \binom{n}{i} [(\text{ad}_x - \alpha 1)^{n-i}(y), (\text{ad}_x - \beta 1)^i(z)] = 0.$$

Therefore, if $[y, z] \neq 0$ then $[y, z] \in L_{\alpha+\beta}$. $\square$

Recall that a root of L is a nonzero weight. For α and β roots of L , the α -string through β is the longest string of roots

$$\beta - s\alpha, \dots, \beta - i\alpha, \dots, \beta, \dots, \beta + j\alpha, \dots, \beta + t\alpha.$$

That is, all the maps in the string are roots, but $\beta - (s+1)\alpha$ and $\beta + (t+1)\alpha$ are not roots.

We have a first application of this concept.

(5.12). PROPOSITION. *Let L be a finite dimensional Lie algebra over the algebraically closed field $\mathbb{K}$ of characteristic 0, and let α and β be roots of L . Then β is a rational multiple of α when restricted to the subspace $[L_\alpha, L_{-\alpha}]$.* prop-13.2

PROOF. This is [Ste70, Lemma 3.2].

The result is trivial if $-\alpha$ is not a root, so we may assume it is. Let

$$\beta - s\alpha, \dots, \beta, \dots, \beta + t\alpha$$

be the α -string through β and M the corresponding subspace

$$M = L_{\beta-s\alpha} \oplus \dots \oplus L_\beta \oplus \dots \oplus L_{\beta+t\alpha}.$$

By Theorem (5.11) we have $[M, L_{-\alpha}] \leq M$ and $[M, L_\alpha] \leq M$.

Let $y \in L_\alpha$ and $z \in L_{-\alpha}$, and set $x = [y, z]$. As y and z normalize M , so does x . We have

$$\text{tr}(\text{ad}_x|_M) = \sum_{i=-s}^t d_i(\beta + i\alpha)(x),$$

where $d_i = \dim_{\mathbb{K}}(L_{\beta+i\alpha})$. But $\text{ad}_x = [\text{ad}_y, \text{ad}_z]$ and so it has trace 0. Therefore

$$0 = \sum_{i=-s}^t d_i(\beta + i\alpha)(x),$$

hence

$$\beta(x) = \frac{d}{e}\alpha(x)$$

for $d = -\sum_{i=-s}^t id_i$. and $e = \sum_{i=-s}^t d_i \neq 0$. By linearity, this holds for all x in $[L_\alpha, L_{-\alpha}]$. $\square$

5.3 Problems

prop-self-norm **(5.13).** PROBLEM. *Prove that any subalgebra of the Lie algebra L that contains $L_{x,0}$ is self-normalizing.*

Chapter 6

Killing forms and semisimple Lie algebras

6.1 Killing forms

Let L be a finite dimensional Lie $\mathbb{K}$ -algebra and V an L -module. The *Killing form* of L on V , $\kappa_L^V: L \times L \rightarrow \mathbb{K}$ is given by

$$\kappa_L^V(x, y) = \text{tr}(\text{ad}_x^V \text{ad}_y^V),$$

where we recall our convention that ad_x^V is the image of $x \in L$ in $\text{End}_{\mathbb{K}}(V)$.

If the relevant Lie algebra L should be evident from the context, then we write κ^V . Finally, if $V = L$, the representation being the adjoint, we may drop reference to V as well, since we then have the usual definition of the Killing form

$$\kappa(x, y) = \text{tr}(\text{ad}_x \text{ad}_y).$$

(6.1). PROPOSITION.

prop-prop-kil

(a) The Killing form κ_L^V is a symmetric, bilinear form on L .

(b) If W is an L -submodule of V , then

$$\kappa_L^V = \kappa_L^W + \kappa_L^{V/W}.$$

Class 18: 14 October '15

(c) The Killing form is an invariant form (or associative form): for all $x, y, z \in L$

$$\kappa_L^V([x, y], z) = \kappa_L^V(x, [y, z]).$$

(d) If I is an ideal of L , then $I^\perp = \{x \in L \mid \kappa_L^V(x, y) = 0, \text{ for all } y \in I\}$ is also an ideal of L .

PROOF.

- (a) The trace is linear in its argument with target $\mathbb{K}$, and multiplication in $\text{End}_{\mathbb{K}}(V)$ is bilinear; so κ_L^V is a bilinear form on L . It is symmetric since $\text{tr}(ab) = \text{tr}(ba)$ in $\text{End}_{\mathbb{K}}(V)$.
- (b) This is clear.
- (c) This is evident if we write the module action in matrix form, using a basis that extends a basis of W to one for all V .
- (d)

$$\begin{aligned}
 \kappa_L^V([x, y], z) &= \text{tr}(\text{ad}_{[x, y]}^V \text{ad}_z^V) \\
 &= \text{tr}((\text{ad}_x^V \text{ad}_y^V - \text{ad}_y^V \text{ad}_x^V) \text{ad}_z^V) \\
 &= \text{tr}(\text{ad}_x^V \text{ad}_y^V \text{ad}_z^V - \text{ad}_y^V \text{ad}_x^V \text{ad}_z^V) \\
 &= \text{tr}(\text{ad}_x^V \text{ad}_y^V \text{ad}_z^V) - \text{tr}(\text{ad}_y^V \text{ad}_x^V \text{ad}_z^V) \\
 &= \text{tr}(\text{ad}_x^V \text{ad}_y^V \text{ad}_z^V) - \text{tr}(\text{ad}_x^V \text{ad}_z^V \text{ad}_y^V) \\
 &= \text{tr}(\text{ad}_x^V (\text{ad}_y^V \text{ad}_z^V - \text{ad}_z^V \text{ad}_y^V)) \\
 &= \text{tr}(\text{ad}_x^V \text{ad}_{[y, z]}^V) \\
 &= \kappa_L^V(x, [y, z]).
 \end{aligned}$$

- (e) This is clear.
- (f) For all $a \in I$, $y \in L$, and $b \in I^\perp$ we have by (c)

$$0 = \kappa_L^V([a, y], b) = \kappa_L^V(a, [y, b]).$$

That is, $[y, b] \in I^\perp$ for all $y \in L$ and $b \in I^\perp$; so $I^\perp$ is an ideal. $\square$

(6.2). COROLLARY. *Let I be an ideal of the finite dimensional Lie algebra L . Then $I \leq \text{Rad}(\kappa_L^{L/I})$ and $\kappa_I^I = \kappa_L^I|_{I \times I} = \kappa_L^L|_{I \times I}$*

PROOF. From the second part of the proposition

$$\kappa_L^L = \kappa_L^I + \kappa_L^{L/I}.$$

As I acts as 0 on L/I , we certainly have $I \leq \text{Rad}(\kappa_L^{L/I})$. The rest of the corollary follows easily. $\square$

Some care must be taken in the use of this result. In $\mathfrak{gl}_n(\mathbb{K})$ the Borel algebra $\mathfrak{b}_n(\mathbb{K}) = \mathfrak{n}_n^+(\mathbb{K}) \oplus \mathfrak{h}_n(\mathbb{K})$ is the split extension of its derived subalgebra $\mathfrak{n}_n^+(\mathbb{K}) = [\mathfrak{b}_n(\mathbb{K}), \mathfrak{b}_n(\mathbb{K})]$ by the Cartan subalgebra $\mathfrak{h}_n(\mathbb{K})$. Let $L = \mathfrak{b}_n(\mathbb{K})$ and $I = \mathfrak{n}_n^+(\mathbb{K})$. Then nilpotent I consists of strictly upper triangular matrices, so κ_I^I is identically 0; $L/I \simeq \mathfrak{h}_n(\mathbb{K})$ is abelian and so $\kappa_{L/I}^{L/I}$ is identically 0. Nevertheless $\kappa_L^L = \kappa_L^I + \kappa_L^{L/I}$ is not identically 0 on solvable $\mathfrak{b}_n(\mathbb{K})$ provided $n \geq 2$.

Wowwo!!

thm-perfect-kil **(6.3). THEOREM.** *Let $L (\neq 0)$ be a finite dimensional Lie algebra over the field $\mathbb{K}$ of characteristic 0. If $L = [L, L]$, then the Killing form κ is not identically 0.*

PROOF. For any extension field $\mathbb{E}$ of $\mathbb{K}$, if κ^L is identically 0, then so is $\kappa^{\mathbb{E} \otimes_{\mathbb{K}} L}$. Therefore in proving the theorem we may assume that $\mathbb{K}$ is algebraically closed.

Let $L = \bigoplus_{\lambda \in \Phi_0} L_\lambda$ be the Cartan decomposition for L relative to the Cartan subgroup $H = L_0$ and finite set of weights Φ_0 . Thus

$$L = [L, L] = \left[\bigoplus_{\lambda \in \Phi_0} L_\lambda, \bigoplus_{\lambda \in \Phi_0} L_\lambda \right] = \bigoplus_{\lambda \neq \mu} [L_\lambda, L_\mu]$$

In particular

$$H = \bigoplus_{\lambda \in \Phi_0} [L_\lambda, L_{-\lambda}].$$

As nonzero nilpotent $H > [H, H]$ and $L = [L, L]$, we have $H < L$; so the set of roots $\Phi = \Phi_0 \setminus \{0\}$ is nonempty.

Let $\beta \in \Phi$. By the definition of roots, $\beta|_H \neq 0$ but $\beta|_{[H, H]} = 0$. Therefore there is an $\alpha \in \Phi_0$ with $\beta|_{[L_\alpha, L_{-\alpha}]} \neq 0$. Furthermore α is not the zero weight as again $\beta|_{[H, H]} = 0$. Thus by Proposition (5.12) there is a rational number $r_{\beta, \alpha}$ with

$$\beta|_{[L_\alpha, L_{-\alpha}]} = r_{\beta, \alpha} \alpha|_{[L_\alpha, L_{-\alpha}]}.$$

Choose an $x \in [L_\alpha, L_{-\alpha}]$ with $\beta(x) \neq 0$, hence $r_{\beta, \alpha} \neq 0$ and $\alpha(x) \neq 0$. Then

$$\begin{aligned} \kappa(x, x) &= \text{tr}(\text{ad}_x \text{ad}_x) \\ &= \sum_{\lambda \in \Phi_0} \lambda(x)^2 \dim_{\mathbb{K}}(L_\lambda) \\ &= 0 + \sum_{\lambda \in \Phi} \lambda(x)^2 \dim_{\mathbb{K}}(L_\lambda) \\ &= \alpha(x)^2 \sum_{\lambda \in \Phi} r_{\lambda, \alpha}^2 \dim_{\mathbb{K}}(L_\lambda), \end{aligned}$$

which is not equal to 0, as all $r_{\lambda, \alpha}$ are nonzero and all $\dim_{\mathbb{K}}(L_\lambda)$ are positive integers. Since $\kappa(x, x) \neq 0$, the form κ is not identically 0 on L , as desired. $\square$

(6.4). COROLLARY. (CARTAN'S SOLVABILITY CRITERION) *Let L be a finite dimensional Lie algebra over the field $\mathbb{K}$ of characteristic 0. If the Killing form is identically 0, then L is solvable.* cor-cart-solv

PROOF. Assume the Killing form κ is identically 0. The proof is by induction on $\dim_{\mathbb{K}}(L)$, with the dimension 0 and 1 cases clear. By the Theorem $L \neq [L, L]$. By Corollary (6.2) the Killing form for $[L, L]$ comes from restriction of the Killing form for L and so is also identically 0. Therefore by induction $[L, L]$ is solvable, and then L is as well by Lemma (4.6). $\square$

A slightly more complicated condition on κ is both necessary and sufficient for solvability; see [Eld15]: L is solvable if and only if $\kappa|_{L \times [L, L]}$ is identically 0.

A case in point is that of the Borel algebras $\mathfrak{b}_n(\mathbb{K})$, mentioned above, which, are solvable with a nonzero Killing form whose restriction to the derived subalgebra $[\mathfrak{b}_n(\mathbb{K}), \mathfrak{b}_n(\mathbb{K})]$ is identically 0.

Class 19: 16 October '15

We then have the natural result that lives at the opposite end of the solvability and degeneracy spectrum.

(6.5). THEOREM. (CARTAN'S SEMISIMPLICITY CRITERION) *Let L be a finite dimensional Lie algebra over the field $\mathbb{K}$ of characteristic 0. Then L is semisimple if and only if its Killing form is nondegenerate.*

PROOF. Let κ be the Killing form and $R = \text{Rad}(\kappa)$, an ideal by Proposition (6.1). But $\kappa|_{R \times R} = \kappa_R^R$ is identically 0, so R is solvable by Cartan's Solvability Criterion (6.4). If L is semisimple, then $R = 0$ and κ is nondegenerate.

Now let S be a nonzero solvable ideal of L , and take I to be the last nonzero term in its derived series. Therefore abelian I is in $\text{Rad}(\kappa_L^I)$, and also $I \leq \text{Rad}(\kappa_L^{L/I})$ by Corollary (6.2). Hence by Proposition (6.1)

$$0 \neq I \leq \text{Rad}(\kappa_L^I) \cap \text{Rad}(\kappa_L^{L/I}) = \text{Rad}(\kappa_L^I + \kappa_L^{L/I}) = \text{Rad}(\kappa)$$

and κ is degenerate. $\square$

We also have a result which resolves a possible confusion involving terminology.

(6.6). THEOREM. *Let L be a finite dimensional Lie algebra over the field $\mathbb{K}$ of characteristic 0. Then L is semisimple if and only if, as L -module, it is completely reducible with no trivial 1-dimensional ideals.*

In this case all minimal ideals (irreducible submodules) are nontrivial simple subalgebras.

PROOF. Let κ be the Killing form on L , and let I be an ideal in semisimple L . Then $I \cap I^\perp$ is an ideal by Proposition (6.1), and the restriction of κ to it is identically 0. Therefore by Cartan's Solvability Criterion (6.4) the ideal $I \cap I^\perp$ is solvable and hence 0 in semisimple L . Therefore finite dimensional $L = I \oplus I^\perp$, and every ideal I is complemented in L . By Theorem (4.9), L is completely reducible as L -module. In particular, minimal ideals and irreducible submodules are the same and are simple. If any of these were trivial simple submodules, they would be solvable ideals, which is not the case.

Conversely, assume that L is completely reducible with the decomposition $L = \bigoplus_{i=0}^m S_i$ into simple ideals with no summand trivial. Any solvable ideal I projects onto each summand S_i as a solvable subideal. Since no summand is trivial, each of these projections is onto the zero ideal; so I itself is zero. Therefore L is semisimple. $\square$

6.2 Semisimple algebras I

The classification of finite dimensional, semisimple Lie algebras over algebraically closed fields of characteristic 0 has two basic parts:

- (i) the reduction of the classification to that of root systems;
- (ii) the use of the classification of root systems to complete the Lie algebra classification.

In this section we do most of the first part.

We first set some notation to be used throughout this section. In particular $L (\neq 0)$ will be a finite dimensional, semisimple Lie algebra over the algebraically closed field $\mathbb{K}$ of characteristic 0.

By Theorem (5.7) we may choose a Cartan subalgebra H in L . By Proposition (5.8) we have $H = L_{H,0} = L_0$, the zero weight space. Let Φ be the set of all roots for H on L , a finite set by Theorem (5.10). The set of all weights is $\Phi_0 = \{0\} \cup \Phi$.

For each $\lambda \in \Phi$, we have the weight space $L_\lambda = L_{H,\lambda}$, giving the Cartan decomposition

$$L = H \oplus \bigoplus_{\lambda \in \Phi} L_\lambda.$$

Since L is nonzero and semisimple, the nilpotent Cartan subalgebra $H = L_0$ is proper in L , hence the root set Φ is nonempty.

The Killing form $\kappa = \kappa^L = \kappa_L = \kappa_L^L$ is nondegenerate by Cartan's Semisimplicity Criterion (6.5).

(6.7). PROPOSITION. *Let α and β be weights.*

prop-ss-prop

- (a) $\kappa(L_\alpha, L_\beta) = 0$ if $\alpha + \beta \neq 0$.
- (b) $\kappa|_{H \times H}$ is nondegenerate, and $H^\perp = \bigoplus_{\lambda \in \Phi} L_\lambda$.
- (c) If $0 \neq x \in L_\alpha$, then $\kappa(x, L_{-\alpha}) \neq 0$. Especially, $\alpha \in \Phi$ implies $-\alpha \in \Phi$.

PROOF. (a) Recall that for all weights μ, ν we have $[L_\mu, L_\nu] \leq L_{\mu+\nu}$ by Theorem (5.11), and this extends to all $\lambda, \mu \in (H/[H, H])^*$ when we define L_λ to be 0 whenever λ is not a root.

For $x \in L_\alpha$, $y \in L_\beta$, and $\gamma \in \Phi_0$,

$$\text{ad}_x \text{ad}_y L_\gamma = [x, [y, L_\gamma]] \leq [L_\alpha, [L_\beta, L_\gamma]] \leq [L_\alpha, L_{\beta+\gamma}] \leq L_{\alpha+\beta+\gamma}.$$

Therefore $\text{tr}(\text{ad}_x \text{ad}_y)$ is 0 if $\alpha + \beta$ is not equal to 0.

(b) By (a) $H^\perp = L_0^\perp \geq \bigoplus_{\lambda \in \Phi} L_\lambda$. Therefore $H^\perp \cap H \leq \text{Rad}(\kappa) = 0$, and $\kappa|_{H \times H}$ is nondegenerate. Then $L = H \oplus \bigoplus_{\lambda \in \Phi} L_\lambda$ yields $H^\perp = \bigoplus_{\lambda \in \Phi} L_\lambda$.

(c) If $\kappa(x, L_{-\alpha}) = 0$, then by (a) we have $x \in \text{Rad}(\kappa) = 0$. Therefore if α is a root, then $\kappa(L_\alpha, L_{-\alpha}) \neq 0$, hence $-\alpha$ is also a root. $\square$

(6.8). THEOREM. *The Cartan subalgebra H is abelian.*

thm-csa-ab

Class 20: 19 October '15

PROOF. Let $x, y \in H$. Then

$$\kappa(x, y) = \text{tr}(\text{ad}_x \text{ad}_y) = \sum_{\lambda \in \Phi} \lambda(x) \lambda(y) \dim L_\lambda.$$

If $w \in [H, H]$, then $\lambda(w) = 0$ for all $\lambda \in \Phi$; so $\kappa(w, y) = 0$ for all $w \in [H, H]$ and $y \in H$. That is, $[H, H] \leq H \cap H^\perp = 0$ by Proposition (6.7)(b). Therefore H is abelian. $\square$

As abelian $H = H/[H, H]$ is finite dimensional and nondegenerate under κ , for every linear functional $\mu \in H^*$ there is a unique $t_\mu \in H$ with $\kappa(t_\mu, h) = \mu(h)$ for all $h \in H$. Especially $t_{-\mu} = -t_\mu$.

Class 21: 21 October '15

prop-t-prop **(6.9). PROPOSITION.**

- (a) $H = \sum_{\lambda \in \Phi} \mathbb{K}t_\lambda$.
- (b) For each $\alpha \in \Phi$ we have $\alpha(t_\alpha) = \kappa(t_\alpha, t_\alpha) \neq 0$.
- (c) For $\alpha \in \Phi$, $x \in L_\alpha$, and $y \in L_{-\alpha}$ we have $[x, y] = \kappa(x, y)t_\alpha$. Especially $[L_\alpha, L_{-\alpha}] = \mathbb{K}t_\alpha$.
- (d) For each $\alpha \in \Phi$ we have $\dim_{\mathbb{K}}(L_\alpha) = 1$ and $k\alpha \notin \Phi$ for $k \notin \{-1, 1\}$.

PROOF. (a) Let $J = \sum_{\lambda \in \Phi} \mathbb{K}t_\lambda \leq H$ and choose $h \in J^\perp \cap H$. Then $\lambda(h) = \kappa(t_\lambda, h) = 0$ for all $\lambda \in \Phi$ and indeed for all $\lambda \in \Phi_0$ since $H = L_0$. Thus for a basis of L consisting of bases for the various L_λ (ordered appropriately using Theorem (5.10)) every ad_x , for $x \in H$, is represented by a matrix that is upper triangular and ad_h itself is strictly upper triangular. But then $\text{ad}_h \text{ad}_x$ is always strictly upper triangular, hence $h \in \text{Rad}(\kappa) = 0$. Therefore $J^\perp \cap H = 0$ with $J \leq H$, so $J = H$ because L has finite dimension.

(b) By nondegeneracy of κ on H and (a), there is a root β with $0 \neq \kappa(t_\beta, t_\alpha) = \beta(t_\alpha)$. Then Proposition (5.12) yields

$$0 \neq \beta(t_\alpha) = r_{\beta, \alpha} \alpha(t_\alpha) = r_{\beta, \alpha} \kappa(t_\alpha, t_\alpha),$$

and $\kappa(t_\alpha, t_\alpha) \neq 0$.

(c) First choose x so that $\mathbb{K}x$ is a 1-dimensional H -submodule of L_α , possible by Theorem (5.10). For all $h \in H$ and $y \in L_{-\alpha}$

$$\begin{aligned} \kappa(h, [x, y]) &= \kappa([h, x], y) = \alpha(h) \kappa(x, y) \\ &= \kappa(t_\alpha, h) \kappa(x, y) = \kappa(h, \kappa(x, y)t_\alpha). \end{aligned}$$

We thus have $\kappa(h, [x, y] - \kappa(x, y)t_\alpha) = 0$ for all h . By the nondegeneracy of κ on H , this gives $[x, y] = \kappa(x, y)t_\alpha \leq \mathbb{K}t_\alpha$. By Proposition (6.7)(c) we may choose $y \in L_{-\alpha}$ with $\kappa(x, y) \neq 0$, hence $[x, L_{-\alpha}] = \mathbb{K}t_\alpha$. Once we have proven (d) we will have $\mathbb{K}x = L_\alpha$, and this part will be done as well.

(d) Let x and y be as in the previous paragraph, and consider the subspace

$$M = \mathbb{K}x \oplus H \oplus \bigoplus_{i \in \mathbb{Z}^+} L_{-i\alpha}$$

where $\mathbb{K}x \leq L_\alpha$ and $H = L_0$. By Theorem (5.11) the space M is an $L_{-\alpha}$ -submodule of L and especially is y -invariant. Similarly it is a $\mathbb{K}x$ -submodule as

$$[x, x] = 0, [x, H] \leq \mathbb{K}x, [x, L_{-i\alpha}] \leq L_{-(i-1)\alpha}.$$

Therefore M is a submodule for the subalgebra generated by $\mathbb{K}x \oplus L_{-\alpha}$, which contains x , y and $[x, y] = \kappa(x, y)t_\alpha \neq 0$. In particular

$$\begin{aligned} 0 &= \kappa(x, y)^{-1} \operatorname{tr}(\operatorname{ad}_{[x, y]}|_M) = \operatorname{tr}(\operatorname{ad}_{t_\alpha}|_M) \\ &= \alpha(t_\alpha) + 0 + \sum_{i \in \mathbb{Z}^+} -i\alpha(t_\alpha) \dim_{\mathbb{K}}(L_{-i\alpha}). \end{aligned}$$

By (b) the constant $\alpha(t_\alpha)$ is nonzero, so it can be canceled to give

$$1 = \sum_{i \in \mathbb{Z}^+} i \dim_{\mathbb{K}}(L_{-i\alpha}) = \dim_{\mathbb{K}}(L_{-\alpha}) + \sum_{2 \leq i \in \mathbb{Z}^+} i \dim_{\mathbb{K}}(L_{-i\alpha}).$$

Since all the summands are nonnegative integers and $\dim_{\mathbb{K}}(L_{-\alpha}) \neq 0$, we conclude that $\dim_{\mathbb{K}}(L_{-\alpha}) = 1$ and $\dim_{\mathbb{K}}(L_{-i\alpha}) = 0$ for all integers i at least two. As α runs through Φ , so does $-\alpha$. This completes the result. $\square$

Define $h_\alpha = \frac{2}{\kappa(t_\alpha, t_\alpha)} t_\alpha$, possible by Proposition (6.9)(c).

(6.10). THEOREM. *For each $\alpha \in \Phi$*

thm-sltwo-cover

$$\mathbb{K}t_\alpha \oplus L_\alpha \oplus L_{-\alpha} = \mathbb{K}h_\alpha \oplus L_\alpha \oplus L_{-\alpha}$$

is a subalgebra isomorphic to $\mathfrak{sl}_2(\mathbb{K})$.

PROOF. For any nonzero $x \in L_\alpha$, the previous proposition allows us to choose a $y \in L_{-\alpha}$ with $\mathbb{K}x = L_\alpha$, $\mathbb{K}y = L_{-\alpha}$, and $\kappa(x, y) = \frac{2}{\kappa(t_\alpha, t_\alpha)}$. Then

$$\begin{aligned} [x, y] &= \kappa(x, y)t_\alpha = \frac{2}{\kappa(t_\alpha, t_\alpha)} t_\alpha = h_\alpha; \\ [h_\alpha, x] &= \frac{2}{\kappa(t_\alpha, t_\alpha)} [t_\alpha, x] \\ &= \frac{2}{\kappa(t_\alpha, t_\alpha)} \alpha(t_\alpha)x = \frac{2}{\kappa(t_\alpha, t_\alpha)} \kappa(t_\alpha, t_\alpha)x = 2x; \\ [h_\alpha, y] &= \frac{2}{\kappa(t_\alpha, t_\alpha)} [t_\alpha, y] = \frac{-2}{\kappa(t_\alpha, t_\alpha)} [t_{-\alpha}, y] = \frac{-2}{\kappa(t_\alpha, t_\alpha)} (-\alpha)(t_{-\alpha})y \\ &= \frac{-2}{\kappa(t_\alpha, t_\alpha)} \alpha(t_\alpha)y = \frac{-2}{\kappa(t_\alpha, t_\alpha)} \kappa(t_\alpha, t_\alpha)y = -2y. \end{aligned}$$

Therefore

$$\mathbb{K}x \oplus \mathbb{K}y \oplus \mathbb{K}t_\alpha = \mathbb{K}x \oplus \mathbb{K}y \oplus \mathbb{K}h_\alpha$$

is a subalgebra, and by Section 2.3 it is a copy of $\mathfrak{sl}_2(\mathbb{K})$. $\square$

(6.11). COROLLARY. *For each $\alpha \in \Phi$ there is a subalgebra S_α in L that is isomorphic to $\mathfrak{sl}_2(\mathbb{K})$. Furthermore $L = \sum_{\alpha \in \Phi} S_\alpha$.*

cor-sltwo-cover

PROOF. By the theorem, for each $\alpha \in \Phi$ there is an S_α with $\mathbb{K}t_\alpha + L_\alpha \leq S_\alpha$. Thus by Proposition (6.9)(a) and the Cartan Decomposition

$$L \geq \sum_{\alpha \in \Phi} S_\alpha \geq \sum_{\alpha \in \Phi} \mathbb{K}t_\alpha + \sum_{\alpha \in \Phi} L_\alpha = H \oplus \bigoplus_{\alpha \in \Phi} L_\alpha = L. \quad \square$$

6.3 Problems

prob-cart-solv **(6.12).** PROBLEM. *Let L be a finite dimensional Lie algebra in characteristic 0. Prove that L is solvable if $\kappa|_{L \times [L, L]}$ is identically 0.*

Chapter 7

Representations of $\mathfrak{sl}_2(\mathbb{K})$

We have seen at the end of the last chapter that a finite dimensional semisimple Lie algebra L over an algebraically closed field $\mathbb{K}$ of characteristic 0 is sewed together out of copies of $\mathfrak{sl}_2(\mathbb{K})$. We could have proceeded with the program outlined at the beginning of Section 6.2 toward a refined description of the Cartan decomposition, leading to the introduction of root systems. But the relevant calculations essentially come from the structure of L as an $\mathfrak{sl}_2(\mathbb{K})$ -module. Therefore in this chapter we take some time off to describe the representation theory of $\mathfrak{sl}_2(\mathbb{K})$ in a manner more detailed than actually needed for the program. (For the semisimple classification, we only need the much easier Theorem (7.22).)

In fact, $\mathfrak{sl}_2(\mathbb{K})$ is the only semisimple Lie algebra whose irreducible representations have been completely cataloged, but we do not do it in its entirety. For that, one should consult the excellent book [Maz10], which is the motivation for much in this chapter.

The irreducible, finite dimensional $\mathfrak{sl}_2(\mathbb{K})$ -modules can be described quickly (and we have already seen them on page 19), but we shall also pursue certain (possibly) infinite dimensional modules. For the module properties considered, we will be guided by the desire to include all finite dimensional modules; so their properties will motivate our definitions.

Throughout this chapter, we let $\mathbb{K}$ be an algebraically closed field of characteristic 0.

7.1 Weight modules

Within the Lie algebra $\mathfrak{sl}_2(\mathbb{K})$, we have focused on three elements

$$h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

which provide the presentation

$$[e, f] = h, \quad [h, e] = 2e, \quad [h, f] = -2f,$$

for the 3-dimensional algebra $\mathfrak{sl}_2(\mathbb{K}) = \mathbb{K}h \oplus \mathbb{K}e \oplus \mathbb{K}f$. The following lemma tells us these 1-spaces can be characterized extrinsically (as a Cartan subalgebra and the nilpotent radicals of the two Borel subalgebras containing it), so this presentation is in a sense canonical.

lem-ss-sltwo **(7.1). LEMMA.** *The algebra $\mathfrak{sl}_2(\mathbb{K})$ is simple. Every Cartan subalgebra of $\mathfrak{sl}_2(\mathbb{K})$ is equivalent under $\text{Aut}(\mathfrak{sl}_2(\mathbb{K}))$ to $\mathbb{K}h$, and the only subalgebras containing $\mathbb{K}h$ are*

$$\mathfrak{sl}_2(\mathbb{K}), \quad \mathbb{K}h, \quad B^+ = \mathbb{K}h \oplus \mathbb{K}e, \quad B^- = \mathbb{K}h \oplus \mathbb{K}f.$$

PROOF.¹ With respect to the basis $\{h, e, f\}$ the Gram matrix of the Killing form κ is

$$\begin{pmatrix} 8 & 0 & 0 \\ 0 & 0 & 4 \\ 0 & 4 & 0 \end{pmatrix}$$

so κ is nondegenerate as $\text{char } \mathbb{K} \neq 2$. Therefore by Cartan's Semisimplicity Criterion (6.5) the algebra $\mathfrak{sl}_2(\mathbb{K})$ of dimension 3 is semisimple. Were it not simple, it would have an abelian ideal of dimension 1, which is not the case.

A Cartan subalgebra is thus abelian (Theorem (6.8)) and so is contained in the normalizer of the subalgebra generated by each of its elements. By Jordan Canonical Form, every nonzero element of $\mathfrak{sl}_2(\mathbb{K})$ is conjugate under $\text{GL}_2(\mathbb{K}) \leq \text{Aut}(\mathfrak{sl}_2(\mathbb{K}))$ to one of

$$\begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} b & 0 \\ 0 & -b \end{pmatrix}$$

for nonzero $a, b \in \mathbb{K}$. The normalizer of the subalgebra with the first shape contains elements of the second type and is not nilpotent. The abelian subalgebra $\mathbb{K}h$ of all matrices of the second type is self-normalizing and so is a Cartan subalgebra by Proposition (5.8).

The rest of the Lemma follows by Problem (2.7) or by calculation with respect to the Cartan decomposition given by $\mathbb{K}h$. $\square$

Class 22: 23 October '15

The Cartan subalgebra $\mathbb{K}h$ of $\mathfrak{sl}_2(\mathbb{K})$ has dimension 1, so the elements λ of its dual space are described entirely by the element $\lambda(h) \in \mathbb{K}$ via $\lambda(kh) = k\lambda(h)$ for all $k \in \mathbb{K}$. We therefore abuse notation and terminology somewhat by setting $\lambda(h) = \lambda \in \mathbb{K}$ and saying that the element λ of $\mathbb{K}$ is a weight when, more properly, it is the associated linear functional $\lambda: \mathbb{K}h \rightarrow \mathbb{K}$ given by $kh \mapsto \lambda k$ that is the weight.

Let V be an $\mathfrak{sl}_2(\mathbb{K})$ -module. We let the images of h, e , and f in $\text{End}_{\mathbb{K}}^-(V)$ be, respectively, H^V, E_+^V, E_-^V , usually abbreviated to H, E_+, E_- . (In the notation

¹Exercise: This proof is overkill. Do the calculations needed to make it elementary—that is, free of reference to results like Proposition (5.8) and Theorems (6.5) and (6.8).

introduced on page 44, these are ad_h^V , ad_e^V , and ad_f^V .)

In the following, we write V_λ^w for those $x \in L_\lambda$ with $[h, x] = \lambda(h)x$. Such a nonzero x is called a *weight vector* for the weight λ . The space of weight vectors V_λ^w is the actual $\mathbf{H}$ -eigenspace, a subspace of the generalized $\mathbf{H}$ -eigenspace V_λ . If finite dimensional $V_\lambda \neq 0$ then $V_\lambda^w \neq 0$ by Theorem (5.10).

The $\mathfrak{sl}_2(\mathbb{K})$ -module V is *torsion-free* if, for all nonzero $f(x) \in \mathbb{K}[x]$ and all nonzero $v \in V$, we have $f(\mathbf{H})v \neq 0$.

Torsion-free modules and weight vectors provide us with a basic dichotomy.

(7.2). PROPOSITION. *If V is an irreducible $\mathfrak{sl}_2(\mathbb{K})$ -module, then either V is torsion-free or V contains a weight vector.* prop-dichotomy

PROOF. Assume V has torsion, so that there are $0 \neq v \in V$ and $0 \neq f(x) \in \mathbb{K}[x]$ with $f(\mathbf{H})v = 0$. Choose $f(x)$ to be monic and of minimal degree subject to this. Thus $f(x) = \prod_{i=1}^n (x - \alpha_i)$ for distinct $\alpha_i \in \mathbb{K}$. With $w = \prod_{i=1}^{n-1} (\mathbf{H} - \alpha_i)v$, we then have w a weight vector for $\lambda = \alpha_n$. $\square$

Our goal in this chapter is to classify all irreducible modules containing a weight vector. As $v, \mathbf{H}v, \mathbf{H}^2v, \dots, \mathbf{H}^i v, \dots$ must be linearly dependent in finite dimensional V , this classification will include the classification of all irreducible finite dimensional $\mathfrak{sl}_2(\mathbb{K})$ -modules.

Class 23: 26 October '15

(7.3). LEMMA. *Let V be an $\mathfrak{sl}_2(\mathbb{K})$ -module.*

lem-hef-action

- (a) $\mathbf{E}_\epsilon \mathbf{E}_{-\epsilon} - \mathbf{E}_{-\epsilon} \mathbf{E}_\epsilon = \epsilon \mathbf{H}$; $\mathbf{H} \mathbf{E}_\epsilon - \mathbf{E}_\epsilon \mathbf{H} = 2\epsilon \mathbf{E}_\epsilon$.
- (b) $\mathbf{E}_\epsilon \mathbf{E}_{-\epsilon} = \epsilon \mathbf{H} + \mathbf{E}_{-\epsilon} \mathbf{E}_\epsilon$; $\mathbf{H} \mathbf{E}_\epsilon = \mathbf{E}_\epsilon (\mathbf{H} + 2\epsilon)$; $\mathbf{E}_\epsilon \mathbf{H} = (\mathbf{H} - 2\epsilon) \mathbf{E}_\epsilon$.
- (c) If $f(x) \in \mathbb{K}[x]$ is a polynomial and $n \in \mathbb{N}$, then

$$f(\mathbf{H}) \mathbf{E}_\epsilon^n = \mathbf{E}_\epsilon^n f(\mathbf{H} + 2\epsilon n) \quad \text{and} \quad \mathbf{E}_\epsilon^n f(\mathbf{H}) = f(\mathbf{H} - 2\epsilon n) \mathbf{E}_\epsilon^n.$$

- (d) If $v \in V_\lambda$ then $\mathbf{E}_\epsilon v \in V_{\lambda+2\epsilon}$.
- (e) If $v \in V_\lambda^w$ then $\mathbf{E}_\epsilon v \in V_{\lambda+2\epsilon}^w$ and $\mathbf{E}_{-\epsilon} \mathbf{E}_\epsilon v \in V_\lambda^w$.

PROOF. The first part consists of the equations demonstrating that V is an $\mathfrak{sl}_2(\mathbb{K})$ -module. The second part is then a rewritten version of the first, and the third part follows by induction. (Exercise.)

Let $v \in V_\lambda$. Then by the previous part

$$(\mathbf{H} - (\lambda + 2\epsilon))^k (\mathbf{E}_\epsilon v) = ((\mathbf{H} - (\lambda + 2\epsilon))^k \mathbf{E}_\epsilon) v = (\mathbf{E}_\epsilon (\mathbf{H} - \lambda)^k) v = \mathbf{E}_\epsilon ((\mathbf{H} - \lambda)^k v),$$

which is 0 for large enough k , hence $\mathbf{E}_\epsilon v \in V_{\lambda+2\epsilon}$. Further let $v \in V_\lambda^w$. Then the equalities holds and are 0 for $k = 1$, proving that $\mathbf{E}_\epsilon v \in V_{\lambda+2\epsilon}^w$, as claimed at the beginning of the last part. The end follows directly. $\square$

A *weight module* for $\mathfrak{sl}_2(\mathbb{K})$ is a module that is generated by weight vectors, which for an irreducible module is equivalent to containing a weight vector. By the previous result, the sum of all the spaces V_λ^w of weight vectors is a submodule

of V , so V is a weight module if and only if $V_\lambda^w = V_\lambda$ for all λ . This in turn is equivalent to the statement that $H^V = H$ is “diagonal” or “semisimple” in its action on V . These remarks include the usual definitions of a weight module; see [Maz10, p. 59].

If v is a λ -weight vector, then always

$$\lambda v = H v = (E_+ E_- - E_- E_+) v = E_+ E_- v - E_- E_+ v$$

where by the lemma $E_+ E_- v$ and $E_- E_+ v$ both belong to V_λ^w . We say that the weight vector v is *coherent* if there are constants λ_+ and λ_- in $\mathbb{K}$ with

$$E_+ E_- v = \lambda_+ v \quad \text{and} \quad E_- E_+ v = \lambda_- v,$$

where necessarily $\lambda_+ - \lambda_- = \lambda$.

lem-coh-dim-1 **(7.4).** LEMMA. *If $\dim_{\mathbb{K}}(V_\lambda^w) = 1$, then each weight vector $v \in V_\lambda^w$ is coherent.* $\square$

We shall see (in Proposition (7.24)) that weight vectors in irreducible $\mathfrak{sl}_2(\mathbb{K})$ -modules are always coherent. Irreducible modules are always cyclic, so our classification results will come from careful study of cyclic modules generated by coherent weight vectors.

Two special types of coherent weight vectors are of particular note. In the $\mathfrak{sl}_2(\mathbb{K})$ -module V , we will call the nonzero vector v a *highest weight vector* for the weight $\lambda \in \mathbb{K}$ provided v is an eigenvector, a weight vector, in V_λ for H but $E_+(v) = 0$ (so $E_- E_+(v) = 0$, $\lambda_- = 0$, and $\lambda_+ = \lambda$). Equivalently, $\mathbb{K}v \leq V_\lambda$ is not just a $\mathbb{K}h$ -submodule but is also a submodule for the Borel subalgebra $B^+ = \mathbb{K}h \oplus \mathbb{K}e$. Similarly nonzero v is a *lowest weight vector* for the weight $\lambda \in K$ provided $\mathbb{K}v \leq V_\lambda$ is a B^- -submodule; that is, v is weight vector of V_λ^w for H but $E_-(v) = 0$ (hence $E_+ E_-(v) = 0$, $\lambda_+ = 0$, and $\lambda_- = \lambda$).

thm-fd-weight-vectors **(7.5).** THEOREM. *Let V be a finite dimensional module for $\mathfrak{sl}_2(\mathbb{K})$. Then V contains highest weight vectors and lowest weight vectors.*

PROOF. Choose a weight λ with $0 \neq v \in V_\lambda^w$. By the lemma and finite dimensionality, there are integers t with $0 \neq E_+^t v \in V_{\lambda+2t}^w$ but $0 = E_+^{t+1} v$, and s with $0 \neq E_-^s v \in V_{\lambda-2s}^w$ but $0 = E_-^{s+1} v$. Thus $v_+ = E_+^t v$ is a highest weight vector and $v_- = E_-^s v$ is a lowest weight vector. $\square$

An elementary arithmetic lemma will be of help.

lem-a-calc **(7.6).** LEMMA. *Let $\lambda, \lambda_+, \lambda_- \in \mathbb{K}$ with $\lambda_+ - \lambda_- = \lambda$. Set $\epsilon \in \{\pm\} = \{\pm 1\}$, and consider the two sequences $a_\epsilon(i)$, for $i \in \mathbb{Z}$, where $a_\epsilon(i) = 1$ for $-\epsilon i \in \mathbb{N}$. The following are equivalent:*

- (1) $a_\epsilon(i) = (i - \epsilon)(\lambda - i) + \lambda_\epsilon = i(\lambda - i + \epsilon) + \lambda_{-\epsilon}$ for all $\epsilon i \in \mathbb{Z}^+$.
- (2) $a_\epsilon(\epsilon) = \lambda_\epsilon$ and $a_\epsilon(i) - a_\epsilon(i - \epsilon) = \epsilon(\lambda - 2(i - \epsilon))$ for all $\epsilon i \in \mathbb{Z}^+$.
- (3) $a_\epsilon(\epsilon) = \lambda_\epsilon$ and $a_-(i)a_+(i+1) - a_+(i)a_-(i-1) = \lambda - 2i$ for all $i \in \mathbb{Z}$.

PROOF. (Exercise.) □

cor-a-calc **(7.7).** COROLLARY. *Let $\lambda = \lambda_+ - \lambda_-$ and $\epsilon \in \{\pm 1\} = \{\pm 1\}$. From the two half-infinite sequences*

$$a_\epsilon(i) = (i - \epsilon)(\lambda - i) + \lambda_\epsilon = i(\lambda - i + \epsilon) + \lambda_{-\epsilon} \text{ for all } \epsilon i \in \mathbb{Z}^+$$

create the new doubly infinite sequence

$$b(j) = a_-(j) \text{ for } j \in \mathbb{Z}^- \text{ and } b(j) = a_+(j+1) \text{ for } j \in \mathbb{N}.$$

Then $b(j) = j(\lambda - j - 1) + \lambda_+$ for all $j \in \mathbb{Z}$. □

We now investigate the structure of a cyclic submodule.

(7.8). THEOREM. *Let v be a weight vector for the weight λ in the $\mathfrak{sl}_2(\mathbb{K})$ -thm-internal-verma module V . Set $v_0 = v$, and let $\epsilon \in \{\pm 1\} = \{\pm 1\}$.*

For all $i \in \mathbb{Z}^+$ define

$$v_{\epsilon i} = E_{-\epsilon}^i(v) = E_{-\epsilon}^i(v_0) = E_{-\epsilon}(v_{\epsilon(i-1)}),$$

which we may rewrite as:

$$E_\epsilon(v_i) = v_{i-\epsilon} \text{ for } \epsilon i \in -\mathbb{N}$$

Then

(a) $H(v_i) = (\lambda - 2i)v_i$ for all $i \in \mathbb{Z}$;

(b) *Assume v is coherent. Set $E_\epsilon E_{-\epsilon}v = \lambda_\epsilon v$, hence $\lambda_+ - \lambda_- = \lambda$. Then*

$$E_\epsilon(v_i) = a_\epsilon(i)v_{i-\epsilon} \text{ for } \epsilon i \in \mathbb{Z}^+$$

where $a_\epsilon(i) = (i - \epsilon)(\lambda - i) + \lambda_\epsilon$, as in Lemma (7.6).

PROOF. (a) We prove this by induction on $|i|$, the case $|i| = 0$ holding by definition. For an i with $|i| > 0$, define $\epsilon \in \{\pm 1\} = \{\pm 1\}$ by $i = \epsilon|i|$. We use Lemma (7.3) and induction to calculate

$$\begin{aligned} H v_i &= H E_{-\epsilon} v_{i-\epsilon} \\ &= E_{-\epsilon} (H - 2\epsilon) v_{i-\epsilon} \\ &= E_{-\epsilon} ((\lambda - 2(i - \epsilon)) - 2\epsilon) v_{i-\epsilon} \\ &= (\lambda - 2i) E_{-\epsilon} v_{i-\epsilon} \\ &= (\lambda - 2i) v_i. \end{aligned}$$

(b) We proceed by induction on the positive integer ϵi . The case $\epsilon i = 1$ serves to define the two constants $a_\epsilon(\epsilon) = \lambda_\epsilon$, where $\lambda_\epsilon - \lambda_{-\epsilon} = \epsilon\lambda$ by coherence

of v . Assume $\epsilon i \geq 2$. Then

$$\begin{aligned}
 E_\epsilon(v_i) &= E_\epsilon E_{-\epsilon} v_{i-\epsilon} \\
 &= (\epsilon H + E_{-\epsilon} E_\epsilon) v_{i-\epsilon} && \text{by Lemma (7.3)} \\
 &= \epsilon H v_{i-\epsilon} + E_{-\epsilon} E_\epsilon v_{i-\epsilon} \\
 &= \epsilon(\lambda - 2(i - \epsilon)) v_{i-\epsilon} + E_{-\epsilon} a_\epsilon(i - \epsilon) v_{i-2\epsilon} && \text{by (a) and induction} \\
 &= \epsilon(\lambda - 2(i - \epsilon)) v_{i-\epsilon} + a_\epsilon(i - \epsilon) v_{i-\epsilon} && \text{as } -\epsilon(i - 2\epsilon) \leq 0 \\
 &= a_\epsilon(i) v_{i-\epsilon} && \text{by Lemma (7.6).}
 \end{aligned}$$

This completes the induction and proof. $\square$

7.2 Verma modules

Class 24: 28 October '15

Theorem (7.8) motivates the following construction:

(7.9). DEFINITION. Let $\lambda, \lambda_+, \lambda_- \in \mathbb{K}$ with $\lambda_+ - \lambda_- = \lambda$. Set $M(\lambda, \lambda_+, \lambda_-) = \bigoplus_{i \in \mathbb{Z}} \mathbb{K} v_i$. Define the linear transformations H , E_+ , and E_- on $M(\lambda, \lambda_+, \lambda_-)$ by

$$\begin{aligned}
 H(v_i) &= (\lambda - 2i)v_i, \\
 E_\epsilon(v_i) &= a_\epsilon(i)v_{i-\epsilon},
 \end{aligned}$$

for $i \in \mathbb{Z}$ and $\epsilon \in \{\pm 1\} = \{\pm 1\}$ with

$$\begin{aligned}
 a_\epsilon(i) &= (i - \epsilon)(\lambda - i) + \lambda_\epsilon && \text{for } \epsilon i \in \mathbb{Z}^+ \text{ and} \\
 a_\epsilon(i) &= 1 && \text{for } \epsilon i \in -\mathbb{N}.
 \end{aligned}$$

In this action and with respect to the basis $\{v_i \mid i \in \mathbb{Z}\}$, the linear transformation H is “diagonal,” in the sense that it takes each 1-space $\mathbb{K}v_i$ to itself. Similarly E_+ is “lower diagonal,” always taking $\mathbb{K}v_i$ to $\mathbb{K}v_{i-1}$, and E_- is “upper diagonal,” taking $\mathbb{K}v_i$ to $\mathbb{K}v_{i+1}$. The corresponding nonzero coefficients, the *transformation coefficients*, are

$$H: v_i \xrightarrow{\lambda-2i} v_i \quad E_+: v_i \xrightarrow{a_+(i)} v_{i-1} \quad E_-: v_i \xrightarrow{a_-(i)} v_{i+1}$$

We exhibit the actions on $M(\lambda, \lambda_+, \lambda_-)$ pictorially as below. In the picture every basis vector appears twice—once on the line displaying the action of E_- and once on the line displaying the action of E_+ . Passage between the two lines gives the action of H .

$$\begin{array}{ccccccccccc}
 E_- & \xrightarrow{\quad} & v_{-2} & \xrightarrow{a_-(-2)} & v_{-1} & \xrightarrow{\lambda_-} & v_0 & \xrightarrow{1} & v_1 & \xrightarrow{1} & v_2 & \xrightarrow{\quad} \\
 & & \downarrow \lambda+4 & & \downarrow \lambda+2 & & \downarrow \lambda & & \downarrow \lambda-2 & & \downarrow \lambda-4 & \\
 E_+ & \xleftarrow{\quad} & v_{-2} & \xleftarrow{1} & v_{-1} & \xleftarrow{1} & v_0 & \xleftarrow{\lambda_+} & v_1 & \xleftarrow{a_+(2)} & v_2 & \xleftarrow{\quad}
 \end{array}$$

thm-verma **(7.10).** THEOREM. *The maps*

$$h \mapsto H, \quad e \mapsto E_+, \quad f \mapsto E_-$$

give $M(\lambda, \lambda_+, \lambda_-)$, as defined in (7.9) above, the structure of a cyclic $\mathfrak{sl}_2(\mathbb{K})$ -module generated by the coherent weight vector v_0 for the weight λ and having $E_\epsilon E_{-\epsilon} v_0 = \lambda_\epsilon v_0$.

PROOF. Within the definition (7.9) we find $E_\epsilon E_{-\epsilon} v_0 = a_\epsilon(\epsilon) v_0 = \lambda_\epsilon v_0$.

To verify $[H, E_\epsilon] = H E_\epsilon - E_\epsilon H$ we check equality on the basis vectors v_i :

$$\begin{aligned} [H, E_\epsilon] v_i &= (H E_\epsilon - E_\epsilon H) v_i \\ &= H E_\epsilon v_i - E_\epsilon H v_i \\ &= H a_\epsilon(i) v_{i-\epsilon} - E_\epsilon (\lambda - 2i) v_i \\ &= a_\epsilon(i) H v_{i-\epsilon} - (\lambda - 2i) E_\epsilon v_i \\ &= a_\epsilon(i) (\lambda - 2(i - \epsilon)) v_{i-\epsilon} - (\lambda - 2i) a_\epsilon(i) v_{i-\epsilon} \\ &= ((\lambda - 2(i - \epsilon)) - (\lambda - 2i)) a_\epsilon(i) v_{i-\epsilon} \\ &= 2\epsilon E_\epsilon v_i; \end{aligned}$$

so $[H, E_\epsilon] = 2\epsilon E_\epsilon$, as desired.

We must also verify $[E_+, E_-] = H$:

$$\begin{aligned} [E_+, E_-] v_i &= E_+ E_- v_i - E_- E_+ v_i \\ &= E_+ a_-(i) v_{i+1} - E_- a_+(i) v_{i-1} \\ &= a_-(i) a_+(i+1) v_i - a_+(i) a_-(i-1) v_i \\ &= (a_-(i) a_+(i+1) - a_+(i) a_-(i-1)) v_i \\ &= (\lambda - 2i) v_i = H v_i, \end{aligned}$$

where between the last two lines we have used Lemma (7.6). Accordingly $[E_+, E_-] = H$, which together with the preceding paragraph proves that we have a representation and module. $\square$

The module $M_+(\lambda, \lambda_+, \lambda_-)$ is a *generalized Verma module* for $\mathfrak{sl}_2(\mathbb{K})$ with weight λ . As an immediate consequence of the previous two results we have a universal property for generalized Verma modules:

(7.11). COROLLARY. *Let v be a coherent weight vector for the weight λ in the $\mathfrak{sl}_2(\mathbb{K})$ -module V with $E_\epsilon^V E_{-\epsilon}^V v = \lambda_\epsilon v$. Then the map $v_0 \mapsto v$ extends to a Lie module homomorphism taking the generalized Verma module $M(\lambda, \lambda_+, \lambda_-)$ onto the $\mathfrak{sl}_2(\mathbb{K})$ -submodule of V generated by v .* $\square$

(7.12). COROLLARY.

cor-gverm-struct

(a) $V = M(\lambda, \lambda_+, \lambda_-) = \bigoplus_{\mu \in \lambda + 2\mathbb{Z}} V_\mu$ with $\dim_{\mathbb{K}}(V_\mu) = 1$ for all $\mu \in \lambda + 2\mathbb{Z}$.

(b) Every weight vector in V is coherent.

- (c) Any $\mathbf{H}$ -submodule, and so any $\mathfrak{sl}_2(\mathbb{K})$ -submodule, of V is a sum $\bigoplus_{\mu \in I} V_\mu$, for some subset $I \subseteq \lambda + 2\mathbb{Z}$. $\square$

As already mentioned, each of the operators $\mathbf{H}$ and $\mathbf{E}_\epsilon$ is “nearly diagonal” on the basis $\{v_i \mid i \in \mathbb{Z}\}$. If we replace various of the v_i with nonzero scalar multiples, this will not change the near-diagonal structures, but it will change the values of the certain of the transition coefficients. A particular case is of interest.

(7.13). PROPOSITION. *For constants $0 \neq b, d \in \mathbb{K}$ with $bd = 1$, and an integer j , consider the new basis $\{v'_i \mid i \in \mathbb{Z}\}$ for $M_+(\lambda, \lambda_+, \lambda_-)$ given by*

$$\begin{aligned} v'_i &= v_i & \text{if } i > j \\ &= bv_i & \text{if } i \leq j. \end{aligned}$$

Then the transition parameters for this basis are equal to the transition parameters with respect to the original basis with only two exceptions:

$$a'_-(j) = ba_-(j) \text{ and } a'_+(j+1) = da_+(j+1).$$

PROOF. If $i > j+1$ or $i = j+1$ and $\epsilon = -$

$$\mathbf{E}_\epsilon(v'_i) = \mathbf{E}_\epsilon(v_i) = a_\epsilon(i)v_{i-\epsilon} = a_\epsilon(i)v'_{i-\epsilon}.$$

Similarly if $i < j$ or $i = j$ and $\epsilon = +$

$$\mathbf{E}_\epsilon(v'_i) = \mathbf{E}_\epsilon(bv_i) = a_\epsilon(i)bv_{i-\epsilon} = a_\epsilon(i)v'_{i-\epsilon}.$$

On the other hand,

$$\mathbf{E}_+(v'_{j+1}) = \mathbf{E}_+(v_{j+1}) = a_+(j+1)v_j = bda_+(j+1)v_j = da_+(j+1)v'_j$$

and

$$\mathbf{E}_-(v'_j) = \mathbf{E}_-(bv_j) = b\mathbf{E}_-(v_j) = ba_-(j)v_{j+1} = ba_-(j)v'_{j+1}. \quad \square$$

(7.14). THEOREM. *For a given $\delta \in \{\pm 1\}$, if $a_\delta(\delta) = \lambda_\delta$ is not equal to 0, then the map $\lambda_\delta^{-1}v_\delta \mapsto v''_0$ extends to an isomorphism of $M(\lambda, \lambda_+, \lambda_-)$ with $M(\lambda'', \lambda''_+, \lambda''_-)$ where $\lambda'' = \lambda + 2\delta$ and $\lambda''_{-\delta} = \lambda_\delta$.*

PROOF. For $M(\lambda, \lambda_+, \lambda_-)$ we start with

$$\begin{array}{ccccccccccc} \mathbf{E}_- & \dashrightarrow & v_{-2} & \xrightarrow{a_-(-2)} & v_{-1} & \xrightarrow{\lambda_-} & v_0 & \xrightarrow{1} & v_1 & \xrightarrow{1} & v_2 & \longrightarrow \\ & & \downarrow \lambda+4 & & \downarrow \lambda+2 & & \downarrow \lambda & & \downarrow \lambda-2 & & \downarrow \lambda-4 & \\ \mathbf{E}_+ & \longleftarrow & v_{-2} & \xleftarrow{1} & v_{-1} & \xleftarrow{1} & v_0 & \xleftarrow{\lambda_+} & v_1 & \xleftarrow{a_+(2)} & v_2 & \longleftarrow \end{array}$$

Assume $\lambda_- \neq 0$, and in the proposition set $j = -1$, $b = \lambda_-^{-1}$, and $d = \lambda_-$. We then find

$$\begin{array}{ccccccccccc}
 E_- & & \xrightarrow{\quad} & v'_{-2} & \xrightarrow{a_-(-2)} & v'_{-1} & \xrightarrow{1} & v'_0 & \xrightarrow{1} & v'_1 & \xrightarrow{1} & v'_2 & \longrightarrow \\
 & & & \downarrow \lambda+4 & & \downarrow \lambda+2 & & \downarrow \lambda & & \downarrow \lambda-2 & & \downarrow \lambda-4 & \\
 E_+ & \longleftarrow & v'_{-2} & \xleftarrow{1} & v'_{-1} & \xleftarrow{\lambda_-} & v'_0 & \xleftarrow{\lambda_+} & v'_1 & \xleftarrow{a_+(2)} & v'_2 & \longleftarrow
 \end{array}$$

Therefore the map $v'_i \mapsto v''_{i+1}$ gives an isomorphism of $M(\lambda, \lambda_+, \lambda_-)$ with the generalized Verma module $M(\lambda+2, \lambda_-, a_-(-2))$.

On the other hand, if $\lambda_+ \neq 0$ then in the proposition we set $j = 0$, $b = \lambda_+$, and $d = \lambda_+^{-1}$ to reveal an isomorphism of $M(\lambda, \lambda_+, \lambda_-)$ with the generalized Verma module $M(\lambda-2, a_+(2), \lambda_+)$. □ Class 25: 30 October '15

(7.15). COROLLARY.

cor-gverm-irred

- (a) If $a_\epsilon(i)$ is nonzero for all pairs (ϵ, i) , then $M(\lambda, \lambda_+, \lambda_-)$ is irreducible.
- (b) If $M(\lambda, \lambda_+, \lambda_-)$ is reducible, then there is a $\mu \in \lambda + 2\mathbb{Z}$ with $M(\lambda, \lambda_+, \lambda_-)$ isomorphic to $M(\mu, \mu, 0)$ or to $M(\mu, 0, -\mu)$.

PROOF. (a) By the theorem, under these circumstances $M(\lambda, \lambda_+, \lambda_-)$ is cyclically generated by every $\mathbb{K}v_i$; so by Corollary (7.12) there are no nonzero, proper submodules.

(b) By (a) if $M(\lambda, \lambda_+, \lambda_-)$ is reducible, then there is at least one pair (ϵ, i) with $a_\epsilon(i) = 0$. Choose the smallest $|i| = \epsilon i$ for which this is true. Then $a_\epsilon(\epsilon j) \neq 0$ for $1 \leq j < |i|$, and by the theorem $M(\lambda, \lambda_+, \lambda_-)$ is isomorphic to $M(\mu, \mu_+, \mu_-)$ with $\mu = \lambda - 2i$ and $\mu_\epsilon = a_\epsilon(i) = 0$. □

We therefore must analyse the submodule structure of the modules $M(\mu, \mu, 0)$ and $M(\mu, 0, -\mu)$. For this we have two important definitions.

(7.16). DEFINITION. Let $\lambda \in \mathbb{K}$ and set $M_+(\lambda) = \bigoplus_{i \in \mathbb{N}} \mathbb{K}v_i$ with $v_{-1} = 0 \in$ defn-high-verma $M_+(\lambda)$. Define the linear transformations H , E_+ , and E_- on $M_+(\lambda)$ by

$$\begin{aligned}
 H(v_i) &= (\lambda - 2i)v_i, \\
 E_\epsilon(v_i) &= a_\epsilon(i)v_{i-\epsilon},
 \end{aligned}$$

for $i \in \mathbb{N}$, $\epsilon \in \{\pm 1\}$, and $a_\epsilon(i)$ by

$$\begin{aligned}
 a_+(i) &= i(\lambda - i + 1) & \text{for } i \in \mathbb{Z}^+ \text{ and} \\
 a_-(i) &= 1 & \text{for } i \in \mathbb{N} :
 \end{aligned}$$

$$\begin{array}{ccccccc}
 E_- & & v_0 & \xrightarrow{1} & v_1 & \xrightarrow{1} & v_2 & \longrightarrow \\
 & & \downarrow \lambda & & \downarrow \lambda-2 & & \downarrow \lambda-4 & \\
 E_+ & & v_0 & \xleftarrow{\lambda} & v_1 & \xleftarrow{a_+(2)} & v_2 & \xleftarrow{\quad}
 \end{array}$$

(7.17). DEFINITION. Let $\lambda \in \mathbb{K}$ and set $M_-(\lambda) = \bigoplus_{i \in \mathbb{N}} \mathbb{K}v_{-i}$ with $v_1 = 0 \in$ defn-low-verma $M_-(\lambda)$. Define the linear transformations H , E_+ , and E_- on $M_-(\lambda)$ by

$$\begin{aligned} H(v_i) &= (\lambda - 2i)v_i, \\ E_\epsilon(v_i) &= a_\epsilon(i)v_{i-\epsilon}, \end{aligned}$$

for $i \in -\mathbb{N}$, $\epsilon \in \{\pm\} = \{\pm 1\}$, and $a_\epsilon(i)$ by

$$\begin{aligned} a_-(i) &= i(\lambda - i - 1) & \text{for } i \in -\mathbb{Z}^+ \text{ and} \\ a_+(i) &= 1 & \text{for } i \in -\mathbb{N} : \end{aligned}$$

$$\begin{array}{ccccccc} E_- & & \cdots \cdots \cdots & v_{-2} & \xrightarrow{a_-(-2)} & v_{-1} & \xrightarrow{-\lambda} v_0 \\ & & & \downarrow \lambda+4 & & \downarrow \lambda+2 & \downarrow \lambda \\ E_+ & \longleftarrow & v_{-2} & \xleftarrow{1} & v_{-1} & \xleftarrow{1} & v_0 \end{array}$$

The space $M_+(\lambda)$, defined above and with the described action, is the *Verma module* for $\mathfrak{sl}_2(\mathbb{K})$ with highest weight λ . Similarly $M_-(\lambda)$ with the described action is the *Verma module* for $\mathfrak{sl}_2(\mathbb{K})$ with lowest weight λ . At this point, these names are presumptive, since we have not proven that the maps $h \mapsto H$, $e \mapsto E_+$, and $f \mapsto E_-$ give $M_+(\lambda)$ or $M_-(\lambda)$ the structure of a cyclic $\mathfrak{sl}_2(\mathbb{K})$ -module generated by the weight vector v_0 with highest or lowest weight λ . This will be a consequence of the next theorem, where we will see each of these as quotient modules and submodules of appropriate generalized Verma modules.

thm-verma-submod **(7.18). THEOREM.** For each $\lambda \in \mathbb{K}$ we have following nonsplit exact sequences of $\mathfrak{sl}_2(\mathbb{K})$ -modules:

(a)

$$0 \longrightarrow M_-(\lambda + 2) \longrightarrow M(\lambda, \lambda, 0) \longrightarrow M_+(\lambda) \longrightarrow 0;$$

(b)

$$0 \longrightarrow M_+(\lambda - 2) \longrightarrow M(\lambda, 0, -\lambda) \longrightarrow M_-(\lambda) \longrightarrow 0.$$

PROOF. For $M(\lambda, \lambda, 0)$ we have

$$\begin{array}{ccccccccccc} E_- & & \cdots \cdots \cdots & v_{-2} & \xrightarrow{a_-(-2)} & v_{-1} & & v_0 & \xrightarrow{1} & v_1 & \xrightarrow{1} & v_2 & \longrightarrow \\ & & & \downarrow \lambda+4 & & \downarrow \lambda+2 & & \downarrow \lambda & & \downarrow \lambda-2 & & \downarrow \lambda-4 & \\ E_+ & \longleftarrow & v_{-2} & \xleftarrow{1} & v_{-1} & \xleftarrow{1} & v_0 & \xleftarrow{\lambda} & v_1 & \xleftarrow{a_+(2)} & v_2 & \longleftarrow \end{array}$$

Thus $K_- = \sum_{i \in \mathbb{Z}^-} \mathbb{K}v_i$ is a proper submodule and is, in fact, isomorphic to $M_-(\lambda+2)$. The quotient $M(\lambda, \lambda, 0)/K_-$ is next seen to be a copy of $M_+(\lambda)$. The extension is nonsplit since, by Corollary (7.12)(c), the only possible submodule complement to K_- would be $\sum_{i \in \mathbb{N}} \mathbb{K}v_i$ whereas $E_+v_0 \in K_-$.

For $M(\lambda, 0, -\lambda)$ instead

$$\begin{array}{ccccccccccc}
 E_- & \dashrightarrow & v_{-2} & \xrightarrow{a-(-2)} & v_{-1} & \dashrightarrow & v_0 & \xrightarrow{1} & v_1 & \xrightarrow{1} & v_2 & \longrightarrow \\
 & & \downarrow \lambda+4 & & \downarrow \lambda+2 & & \downarrow \lambda & & \downarrow \lambda-2 & & \downarrow \lambda-4 & \\
 E_+ & \longleftarrow & v_{-2} & \xleftarrow{1} & v_{-1} & \xleftarrow{1} & v_0 & & v_1 & \xleftarrow{a_+(2)} & v_2 & \longleftarrow
 \end{array}$$

Here $K_+ = \sum_{i \in \mathbb{Z}^+} \mathbb{K}v_i$ is a proper submodule and is isomorphic to $M_+(\lambda - 2)$. The quotient $M(\lambda, 0, \lambda)/K_+$ is a copy of $M_-(\lambda)$. Again the extension is nonsplit as the only possible complement to K_+ would be $\sum_{i \in \mathbb{N}} \mathbb{K}v_{-i}$ but $E_-v_0 \in K_+$. $\square$

Especially, the spaces $M_\epsilon(\lambda)$ are indeed $\mathfrak{sl}_2(\mathbb{K})$ -modules, as presumed above.

(7.19). COROLLARY.

cor-univ-verma-bis

- (a) Let v be a highest weight vector for the weight λ in the $\mathfrak{sl}_2(\mathbb{K})$ -module V . Then the map $v_0 \mapsto v$ extends to a Lie module homomorphism taking the Verma module $M_+(\lambda)$ with highest weight λ onto the $\mathfrak{sl}_2(\mathbb{K})$ -submodule of V generated by v .
- (b) Let v be a lowest weight vector for the weight λ in the $\mathfrak{sl}_2(\mathbb{K})$ -module V . Then the map $v_0 \mapsto v$ extends to a Lie module homomorphism taking the Verma module $M_-(\lambda)$ with lowest weight λ onto the $\mathfrak{sl}_2(\mathbb{K})$ -submodule of V generated by v . $\square$

Class 26: 2 November '15

(7.20). THEOREM.

thm-verm-struct

- (a) The module $M_\epsilon(\lambda)$ is irreducible if and only if $\epsilon\lambda \notin \mathbb{N}$.
- (b) If $\lambda = n \in \epsilon\mathbb{N}$, then $M_\epsilon(n)$ is indecomposable with two composition factors:

$$0 \longrightarrow M_\epsilon(-n - 2\epsilon) \longrightarrow M_\epsilon(n) \longrightarrow M_0(\epsilon n + 1) \longrightarrow 0,$$

with $M_\epsilon(-n - 2\epsilon)$ irreducible of infinite dimension and $M_0(\epsilon n + 1)$ irreducible of finite dimension $\epsilon n + 1$.

PROOF. By Corollary (7.12) every submodule is $\sum_{i \in I} \mathbb{K}v_i$ for some $I \subseteq \epsilon\mathbb{N}$. As it is E_- -invariant, a nonzero submodule must be $\sum_{k \leq i \in \mathbb{N}} \mathbb{K}v_{\epsilon i}$ for some $k \in \mathbb{N}$. When $a_\epsilon(i)$ is nonzero for all $i \in \epsilon\mathbb{N}$, the module $M_\epsilon(\lambda)$ itself is the only such E_- -invariant subspace. Therefore $M_\epsilon(\lambda)$ is irreducible unless

$$a_\epsilon(i) = i(\lambda - i + \epsilon) = i((\lambda + \epsilon) - i) = 0$$

for some $i \in \epsilon\mathbb{Z}^+$. For a fixed λ and ϵ this can only happen for

$$i = \lambda + \epsilon \in \epsilon\mathbb{Z}^+,$$

which is to say

$$\lambda = i - \epsilon \in \epsilon\mathbb{N}.$$

Especially if $\lambda \notin \epsilon\mathbb{N}$, then $M_\epsilon(\lambda)$ is irreducible.

Now suppose $\lambda = n = \epsilon m \in \epsilon\mathbb{N}$ for $m \in \mathbb{N}$, so that $\epsilon(m+1) = \lambda + \epsilon$ with $a_\epsilon(\lambda + \epsilon) = 0$. Then $K_\epsilon = \bigoplus_{\epsilon j \geq m+1} \mathbb{K}v_j$ is a submodule of $M_\epsilon(\lambda)$, complemented by the $(m+1)$ -subspace $\bigoplus_{\epsilon j=0}^m \mathbb{K}v_j$. In particular $M_\epsilon(\lambda)$ is reducible, completing (a).

For the submodule $K_\epsilon = \bigoplus_{\epsilon j \geq m+1} \mathbb{K}v_j$ we calculate

$$H(v_{\epsilon(m+1)}) = H(v_{\lambda+\epsilon}) = \lambda - 2(\lambda + \epsilon) = -\lambda - 2\epsilon = -n - 2\epsilon.$$

Thus K_ϵ is isomorphic to $M_\epsilon(-n-2\epsilon)$, an infinite dimensional irreducible module as $\epsilon(-n-2\epsilon) = -\epsilon n - 2 \in \mathbb{Z}^-$. The extension is nonsplit since, by Corollary (7.12)(c) and the Third Isomorphism Theorem, the only possible complementary submodule would be $\bigoplus_{\epsilon j=0}^m \mathbb{K}v_j$, whereas $E_{-\epsilon}v_{\epsilon m} \in K_\epsilon$.

For $m \in \mathbb{N}$, the quotients $M_\epsilon(\epsilon m)/K_\epsilon$ have dimension $m+1$ and are irreducible, since there are no further solutions to $i((\lambda + \epsilon) - i) = 0$. By Theorem (7.5) every finite dimensional irreducible $\mathfrak{sl}_2(\mathbb{K})$ -module has both high weight vectors and low weight vectors. In particular each $M_\epsilon(\epsilon m)/K_\epsilon$ must also be a quotient of some $M_{-\epsilon}(-\epsilon k)/K_\epsilon$, dimension considerations forcing $m = k$. That is, the two finite dimensional modules $M_\epsilon(\epsilon m)/K_\epsilon$ are isomorphic. We conclude that, up to isomorphism, there is a unique irreducible $\mathfrak{sl}_2(\mathbb{K})$ -module of each positive dimension $m+1$. This we have denoted $M_0(m+1)$. $\square$

For example, with $\epsilon = +$ and $\lambda = m = 3$ we have

$$\begin{array}{ccccccccc} E_- & v_0 & \xrightarrow{1} & v_1 & \xrightarrow{1} & v_2 & \xrightarrow{1} & v_3 & \xrightarrow{1} & v_4 & \xrightarrow{1} & v_5 \\ & \downarrow 3 & & \downarrow 1 & & \downarrow -1 & & \downarrow -3 & & \downarrow -5 & & \\ E_+ & v_0 & \xleftarrow{-3} & v_1 & \xleftarrow{-4} & v_2 & \xleftarrow{-3} & v_3 & & v_4 & \xleftarrow{-5} & v_5 \end{array}$$

and the submodule $K_+ = \bigoplus_{j \geq 4} \mathbb{K}v_j$ of $M_+(3)$ is revealed as a copy of $M_+(-5)$, while the quotient module is $\bar{M}_0(4)$ with dimension 4 and weights $\{3, 1, -1, -3\}$.

We have already seen on page 19 a version of the module $M_0(m+1)$. Let

$$e \mapsto x \frac{\partial}{\partial y} \quad \text{and} \quad f \mapsto y \frac{\partial}{\partial x} \quad \text{and} \quad h \mapsto x \frac{\partial}{\partial x} - y \frac{\partial}{\partial y}$$

in $\text{Der}_{\mathbb{K}}(\mathbb{K}[x, y])$ acting on $\mathbb{K}[x, y]_m$, the space of homogeneous polynomials of total degree m in $\mathbb{K}[x, y]$. The polynomial x^m is a weight vector with highest weight m , while y^m is a lowest weight vector with weight $-m$; we have a module isomorphism with $M_+(m)/K_+$ given by $v_i + K_+ \mapsto x^{m-i}y^i$, for $0 \leq i \leq m$.

Class 27: 4 November '15

The next result is a corollary of the previous one and is of fundamental importance. Versions of it are at the heart of the representation theory for all semisimple Lie algebras over algebraically closed fields of characteristic 0.

thm-class-hw-mod

(7.21). THEOREM. (CLASSIFICATION OF IRREDUCIBLE HIGHEST WEIGHT MODULES) *For every $\lambda \in \mathbb{K}$, up to isomorphism there is a unique irreducible $\mathfrak{sl}_2(\mathbb{K})$ -module $L_+(\lambda)$ with highest weight λ . Indeed*

- (a) if $\lambda \notin \mathbb{N}$, then $L_+(\lambda) = M_+(\lambda)$ of infinite dimension;
 (b) if $\lambda = n \in \mathbb{N}$, then $L_+(\lambda) = M_0(n+1)$ of finite dimension $n+1$.

These are pairwise nonisomorphic. $\square$

There is, of course, a corresponding result for irreducible lowest weight modules.

Every irreducible finite dimensional $\mathfrak{sl}_2(\mathbb{K})$ -module is a highest weight module, so we also have the following result which was mentioned in the introduction to this chapter and will aid us to complete the classification of semisimple Lie algebras over $\mathbb{K}$.

(7.22). THEOREM. *Up to isomorphism the finite dimensional irreducible $\mathfrak{sl}_2(\mathbb{K})$ -modules are the modules $M_0(m+1)$ for $m \in \mathbb{N}$. Especially, these modules are self-dual. In $M_0(m+1)$ the weights are $-m, -m+2, \dots, m-2, m$, and each weight space has dimension 1.* $\square$ thm-fd-irred-sl2-mod

7.3 The Casimir operator

(7.23). PROPOSITION. *Let V be an $\mathfrak{sl}_2(\mathbb{K})$ -module, and in $\text{End}_{\mathbb{K}}(V)$ define the element* prop-sltwo-casimir

$$C = C^V = (H^V)^2 + 1 + 2(E_+^V E_-^V + E_-^V E_+^V).$$

- (a) $C = (H - \epsilon)^2 + 4E_{\epsilon}E_{-\epsilon}$.
 (b) $CH = HC$ and $CE_{\epsilon} = E_{\epsilon}C$.

PROOF. We make frequent use of Lemma (7.3):

$$\begin{aligned} C &= H^2 + 1 + 2(E_{\epsilon}E_{-\epsilon} + E_{-\epsilon}E_{\epsilon}) \\ &= H^2 + 1 + 2(E_{\epsilon}E_{-\epsilon} - \epsilon H + E_{\epsilon}E_{-\epsilon}) \\ &= (H^2 - 2\epsilon H + 1) + 4E_{\epsilon}E_{-\epsilon} \\ &= (H - \epsilon)^2 + 4E_{\epsilon}E_{-\epsilon}. \end{aligned}$$

$$\begin{aligned} HC &= H(H - \epsilon)^2 + 4(HE_{\epsilon})E_{-\epsilon} \\ &= (H - \epsilon)^2 H + 4E_{\epsilon}((H + 2\epsilon)E_{-\epsilon}) \\ &= (H - \epsilon)^2 H + 4E_{\epsilon}E_{-\epsilon}(H + 2\epsilon - 2\epsilon) \\ &= CH. \end{aligned}$$

$$\begin{aligned} CE_{\epsilon} &= (H - \epsilon)^2 E_{\epsilon} + 4E_{\epsilon}E_{-\epsilon}E_{\epsilon} \\ &= E_{\epsilon}(H - \epsilon + 2\epsilon)^2 + 4E_{\epsilon}E_{-\epsilon}E_{\epsilon} \\ &= E_{\epsilon}((H + \epsilon)^2 + 4E_{-\epsilon}E_{\epsilon}) \\ &= E_{\epsilon}C. \quad \square \end{aligned}$$

The somewhat mysterious $C = C^V$ is the *Casimir operator* on V . It has its uses.

(7.24). PROPOSITION. *In an irreducible weight module for $\mathfrak{sl}_2(\mathbb{K})$, every prop-irred-coherent weight vector is coherent.*

PROOF. If the $\mathfrak{sl}_2(\mathbb{K})$ -module V is irreducible, then by Schur's Lemma (4.11) the Casimir operator C acts as a scalar: there is a $c \in \mathbb{K}$ with $Cv = cv$ for all $v \in V$. Let v be a weight vector in V , say, for the weight λ . As $C = (H - \epsilon)^2 + 4E_\epsilon E_{-\epsilon}$, we have $E_\epsilon E_{-\epsilon} = \frac{1}{4}(C - (H - \epsilon)^2)$. Thus

$$\begin{aligned} E_\epsilon E_{-\epsilon} v &= \frac{1}{4}(C - (H - \epsilon)^2)v \\ &= \frac{1}{4}(Cv - (H - \epsilon)^2 v) \\ &= \frac{c - (\lambda - \epsilon)^2}{4} v. \end{aligned}$$

That is, v is a coherent weight vector for λ with $\lambda_\epsilon = \frac{c - (\lambda - \epsilon)^2}{4}$. $\square$

The *spectrum* of an $\mathfrak{sl}_2(\mathbb{K})$ -module V is the set of weights associated with weight vectors in the module.

thm-irred-sl2-mod **(7.25). THEOREM.** *An irreducible weight module for $\mathfrak{sl}_2(\mathbb{K})$ is isomorphic to one of the following:*

- (1) $M_0(m+1)$ for $m \in \mathbb{N}$ with spectrum $[-m, m] \cap 2\mathbb{Z}$;
- (2) $M_+(\lambda)$ for $\lambda \notin \mathbb{N}$ with spectrum $\lambda - 2\mathbb{N}$;
- (3) $M_-(\lambda)$ for $-\lambda \notin \mathbb{N}$ with spectrum $\lambda + 2\mathbb{N}$;
- (4) $M(\lambda, \lambda_+, \lambda_-)$ with $\lambda_- = \lambda_+ - \lambda$ and spectrum $\lambda + 2\mathbb{Z}$, such that the polynomial $x^2 - (\lambda - 1)x - \lambda_+ \in \mathbb{K}[x]$ has no integer roots.

All these modules are irreducible. No module from one case is isomorphic to a module from one of the other cases. In the first three cases, two modules are isomorphic if and only if they have the same parameter. In the last case, $M(\lambda, \lambda_+, \lambda_-)$ is isomorphic to $M(\mu, \mu_+, \mu_-)$ if and only if $\mu - \lambda \in 2\mathbb{Z}$ and $\mu_+ - \lambda_+ = \frac{1}{4}(\lambda - \mu)(\lambda + \mu - 2)$.

PROOF. By the previous proposition, an irreducible weight module V is generated by a coherent weight vector. If the module is not isomorphic to some $M(\lambda, \lambda_+, \lambda_-)$ (as in the last case), then by Corollary (7.15) it is a quotient of a Verma module $M_\epsilon(\lambda)$. By Theorem (7.20) the module V is then isomorphic to one of the examples in the first three conclusions, all irreducible.

The various spectra are also clear, and show that no module from one case is isomorphic to one from another case, nor can different parameters in any one of the first three cases produce isomorphic modules.

It remains to decide under what circumstances $M(\lambda, \lambda_+, \lambda_-)$ is irreducible and when two such modules can be isomorphic.

By Corollaries (7.7) and (7.15)(a) the module $M(\lambda, \lambda_+, \lambda_-)$ is irreducible if and only if $b(j) = j(\lambda - j - 1) + \lambda_+$ is nonzero for all $j \in \mathbb{Z}$. This is the case precisely when

$$x(\lambda - x - 1) + \lambda_+ = -x^2 + (\lambda - 1)x + \lambda_+ \in \mathbb{K}[x]$$

has no integral roots.

Suppose that $M(\lambda, \lambda_+, \lambda_-)$ and $M(\mu, \mu_+, \mu_-)$ are isomorphic. By spectral considerations, we must have $\lambda + 2\mathbb{Z} = \mu + 2\mathbb{Z}$. By symmetry we may assume that $\mu = \lambda - 2i$ for some $i \in \mathbb{N}$. In that case, isomorphism holds if and only if, in terms of the transformation coefficients for $M(\lambda, \lambda_+, \lambda_-)$, we have

$$\mu_+ = a_+(i + 1) = ((i + 1) - 1)(\lambda - (i + 1)) + \lambda_+ = i(\lambda - i - 1) + \lambda_+.$$

As $\mu = \lambda - 2i$ we have $i = \frac{\lambda - \mu}{2}$, so this becomes $\mu_+ - \lambda_+ = \frac{1}{4}(\lambda - \mu)(\lambda + \mu - 2)$, as claimed. $\square$

If V is an irreducible module containing the (coherent) weight vector v for λ with $E_\epsilon E_{-\epsilon} v = \lambda_\epsilon$, then, as in the proof of Proposition (7.24), we can calculate the Casimir scalar as

$$c = 4\lambda_\epsilon + (\lambda - \epsilon)^2.$$

For instance, for $V = M_0(m + 1)$ we have $\lambda = m$ and $\lambda_- = 0$ so that $c = 4(0) + (m + 1)^2 = (m + 1)^2$.

This also allows effective relabelling of the irreducible modules $M(\lambda, \lambda_+, \lambda_-)$. Indeed this module can now be characterized by the two parameters $\lambda + 2\mathbb{Z} \in (\mathbb{K}, +)/2\mathbb{Z}$ and $c (= 4\lambda_\epsilon + (\lambda - \epsilon)^2)$, two such irreducible modules being isomorphic if and only if they have the same parameter pair. (Exercise.) This is the approach taken in [Maz10, Theorem 3.32]. The corresponding requirement for irreducibility is that $c \neq (\mu + 1)^2$ for all $\mu \in \lambda + 2\mathbb{Z}$. (Exercise.)

We have a second hidden use of the Casimir operator.

Class 28: 6 November '15

(7.26). THEOREM. (WEYL'S THEOREM) *Every finite dimensional $\mathfrak{sl}_2(\mathbb{K})$ -module is completely reducible.* thm-weyl-sltwo

PROOF. Equivalently, an extension V of a finite dimensional irreducible $\mathfrak{sl}_2(\mathbb{K})$ -module $W \simeq M_0(m + 1)$ by a second finite dimensional irreducible module is always split. Consider

$$0 \longrightarrow M_0(m + 1) \longrightarrow V \longrightarrow M_0(n + 1) \longrightarrow 0.$$

By passing to the dual of V , if necessary, we may assume $m \leq n$. If $m < n$, then the weight space V_n has dimension 1. But then (say, by Corollary (7.19) and Theorem (7.20)), V_n generates a submodule U of V that is isomorphic to $M_0(n + 1)$ and splits the extension as $V = W \oplus U$. Therefore we may assume $m = n$.

Each of the weight spaces V_n for $n \in [-m, m] \cap 2\mathbb{Z}$ has dimension 2 with $V_m = \ker E_+$ and $V_{-m} = \ker E_-$. By Lemma (7.3)(d) we have

$$V_m = E_+^m V = E_+^m V_{-m},$$

with $\ker E_+^m = \bigoplus_{i=-m+2}^m V_i$ and $A = E_+^m$ an invertible linear transformation from V_{-m} to V_m .

For $C = C^V$, the Casimir operator, we calculate the map $CA = AC$ from V_{-m} to V_m in two ways. For $v \in V_{-m}$ we first have

$$\begin{aligned} (CA)v &= C(Av) \\ &= ((H+1)^2 + 4E_-E_+)(Av) \\ &= (H+1)^2(Av), \end{aligned}$$

as $Av \in V_m = \ker E_+$. Similarly as $v \in V_{-m} = \ker E_-$,

$$\begin{aligned} (AC)v &= A(Cv) \\ &= A((H-1)^2 + 4E_+E_-)v \\ &= A(H-1)^2v \\ &= E_+^m(H-1)^2v \\ &= (H-1-2m)^2E_+^mv && \text{by Lemma (7.3)(c)} \\ &= (H-1-2m)^2(Av) \end{aligned}$$

Therefore, from V_{-m} to V_m , the map $AC = CA$ is equal to

$$(H+1)^2A = (H-1-2m)^2A.$$

Since A is a bijection from V_{-m} to V_m , this says that, as a map from V_m to V_m ,

$$0 = (H+1)^2 - (H-1-2m)^2 = (4m+4)(H-m).$$

We already knew that $(H-m)^2V_m = 0$, because the weight space V_m has dimension 2. Now we have learned that $(H-m)V_m = 0$, so $V_m = V_m^w$ consists entirely of weight vectors (and 0).

Let $u \in V_m \setminus W$. Then again by Corollary (7.19) and Theorem (7.20), the weight vector u generates a submodule U of V that is isomorphic to $M_0(m+1)$ and splits the extension as $V = W \oplus U$. $\square$

Weyl's Theorem is valid for all finite dimensional semisimple algebras over $\mathbb{K}$, not just $\mathfrak{sl}_2(\mathbb{K})$. The proof uses the appropriate generalization of the current Casimir operator.

7.4 Problems

prob-maz **(7.27).** PROBLEM. Let $\lambda, \lambda_+, \lambda_- \in \mathbb{K}$ with $\lambda_+ - \lambda_- = \lambda$. Set $M_*(\lambda, \lambda_+, \lambda_-) = \bigoplus_{i \in \mathbb{Z}} \mathbb{K}v_i$. Define the linear transformations H , E_+ , and E_- on $M_*(\lambda, \lambda_+, \lambda_-)$ by

$$\begin{aligned} H(v_i) &= (\lambda - 2i)v_i; \\ E_+(v_i) &= v_{i-1}; \\ E_-(v_i) &= b(i)v_{i+1}, \end{aligned}$$

for all $i \in \mathbb{Z}$ with $b(i) = i(\lambda - i - 1) + \lambda_+$ (as in Corollary (7.7)).

(a) *Prove that the maps*

$$h \mapsto H, \quad e \mapsto E_+, \quad f \mapsto E_-$$

give $M_*(\lambda, \lambda_+, \lambda_-)$ the structure of an $\mathfrak{sl}_2(\mathbb{K})$ -module in which each weight space has dimension 1.

(b) *Prove that if the generalized Verma module $M(\lambda, \lambda_+, \lambda_-)$ is irreducible, then it is isomorphic to $M_*(\lambda, \lambda_+, \lambda_-)$.*

(c) *Prove that $M_0(m+1)$ is never a quotient of $M_*(\lambda, \lambda_+, \lambda_-)$. In particular if the generalized Verma module $M(\lambda, \lambda_+, \lambda_-)$ is reducible, then it need not be isomorphic to $M_*(\lambda, \lambda_+, \lambda_-)$.*

REMARK. *The modules $M_*(\lambda, \lambda_+, \lambda_-)$ are those used in Mazorchuk's excellent book [Maz10, Chapter 3].*

Other Stuff as of: 11 November 2015

fake index: *(0)-dummy index 2015-11-11*

Bibliography

- [Alb34] Albert, A. Adrian(1-IASP) On a certain algebra of quantum mechanics. *Ann. of Math. (2)* 35 (1934), no. 1, 6573.
- [Car05] R. Carter, Carter, R. W.(4-WARW-MI) Lie algebras of finite and affine type. *Cambridge Studies in Advanced Mathematics*, 96. Cambridge University Press, Cambridge, 2005.
- [CSM95] Carter, Roger; Segal, Graeme; Macdonald, Ian Lectures on Lie groups and Lie algebras. With a foreword by Martin Taylor. *London Mathematical Society Student Texts*, 32. Cambridge University Press, Cambridge, 1995.
- [Coh54] Cohn, P. M. On homomorphic images of special Jordan algebras. *Canadian J. Math.* 6, (1954). 253264.
- [Dre99] Dresner, Lawrence Applications of Lie's theory of ordinary and partial differential equations. (English summary) *Institute of Physics Publishing*, Bristol, 1999.
- [Eld15] A. Elduque, *Course notes: Lie algebras*, Univ. de Zaragoza, 2015, pp. 1–114.
- [Gle52] Gleason, Andrew M. Groups without small subgroups. *Ann. of Math. (2)* 56, (1952). 193–212.
- [Gor09] I. Gordan, *Notes: Infinite dimensional Lie algebras*, Univ. Edinburgh, 2008-09, pp. 1–55.
- [Hal15] Hall, Brian Lie groups, Lie algebras, and representations. An elementary introduction. Second edition. *Graduate Texts in Mathematics*, 222. Springer, Cham, 2015
- [Hel01] Helgason, Sigurdur Differential geometry, Lie groups, and symmetric spaces. (English summary) Corrected reprint of the 1978

- original. Graduate Studies in Mathematics, 34. American Mathematical Society, Providence, RI, 2001.
- [Hig58] G. Higman, Lie ring methods in the theory of finite nilpotent groups. 1960 Proc. Internat. Congress Math. 1958 pp. 307312 Cambridge Univ. Press, New York
- [How83] Howe, Roger Very basic Lie theory. Amer. Math. Monthly 90 (1983), no. 9, 600623.
- [Hum78] Humphreys, James E. Introduction to Lie algebras and representation theory. Second printing, revised. Graduate Texts in Mathematics, 9. Springer-Verlag, New York-Berlin, 1978.
- [Jac79] N. Jacobson, Jacobson, Nathan Lie algebras. Republication of the 1962 original. Dover Publications, Inc., New York, 1979.
- [Jac89] N. Jacobson, "Basic Algebra II," Second edition, W.H. Freeman and Company, New York, 1989.
- [JvNW34] Jordan, P.; von Neumann, J.; Wigner, E. On an algebraic generalization of the quantum mechanical formalism. Ann. of Math. (2) 35 (1934), no. 1, 2964.
- [Kir08] Kirillov, Alexander, Jr. An introduction to Lie groups and Lie algebras. Cambridge Studies in Advanced Mathematics, 113. Cambridge University Press, Cambridge, 2008.
- [Lie??] S. Lie,
- [Lee13] J.M. Lee, "Introduction to Smooth Manifolds," Second edition. Graduate Texts in Mathematics, 218. Springer, New York, 2013.
- [Maz10] Mazorchuk, Volodymyr(S-UPPS) Lectures on $\mathfrak{sl}_2(\mathbb{C})$ -modules. Imperial College Press, London, 2010.
- [MoZi55] Montgomery, Deane; Zippin, Leo Topological transformation groups. Interscience Publishers, New York-London, 1955.
- [vNe29] v. Neumann, J. ber die analytischen Eigenschaften von Gruppen linearer Transformationen und ihrer Darstellungen. (German) Math. Z. 30 (1929), no. 1, 342
- [Ros02] Rossmann, Wulf Lie groups. An introduction through linear groups. Oxford Graduate Texts in Mathematics, 5. Oxford University Press, Oxford, 2002.
- [Ser06] Serre, Jean-Pierre Lie algebras and Lie groups. 1964 lectures given at Harvard University. Corrected fifth printing of the second (1992) edition. Lecture Notes in Mathematics, 1500. Springer-Verlag, Berlin, 2006.

- [Spi65] M. Spivak, “Calculus on Manifolds,” Mathematics Monograph Series, W.A. Benjamin, 1965.
- [SpV00] T.A. Springer and F.D. Veldkamp, “Octonions, Jordan Algebras and Exceptional Groups,” Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2000.
- [Ste70] Stewart, Ian Lie algebras. Lecture Notes in Mathematics, Vol. 127 Springer-Verlag, Berlin-New York 1970 iv+97 pp.
- [Sti08] Stillwell, John Naive Lie theory. Undergraduate Texts in Mathematics. Springer, New York, 2008. xiv+217
- [Tao14] Tao, Terence Hilbert’s fifth problem and related topics. Graduate Studies in Mathematics, 153. American Mathematical Society, Providence, RI, 2014.
- [Tap05] Tapp, Kristopher Matrix groups for undergraduates. Student Mathematical Library, 29. American Mathematical Society, Providence, RI, 2005
- [Tit??] J. Tits, T-K-K
- [Tu11] Tu, Loring W.(1-TUFT) An introduction to manifolds. Second edition. Universitext. Springer, New York, 2011.
- [Vin03] Vinberg, E. B. A course in algebra. Translated from the 2001 Russian original by Alexander Retakh. Graduate Studies in Mathematics, 56. 2003.
- [Zel97] E. Zelmanov Zelmanov, Efim I.(1-YALE) On the restricted Burnside problem. Fields Medallists’ lectures, 623632, World Sci. Ser. 20th Century Math., 5, World Sci. Publ., River Edge, NJ, 1997.
- [Zor31] M. Zorn, *Theorie der alternativen Ringe*, Abh. Math. Sem. Hamburg Univ., **8** (1931), 123–147

Index

- (0)-dummy index 2015-11-11, 80
- A -module, 8
- $M_0(m+1)$, 19, 74
- $[A, B]$, 40
- $[A; B, n]$, 42
- $M_+(\lambda, \lambda_+, \lambda_-)$, 69
- $M_+(\lambda)$, 72
- $M_-(\lambda)$, 72
- ad_x^V , 44
- α -string through β , 53
- σ -hermitian, 15
- σ -sesquilinear form, 15
- $\mathbb{K}$ -algebra, 1
- $\mathbb{K}$ -form, 1
- $\mathbb{Z}$ -algebra, 2
- $\mathbb{Z}$ -form, 2
- $\mathfrak{f}_4$, 19, 86
- $\mathfrak{f}_4(\mathbb{K})$, 19
- $\mathfrak{g}_2$, 19, 86
- $\mathfrak{g}_2(\mathbb{K})$, 19
- $L^{k+1}(G)$, 21
- PBW Theorem, 11

- abelian Lie algebras, 13
- acts, 8
- adjoint map, 1
- adjoint representation, 33
- Ado-Iwasawa Theorem, 11
- affine Lie algebra, 21
- Albert algebra, 7
- alternative $\mathbb{K}$ -algebras, 3
- associative $\mathbb{K}$ -algebras, 2
- associative form, 55

- bilinear forms, 15

- Campbell-Baker-Hausdorff Theorem, 38
- carries, 8
- Cartan decomposition, 52
- Cartan subalgebra, 50
- Cartan's Semisimplicity Criterion, 58
- Cartan's Solvability Criterion, 57
- Casimir operator, 75
- center, 11
- characteristic, 40
- Classification of irreducible highest weight modules, 74
- closed subgroups, 27
- coherent, 66
- coherent weight vector, 66
- commutator, 40
- composition algebra, 4
- curve, 31

- dense, 36
- derivation, 16
- derivation algebra, 16
- derivative of a matrix function, 28
- derived subalgebra, 14
- differential, 35
- division algebra, 5

- Engel condition, 42
- Engel's Theorem, 48
- equivalence, 35
- exceptional Jordan algebras, 11
- extend scalars, 1
- external semidirect product, 43

- faithful, 8, 35
- First Isomorphism Theorem, 39
- Fourier polynomials, 20
- Fourier series, 20
- full, 36
- general linear algebra, 14
- generalized Verma module, 69
- Gram matrix, 15
- Heisenberg algebra, 23
- hermitian, 15
- highest weight vector, 66
- Hilbert's Fifth Problem, 27
- Hurwitz' Theorem, 5
- ideal, 39
- inner derivation, 17
- inner derivation algebra, 17
- internal semidirect product, 43
- invariant form, 55
- Isomorphism Theorems, 39
- Jacobi Identity, 10, 39
- Jordan algebra, 6
- Jordan product, 6
- Kac-Moody Lie algebra, 21
- kernel, 39
- Killing form, 55
- Krull-Schmidt Theorem, 42
- lattice, 2
- Laurent polynomials, 19
- Leibniz product rule, 16
- Lie algebra, 9, 10, 39
- Lie algebra of G , 34
- Lie group, 26
- Lie homomorphism, 39
- Lie ring, 22
- Lie's Second Theorem, 36
- Lie's Third Theorem, 37
- linear associative algebra, 8
- linear Jordan algebra, 8
- linear Lie algebra, 8
- linear representation, 8
- lower central series, 21
- lowest weight vector, 66
- matrix derivative, 28
- module, 8
- multiplication, 1
- multiplication coefficients, 1
- nil, 47
- nil representation, 47
- nilpotent radical, 41
- nondegenerate, 15
- normalizer, 48
- null identical relation, 9
- octonion algebras, 5
- octonions, 4
- one-parameter subgroup, 29
- opposite algebra, 1
- path connectivity, 36
- product rule, 16
- pure imaginary, 4
- quadratic form, 4
- quadratic Jordan algebras, 12
- quaternions, 4
- quotient algebra, 39
- radical, 41
- rank, 49
- regular, 49
- Restricted Burnside Problem, 22
- root, 51
- Schur's Lemma, 42
- Second Isomorphism Theorem, 40
- self-normalizing, 48
- semisimple, 41
- sesquilinear form, 15
- simply connected, 36
- skew identical relation, 9
- smooth, 27
- smooth vector field, 19
- solvable radical, 41
- special Jordan algebras, 11
- special linear algebra, 14
- spectrum, 76

- split, 5
- split extension, 43
- structure constants, 1
- subalgebra, 39
- symmetric bilinear forms, 15

- tangent space, 31, 32
- Third Isomorphism Theorem, 40
- torsion-free, 65
- transformation coefficients, 68
- trivial weight, 51
- type $\mathfrak{f}_4$, 19
- type $\mathfrak{g}_2$, 19

- variety, 2
- vector field, 19
- Verma module, 72
- Virasoro algebra, 20

- weight, 51
- weight module, 65
- weight space, 51
- weight vector, 65
- Weyl's Theorem, 77
- Witt algebra, 19

- zero weight, 51
- Zorn's vector matrices, 5