Lie Algebras

Shlomo Sternberg

February 18, 2004

Chapter 2

sl(2)

In this chapter (and in most of the succeeding chapters) all Lie algebras and vector spaces are over the complex numbers.

2.1 Low dimensional Lie algebras.

Any one dimensional Lie algebra must be commutative, since $[X, X] = 0$ in any Lie algebra.

If $\mathbf{g}$ is a two dimensional Lie algebra, say with basis X, Y then $[aX+bY, cX+dY] = (ad-bc)[X, Y]$, so that there are two possibilities: $[X, Y] = 0$ in which case $\mathbf{g}$ is commutative, or $[X, Y] \neq 0$, call it B, and the Lie bracket of any two elements of $\mathbf{g}$ is a multiple of B. So if C is not a multiple of B, we have $[C, B] = cB$ for some $c \neq 0$, and setting $A = c^{-1}C$ we get a basis A, B of $\mathbf{g}$ with the bracket relations

$$[A, B] = B.$$

This is an interesting Lie algebra; it is the Lie algebra of the group of all affine transformations of the line, i.e. all transformations of the form

$$x \mapsto ax + b, \quad a \neq 0.$$

For this reason it is sometimes called the "$ax + b$ group". Since

$$\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} ax + b \\ 1 \end{pmatrix}$$

we can realize the group of affine transformations of the line as a group of two by two matrices. Writing

$$a = \exp tA, \quad b = tB$$

so that

$$\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} = \exp t \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} = \exp t \begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix}$$

we see that our algebra $\mathbf{g}$ with basis A, B and $[A, B] = B$ is indeed the Lie algebra of the $ax + b$ group.

In a similar way, we could list all possible three dimensional Lie algebras, by first classifying them according to $\dim[\mathbf{g}, \mathbf{g}]$ and then analyzing the possibilities for each value of this dimension. Rather than going through all the details, we list the most important examples of each type. If $\dim[\mathbf{g}, \mathbf{g}] = 0$ the algebra is commutative so there is only one possibility.

A very important example arises when $\dim[\mathbf{g}, \mathbf{g}] = 1$ and that is the Heisenberg algebra, with basis P, Q, Z and bracket relations

$$[P, Q] = Z, \quad [Z, P] = [Z, Q] = 0.$$

Up to constants (such as Planck's constant and i) these are the famous Heisenberg commutation relations. Indeed, we can realize this algebra as an algebra of operators on functions of one variable x: Let $P = D =$ differentiation, let Q consist of multiplication by x. Since, for any function $f = f(x)$ we have

$$D(xf) = f + xf'$$

we see that $[P, Q] = \mathrm{id}$, so setting $Z = \mathrm{id}$, we obtain the Heisenberg algebra.

As an example with $\dim[\mathbf{g}, \mathbf{g}] = 2$ we have (the complexification of) the Lie algebra of the group of Euclidean motions in the plane. Here we can find a basis h, x, y of $\mathbf{g}$ with brackets given by

$$[h, x] = y, \quad [h, y] = -x, \quad [x, y] = 0.$$

More generally we could start with a commutative two dimensional algebra and adjoin an element h with $\mathrm{ad}\, h$ acting as an arbitrary linear transformation, A of our two dimensional space.

The item of study of this chapter is the algebra $sl(2)$ of **all two by two matrices of trace zero**, where $[\mathbf{g}, \mathbf{g}] = \mathbf{g}$.

2.2 $sl(2)$ and its irreducible representations.

Indeed $sl(2)$ is spanned by the matrices:

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

They satisfy

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

Thus every element of $sl(2)$ can be expressed as a sum of brackets of elements of $sl(2)$, in other words

$$[sl(2), sl(2)] = sl(2).$$

The bracket relations above are also satisfied by the matrices

$$\rho_2(h) := \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \ \rho_2(e) := \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \ \rho_2(f) := \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix},$$

the matrices

$$\rho_3(h) := \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix}, \ \rho_3(e) := \begin{pmatrix} 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \ \rho_3(f) := \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{pmatrix},$$

and, more generally, the $(n+1) \times (n+1)$ matrices given by

$$\rho_n(h) := \begin{pmatrix} n & 0 & \cdots & \cdots & 0 \\ 0 & n-2 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & -n+2 & 0 \\ 0 & 0 & \cdots & \cdots & -n \end{pmatrix}, \ \rho_n(e) = \begin{pmatrix} 0 & n & \cdots & \cdots & 0 \\ 0 & 0 & n-1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & 1 \\ 0 & 0 & \cdots & \cdots & 0 \end{pmatrix},$$

$$\rho_n(f) := \begin{pmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 1 & & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & n & 0 \end{pmatrix}.$$

These representations of $sl(2)$ are all irreducible, as is seen by successively applying $\rho_n(e)$ to any non-zero vector until a vector with non-zero element in the first position and all other entries zero is obtained. Then keep applying $\rho_n(f)$ to fill up the entire space.

These are all the finite dimensional irreducible representations of $sl(2)$ as can be seen as follows: In $U(sl(2))$ we have

$$[h, f^k] = -2kf^k, \quad [h, e^k] = 2ke^k \tag{2.1}$$

$$[e, f^k] = -k(k-1)f^{k-1} + kf^{k-1}h. \tag{2.2}$$

Equation (**??**) follows from the fact that bracketing by any element is a derivation and the fundamental relations in $sl(2)$. Equation (**??**) is proved by induction: For $k = 1$ it is true from the defining relations of $sl(2)$. Assuming it for k, we have

$$\begin{aligned} [e, f^{k+1}] &= [e, f]f^k + f[e, f^k] \\ &= hf^k - k(k-1)f^k + kf^k h \\ &= [h, f^k] + f^k h - k(k-1)f^k + kf^k h \\ &= -2kf^k - k(k-1)f^k + (k+1)f^k h \\ &= -(k+1)kf^k + (k+1)f^k h. \end{aligned}$$

We may rewrite (**??**) as

$$\left[e, \frac{1}{k!}f^k\right] = (-k+1)\frac{1}{(k-1)!}f^{k-1} + \frac{1}{(k-1)!}f^{k-1}h. \tag{2.3}$$

In any finite dimensional module V, the element h has at least one eigenvector. This follows from the fundamental theorem of algebra which assert that any polynomial has at least one root; in particular the characteristic polynomial of any linear transformation on a finite dimensional space has a root. So there is a vector w such that $hw = \mu w$ for some complex number μ. Then

$$h(ew) = [h, e]w + ehw = 2ew + \mu ew = (\mu + 2)(ew).$$

Thus ew is again an eigenvector of h, this time with eigenvalue $\mu+2$. Successively applying e yields a vector v_λ such that

$$hv_\lambda = \lambda v_\lambda, \quad ev_\lambda = 0. \tag{2.4}$$

Then $U(sl(2))v_\lambda$ is an invariant subspace, hence all of V. We say that v is a **cyclic vector** for the action of $\mathbf{g}$ on V if $U(\mathbf{g})v = V$,

We are thus led to study all modules for $sl(2)$ with a cyclic vector v_λ satisfying (**??**). In any such space the elements

$$\frac{1}{k!}f^k v_\lambda$$

span, and are eigenspaces of h of weight $\lambda - 2k$. For any $\lambda \in \mathbf{C}$ we can construct such a module as follows: Let $\mathbf{b}_+$ denote the subalgebra of $sl(2)$ generated by h and e. Then $U(\mathbf{b}_+)$, the universal enveloping algebra of $\mathbf{b}_+$ can be regarded as a subalgebra of $U(sl(2))$. We can make $\mathbf{C}$ into a $\mathbf{b}_+$ module, and hence a $U(\mathbf{b}_+)$ module by

$$h \cdot 1 := \lambda, \quad e \cdot 1 := 0.$$

Then the space

$$U(sl(2)) \otimes_{U(\mathbf{b}^+)} \mathbf{C}$$

with e acting on $\mathbf{C}$ as 0 and h acting via multiplication by λ is a cyclic module with cyclic vector $v_\lambda = 1 \otimes 1$ which satisfies (**??**). It is a "universal" such module in the sense that any other cyclic module with cyclic vector satisfying (**??**) is a homomorphic image of the one we just constructed.

This space $U(sl(2)) \otimes_{U(\mathbf{b}^+)} \mathbf{C}$ is infinite dimensional. It is irreducible unless there is some $\frac{1}{k!}f^k v_\lambda$ with

$$e\left(\frac{1}{k!}f^k v_\lambda\right) = 0$$

where k is an integer ≥ 1. Indeed, any non-zero vector w in the space is a finite linear combination of the basis elements $\frac{1}{k!}f^k v_\lambda$; choose k to be the largest integer so that the coefficient of the corrresponding element does not vanish. Then successive application of the element e (k-times) will yield a multiple of v_λ, and if this multiple is non-zero, then $U(sl(2))w = U(sl(2))v_\lambda$ is the whole space.

But

$$e\left(\frac{1}{k!}f^k v_\lambda\right) = \left[e, \left(\frac{1}{k!}f^k\right)\right] v_\lambda = (1 - k + \lambda) f^{k-1} v_\lambda.$$

This vanishes only if λ is an integer and $k = \lambda + 1$, in which case there is a unique finite dimensional quotient of dimension $k + 1$. QED

The finite dimensional irreducible representations having zero as a weight are all odd dimensional and have only even weights. We will call them "even". They are called "integer spin" representations by the physicists. The others are "odd" or "half spin" representations.

2.3 The Casimir element.

In $U(sl(2))$ consider the element

$$C := \frac{1}{2}h^2 + ef + fe \tag{2.5}$$

called the **Casimir element** or simply the "Casimir" of $sl(2)$.

Since $ef = fe + [e, f] = fe + h$ in $U(sl(2))$ we also can write

$$C = \frac{1}{2}h^2 + h + 2fe. \tag{2.6}$$

This implies that if v is a "highest weight vector" in a $sl(2)$ module satisfying $ev = 0,\ hv = \lambda v$ then

$$Cv = \frac{1}{2}\lambda(\lambda + 2)v. \tag{2.7}$$

Now in $U(sl(2))$ we have

$$\begin{aligned} [h, C] &= 2([h, f]e + f[h, e]) \\ &= 2(-2fe + 2fe) \\ &= 0 \end{aligned}$$

and

$$\begin{aligned} [C, e] &= \frac{1}{2} \cdot 2(eh + he) + 2e - 2he \\ &= eh - he + 2e \\ &= -[h, e] + 2e \\ &= 0. \end{aligned}$$

Similarly

$$[C, f] = 0.$$

In other words, C lies in the **center** of the universal enveloping algebra of $sl(2)$, i.e. it commutes with all elements. If V is a module which possesses a "highest weight vector" v_λ as above, and if V has the property that v_λ is a cyclic vector, meaning that $V = U(L)v_\lambda$ then C takes on the constant value

$$C = \frac{\lambda(\lambda + 2)}{2}\mathrm{Id}$$

since C is central and v_λ is cyclic.

2.4 $sl(2)$ is simple.

An ideal I in a Lie algebra $\mathbf{g}$ is a subspace of $\mathbf{g}$ which is invariant under the adjoint representation. In other words, I is an ideal if $[\mathbf{g}, I] \subset I$. If a Lie algebra $\mathbf{g}$ has the property that its only ideals are 0 and $\mathbf{g}$ itself, and if $\mathbf{g}$ is not commutative, we say that $\mathbf{g}$ is **simple**. Let us prove that $sl(2)$ is simple. Since $sl(2)$ is not commutative, we must prove that the only ideals are 0 and $sl(2)$ itself. We do this by introducing some notation which will allow us to generalize the proof in the next chapter. Let

$$\mathbf{g} = sl(2)$$

and set

$$\mathbf{g}_{-1} := \mathbf{C}f, \quad \mathbf{g}_0 := \mathbf{C}h, \quad \mathbf{g}_1 := \mathbf{C}e$$

so that $\mathbf{g}$, as a vector space, is the direct sum of the three one dimensional spaces

$$\mathbf{g} = \mathbf{g}_{-1} \oplus \mathbf{g}_0 \oplus \mathbf{g}_1.$$

Correspondingly, write any $x \in \mathbf{g}$ as

$$x = x_{-1} + x_0 + x_1.$$

If we let

$$d := \frac{1}{2}h$$

then we have

$$\begin{aligned} x &= x_{-1} + x_0 + x_1, \\ [d, x] &= -x_{-1} + 0 + x_1, \text{ and} \\ [d, [d, x]] &= x_{-1} + 0 + x_1. \end{aligned}$$

Since the matrix

$$\begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

is invertible, we see that we can solve for the "components" x_{-1}, x_0 and x_1 in terms of x, $[d, x]$, $[d, [d, x]]$. This means that if I is an ideal, then

$$I = I_1 \oplus I_0 \oplus I_1$$

where

$$I_{-1} := I \cap \mathbf{g}_{-1}, \ I_0 := I \cap \mathbf{g}_0, \ I_1 := I \cap \mathbf{g}_1.$$

Now if $I_0 \neq 0$ then $d = \frac{1}{2}h \in I$, and hence $e = [d, e]$ and $f = -[d, f]$ also belong to I so $I = sl(2)$. If $I_{-1} \neq 0$ so that $f \in I$, then $h = [e, f] \in I$ so $I = sl(2)$. Similarly, if $I_1 \neq 0$ so that $e \in I$ then $h = [e, f] \in I$ so $I = sl(2)$.

Thus if $I \neq 0$ then $I = sl(2)$ and we have proved that $sl(2)$ is simple.

2.5 Complete reducibility.

We will use the Casimir element C to prove that every finite dimensional representation W of $sl(2)$ is **completely reducible**, which means that if W' is an invariant subspace there exists a complementary invariant subspace W'' so that $W = W' \oplus W''$. Indeed we will prove:

Theorem 1
1. *Every finite dimensional representation of* $sl(2)$ *is completely reducible.*
2. *Each irreducible subspace is a cyclic highest weight module with highest weight* n *where* n *is a non-negative integer.*
3. *When the representation is decomposed into a direct sum of irreducible components, the number of components with even highest weight is the multiplicity of* 0 *as an an eigenvector of* h *and*
4. *the number of components with odd highest weight is the multiplicity of* 1 *as an eigenvalue of* h.

Proof. We know that every irreducible finite dimensional representation is a cyclic module with integer highest weight, that those with even highest weight contain 0 as an eigenvalue of h with multiplicity one and do not contain 1 as an eigenvalue of h, and that those with odd highest weight contain 1 as an eigenvalue of h with multiplicity one, and do not contain 0 as an eigenvalue. So 2), 3) and 4) follow from 1). We must prove 1).

We first prove

Proposition 1 *Let* $0 \to V \to W \to k \to 0$ *be an exact sequence of* $sl(2)$ *modules and such that the action of* $sl(2)$ *on* k *is trivial (as it must be, since* $sl(2)$ *has no non-trivial one dimensional modules). Then this sequence splits, i.e. there is a line in* W *supplementary to* V *on which* $sl(2)$ *acts trivially.*

This proposition is, of course, a special case of the theorem we want to prove. But we shall see that it is sufficient to prove the theorem.

Proof of proposition. It is enough to prove the proposition for the case that V is an irreducible module. Indeed, if V_1 is a submodule, then by induction on dim V we may assume the theorem is known for $0 \to V/V_1 \to W/V_1 \to k \to 0$ so that there is a one dimensional invariant subspace M in W/V_1 supplementary to V/V_1 on which the action is trivial. Let N be the inverse image of M in W. By another application of the proposition, this time to the sequence

$$0 \to V_1 \to N \to M \to 0$$

we find an invariant line, P, in N complementary to V_1. So $N = V_1 \oplus P$. Since $(W/V_1) = (V/V_1) \oplus M$ we must have $P \cap V = \{0\}$. But since dim $W =$ dim $V + 1$, we must have $W = V \oplus P$. In other words P is a one dimensional subspace of W which is complementary to V.

Next we are reduced to proving the proposition for the case that $sl(2)$ acts faithfully on V. Indeed, let $I =$ the kernel of the action on V. Since $sl(2)$ is simple, either $I = sl(2)$ or $I = 0$. Suppose that $I = sl(2)$. For all $x \in sl(2)$ we have, by hypothesis, $xW \subset V$, and for $x \in I = sl(2)$ we have $xV = 0$. Hence

$$[sl(2), sl(2)] = sl(2)$$

acts trivially on all of W and the proposition is obvious. So we are reduced to the case that V is irreducible and the action, ρ, of $sl(2)$ on V is injective. We have our Casimir element C whose image in $\operatorname{End} W$ must map $W \to V$ since every element of $sl(2)$ does. On the other hand, $C = \frac{1}{2}n(n+2)\,\mathrm{Id} \neq 0$ since we are assuming that the action of $sl(2)$ on the irreducible module V is not trivial. In particular, the restriction of C to V is an isomorphism. Hence $\ker C_\rho : W \to V$ is an invariant line supplementary to V. We have proved the proposition.

Proof of theorem from proposition. Let $0 \to E' \to E$ be an exact sequence of $sl(2)$ modules, and we may assume that $E' \neq 0$. We want to find an invariant complement to E' in E. Define W to be the subspace of $\mathrm{Hom}_k(E, E')$ whose restriction to E' is a scalar times the identity, and let $V \subset W$ be the subspace consisting of those linear transformations whose restrictions to E' is zero. Each of these is a submodule of $\mathrm{End}(E)$. We get a sequence

$$0 \to V \to W \to k \to 0$$

and hence a complementary line of invariant elements in W. In particular, we can find an element, T which is invariant, maps $E \to E'$, and whose restriction to E' is non-zero. Then $\ker T$ is an invariant complementary subspace. QED

2.6 The Weyl group.

We have

$$\exp e = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \exp -f = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$$

so

$$(\exp e)(\exp -f)(\exp e) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Since

$$\exp \operatorname{ad} x = Ad(\exp x)$$

we see that

$$\tau := (\exp \operatorname{ad} e)(\exp \operatorname{ad}(-f))(\exp \operatorname{ad} e)$$

consists of conjugation by the matrix

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Thus

$$\tau(h) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = -h,$$

$$\tau(e) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} = -f$$

and similarly $\tau(f) = -e$. In short

$$\tau : e \mapsto -f,\ f \mapsto -e,\ h \mapsto -h.$$

In particular, τ induces the "reflection" $h \mapsto -h$ on $\mathbf{C}h$ and hence the reflection $\mu \mapsto -\mu$ (which we shall also denote by s) on the (one dimensional) dual space. In any finite dimensional module V of $sl(2)$ the action of the element $\tau = (\exp e)(\exp -f)(\exp e)$ is defined, and

$$(\tau)^{-1}h(\tau) = \mathrm{Ad}\left(\tau^{-1}\right)(h) = s^{-1}h = sh$$

so if

$$hu = \mu u$$

then

$$h(\tau u) = \tau(\tau)^{-1}h(\tau)u = \tau s(h)u = -\mu\tau u = (s\mu)\tau u.$$

So if

$$V_\mu : \{u \in V | hu = \mu u\}$$

then

$$\tau(V_\mu) = V_{s\mu}. \tag{2.8}$$

The two element group consisting of the identity and the element s (acting as a reflection as above) is called the Weyl group of $sl(2)$. Its generalization to an arbitrary simple Lie algebra, together with the generalization of formula (**??**) will play a key role in what follows.