# *Chapter II*

# Semisimple Lie Algebras

In Chapter I we looked at Lie algebras over an arbitrary field F. Apart from introducing the basic notions and examples, we were able to prove only one substantial theorem (Engel's Theorem). Virtually all of the remaining theory to be developed in this book will require the assumption that F have characteristic 0. (Some of the exercises will indicate how counterexamples arise in prime characteristic.) Moreover, in order to have available the eigenvalues of ad $x$ for arbitrary $x$ (not just for ad $x$ nilpotent), we shall assume that F is algebraically closed, except where otherwise specified. It is possible to work with a slightly less restrictive assumption on F (cf. Jacobson [1], p. 107), but we shall not do so here.

## 4. Theorems of Lie and Cartan

### *4.1. Lie's Theorem*

The essence of Engel's Theorem for nilpotent Lie algebras is the existence of a common eigenvector for a Lie algebra consisting of nilpotent endomorphisms (Theorem 3.3). The next theorem is similar in nature, but requires algebraic closure, in order to assure that F will contain all required eigenvalues. It turns out to be necessary also to have char F $=$ 0 (Exercise 3).

**Theorem.** *Let $L$ be a solvable subalgebra of $\mathfrak{gl}(V)$, $V$ finite dimensional. If $V \neq 0$, then $V$ contains a common eigenvector for all the endomorphisms in $L$.*

*Proof.* Use induction on dim $L$, the case dim $L = 0$ being trivial. We attempt to imitate the proof of Theorem 3.3 (which the reader should review at this point). The idea is (1) to locate an ideal $K$ of codimension one, (2) to show by induction that common eigenvectors exist for $K$, (3) to verify that $L$ stabilizes a space consisting of such eigenvectors, and (4) to find in that space an eigenvector for a single $z \in L$ satisfying $L = K + \mathsf{F}z$.

Step (1) is easy. Since $L$ is solvable, of positive dimension, $L$ properly includes $[LL]$. $L/[LL]$ being abelian, any subspace is automatically an ideal. Take a subspace of codimension one, then its inverse image $K$ is an ideal of codimension one in $L$ (including $[LL]$).

For step (2), use induction to find a common eigenvector $v \in V$ for $K$ ($K$ is of course solvable; if $K = 0$, then $L$ is abelian of dimension 1 and an eigenvector for a basis vector of $L$ finishes the proof.) This means that for $x \in K$

$x.v = \lambda(x)v$, $\lambda: K \to \mathsf{F}$ some linear function. Fix this $\lambda$, and denote by $W$ the subspace

$$\{w \in V | x.w = \lambda(x)w, \text{ for all } x \in K\}; \text{ so } W \neq 0.$$

Step (3) consists in showing that $L$ leaves $W$ invariant. Assuming for the moment that this is done, proceed to step (4): Write $L = K + \mathsf{F}z$, and use the fact that $\mathsf{F}$ is algebraically closed to find an eigenvector $v_0 \in W$ of $z$ (for some eigenvalue of $z$). Then $v_0$ is obviously a common eigenvector for $L$ (and $\lambda$ can be extended to a linear function on $L$ such that $x.v_0 = \lambda(x)v_0$, $x \in L$).

It remains to show that $L$ stabilizes $W$. Let $w \in W$, $x \in L$. To test whether or not $x.w$ lies in $W$, we must take arbitrary $y \in K$ and examine $yx.w = xy.w - [x, y].w = \lambda(y)x.w - \lambda([x, y])w$. Thus *we have to prove that* $\lambda([x, y]) = 0$. For this, fix $w \in W$, $x \in L$. Let $n > 0$ be the smallest integer for which $w$, $x.w, \ldots, x^n.w$ are linearly dependent. Let $W_i$ be the subspace of $V$ spanned by $w, x.w, \ldots, x^{i-1}.w$ (set $W_0 = 0$), so $\dim W_n = n$, $W_n = W_{n+i}$ $(i \geq 0)$ and $x$ maps $W_n$ into $W_n$. It is easy to check that each $y \in K$ leaves each $W_i$ invariant. Relative to the basis $w, x.w, \ldots, x^{n-1}.w$ of $W_n$, we claim that $y \in K$ is represented by an upper triangular matrix whose diagonal entries equal $\lambda(y)$. This follows immediately from the congruence:

$$(*) \qquad yx^i.w \equiv \lambda(y)x^i.w \pmod{W_i},$$

which we prove by induction on $i$, the case $i = 0$ being obvious. Write $yx^i.w = yxx^{i-1}.w = xyx^{i-1}.w - [x, y]x^{i-1}.w$. By induction, $yx^{i-1}.w = \lambda(y)x^{i-1}.w + w'$ $(w' \in W_{i-1})$; since $x$ maps $W_{i-1}$ into $W_i$ (by construction), (*) therefore holds for all $i$.

According to our description of the way in which $y \in K$ acts on $W_n$, $Tr_{W_n}(y) = n\lambda(y)$. In particular, this is true for elements of $K$ of the special form $[x, y]$ ($x$ as above, $y$ in $K$). But $x$, $y$ both stabilize $W_n$, so $[x, y]$ acts on $W_n$ as the *commutator* of two endomorphisms of $W_n$; its trace is therefore 0. We conclude that $n\lambda([x, y]) = 0$. Since char $\mathsf{F} = 0$, this forces $\lambda([x, y]) = 0$, as required. □

**Corollary A (Lie's Theorem).** *Let L be a solvable subalgebra of* $\mathfrak{gl}(V)$, $\dim V = n < \infty$. *Then L stabilizes some flag in V (in other words, the matrices of L relative to a suitable basis of V are upper triangular).*

*Proof.* Use the theorem, along with induction on dim $V$. □

More generally, let $L$ be any solvable Lie algebra, $\phi: L \to \mathfrak{gl}(V)$ a finite dimensional representation of $L$. Then $\phi(L)$ is solvable, by Proposition 3.1(a), hence stabilizes a flag (Corollary A). For example, if $\phi$ is the adjoint representation, a flag of subspaces stable under $L$ is just a chain of ideals of $L$, each of codimension one in the next. This proves:

**Corollary B.** *Let L be solvable. Then there exists a chain of ideals of L,* $0 = L_0 \subset L_1 \subset \ldots \subset L_n = L$, *such that* $\dim L_i = i$. □

**Corollary C.** *Let L be solvable. Then* $x \in [LL]$ *implies that* $ad_L\, x$ *is nilpotent. In particular,* $[LL]$ *is nilpotent.*

*Proof.* Find a flag of ideals as in Corollary B. Relative to a basis $(x_1, \ldots, x_n)$ of $L$ for which $(x_1, \ldots, x_i)$ spans $L_i$, the matrices of ad $L$ lie in $\mathfrak{t}(n, \mathsf{F})$. Therefore the matrices of $[\text{ad } L, \text{ad } L] = \text{ad}_L [LL]$ lie in $\mathfrak{n}(n, \mathsf{F})$, the derived algebra of $\mathfrak{t}(n, \mathsf{F})$. It follows that $\text{ad}_L\, x$ is nilpotent for $x \in [LL]$; a fortiori $\text{ad}_{[LL]}\, x$ is nilpotent, so $[LL]$ is nilpotent by Engel's Theorem. ▯

### *4.2. Jordan-Chevalley decomposition*

*In this subsection only, char* $\mathsf{F}$ *may be arbitrary.* We digress in order to introduce a very useful tool for the study of linear transformations. The reader may recall that the Jordan canonical form for a single endomorphism $x$ over an algebraically closed field amounts to an expression of $x$ in matrix form as a sum of blocks

$$\begin{bmatrix} a & 1 & & & 0 \\ & a & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ 0 & & & & a \end{bmatrix}$$

Since diag $(a, \ldots, a)$ commutes with the nilpotent matrix having one's just above the diagonal and zeros elsewhere, $x$ is the sum of a diagonal and a nilpotent matrix which commute. We can make this decomposition more precise, as follows.

Call $x \in \text{End } V$ ($V$ finite dimensional) **semisimple** if the roots of its minimal polynomial over $\mathsf{F}$ are all distinct. Equivalently ($\mathsf{F}$ being algebraically closed), $x$ is semisimple if and only if $x$ is diagonalizable. We remark that two commuting semisimple endomorphisms can be simultaneously diagonalized; therefore, their sum or difference is again semisimple (Exercise 2). Also, if $x$ is semisimple and maps a subspace $W$ of $V$ into itself, then obviously the restriction of $x$ to $W$ is semisimple.

**Proposition.** *Let $V$ be a finite dimensional vector space over* $\mathsf{F}$, *$x \in \text{End } V$.*

(*a*) *There exist unique $x_s, x_n \in \text{End } V$ satisfying the conditions: $x = x_s + x_n$, $x_s$ is semisimple, $x_n$ is nilpotent, $x_s$ and $x_n$ commute.*

(*b*) *There exist polynomials $p(T)$, $q(T)$ in one indeterminate, without constant term, such that $x_s = p(x)$, $x_n = q(x)$. In particular, $x_s$ and $x_n$ commute with any endomorphism commuting with $x$.*

(*c*) *If $A \subset B \subset V$ are subspaces, and $x$ maps $B$ into $A$, then $x_s$ and $x_n$ also map $B$ into $A$.*

The decomposition $x = x_s + x_n$ is called the (additive) **Jordan-Chevalley decomposition** of $x$, or just the Jordan decomposition; $x_s$, $x_n$ are called (respectively) the **semisimple part** and the **nilpotent part** of $x$.

*Proof.* Let $a_1, \ldots, a_k$ (with multiplicities $m_1, \ldots, m_k$) be the distinct eigenvalues of $x$, so the characteristic polynomial is $\Pi(T-a_i)^{m_i}$. If $V_i = \text{Ker } (x-a_i \cdot 1)^{m_i}$, then $V$ is the direct sum of the subspaces $V_1, \ldots, V_k$, each stable under $x$. On $V_i$, $x$ clearly has characteristic polynomial $(T-a_i)^{m_i}$. Now apply the Chinese Remainder Theorem (for the ring $\mathsf{F}[T]$) to locate a polynomial $p(T)$ satisfying the congruences, with pairwise relatively prime moduli: $p(T) \equiv a_i \pmod{(T-a_i)^{m_i}}$, $p(T) \equiv 0 \pmod{T}$. (Notice that the last congruence is superfluous if 0 is an eigenvalue of $x$, while otherwise $T$ is relatively prime to the other moduli.) Set $q(T) = T-p(T)$. Evidently each of $p(T)$, $q(T)$ has zero constant term, since $p(T) \equiv 0 \pmod{T}$.

Set $x_s = p(x)$, $x_n = q(x)$. Since they are polynomials in $x$, $x_s$ and $x_n$ commute with each other, as well as with all endomorphisms which commute with $x$. They also stabilize all subspaces of $V$ stabilized by $x$, in particular the $V_i$. The congruence $p(T) \equiv a_i \pmod{(T-a_i)^{m_i}}$ shows that the restriction of $x_s - a_i \cdot 1$ to $V_i$ is zero for all $i$, hence that $x_s$ acts diagonally on $V_i$ with single eigenvalue $a_i$. By definition, $x_n = x - x_s$, which makes it clear that $x_n$ is nilpotent. Because $p(T)$, $q(T)$ have no constant term, (c) is also obvious at this point.

It remains only to prove the uniqueness assertion in (a). Let $x = s+n$ be another such decomposition, so we have $x_s - s = n - x_n$. Because of (b), all endomorphisms in sight commute. Sums of commuting semisimple (resp. nilpotent) endomorphisms are again semisimple (resp. nilpotent), whereas only 0 can be both semisimple and nilpotent. This forces $s = x_s$, $n = x_n$. ☐

To indicate why the Jordan decomposition will be a valuable tool, we look at a special case. Consider the adjoint representation of the Lie algebra $\mathfrak{gl}(V)$, $V$ finite dimensional. If $x \in \mathfrak{gl}(V)$ is nilpotent, then so is ad $x$ (Lemma 3.2). Similarly, *if $x$ is semisimple, then so is ad $x$*. We verify this as follows. Choose a basis $(v_1, \ldots, v_n)$ of $V$ relative to which $x$ has matrix diag $(a_1, \ldots, a_n)$. Let $\{e_{ij}\}$ be the standard basis of $\mathfrak{gl}(V)$ (1.2) relative to $(v_1, \ldots, v_n)$: $e_{ij}(v_k) = \delta_{jk} v_i$. Then a quick calculation (see formula (*) in (1.2)) shows that ad $x\,(e_{ij}) = (a_i - a_j)e_{ij}$. So ad $x$ has diagonal matrix, relative to the chosen basis of $\mathfrak{gl}(V)$.

**Lemma A.** *Let $x \in$ End $V$ (dim $V < \infty$), $x = x_s + x_n$ its Jordan decomposition. Then ad $x$ = ad $x_s$ + ad $x_n$ is the Jordan decomposition of ad $x$ (in End (End $V$)).*

*Proof.* We have seen that ad $x_s$, ad $x_n$ are respectively semisimple, nilpotent; they commute, since [ad $x_s$, ad $x_n$] = ad $[x_s, x_n] = 0$. Then part (a) of the proposition applies. ☐

A further useful fact is the following.

**Lemma B.** *Let $\mathfrak{A}$ be a finite dimensional $\mathsf{F}$-algebra. Then Der $\mathfrak{A}$ contains the semisimple and nilpotent parts (in End $\mathfrak{A}$) of all its elements.*

*Proof.* If $\delta \in$ Der $\mathfrak{A}$, let $\sigma$, $\nu \in$ End $\mathfrak{A}$ be its semisimple and nilpotent parts, respectively. It will be enough to show that $\sigma \in$ Der $\mathfrak{A}$. If $a \in \mathsf{F}$, set $\mathfrak{A}_a = \{x \in \mathfrak{A} \mid (\delta - a.1)^k x = 0$ for some $k$ (depending on $x$)$\}$. Then $\mathfrak{A}$ is the direct sum of those $\mathfrak{A}_a$ for which $a$ is an eigenvalue of $\delta$ (or $\sigma$), and $\sigma$ acts on $\mathfrak{A}_a$ as scalar multiplication by $a$. We can verify, for arbitrary $a, b \in \mathsf{F}$, that $\mathfrak{A}_a \mathfrak{A}_b \subset \mathfrak{A}_{a+b}$, by means of the general formula: (*) $(\delta-(a+b).1)^n(xy)$

$= \sum_{i=0}^{n} \binom{n}{i} ((\delta - a.1)^{n-i}x)\cdot((\delta - b.1)^i y)$, for $x, y \in \mathfrak{A}$. (This formula is easily checked by induction on $n$.) Now if $x \in \mathfrak{A}_a$, $y \in \mathfrak{A}_b$, then $\sigma(xy) = (a+b)xy$, because $xy \in \mathfrak{A}_{a+b}$ (possibly equal to 0); on the other hand, $(\sigma x)y + x(\sigma y) = (a+b)xy$. By directness of the sum $\mathfrak{A} = \coprod \mathfrak{A}_a$, it follows that $\sigma$ is a derivation, as required. □

### *4.3. Cartan's Criterion*

We are now ready to obtain a powerful criterion for solvability of a Lie algebra $L$, based on the traces of certain endomorphisms of $L$. It is obvious that $L$ will be solvable if $[LL]$ is nilpotent (this is the converse of Corollary 4.1C). In turn, Engel's Theorem says that $[LL]$ will be nilpotent if (and only if) each $\text{ad}_{[LL]}x$, $x \in [LL]$, is nilpotent. We begin, therefore, with a "trace" criterion for nilpotence of an endomorphism.

**Lemma.** *Let $A \subset B$ be two subspaces of $\mathfrak{gl}(V)$, dim $V < \infty$. Set $M = \{x \in \mathfrak{gl}(V) | [x, B] \subset A\}$. Suppose $x \in M$ satisfies $Tr(xy) = 0$ for all $y \in M$. Then $x$ is nilpotent.*

*Proof.* Let $x = s+n$ ($s = x_s$, $n = x_n$) be the Jordan decomposition of $x$. Fix a basis $v_1, \ldots, v_m$ of $V$ relative to which $s$ has matrix diag $(a_1, \ldots, a_m)$. Let $E$ be the vector subspace of $\mathsf{F}$ (over the prime field $\mathbf{Q}$) spanned by the eigenvalues $a_1, \ldots, a_m$. We have to show that $s = 0$, or equivalently, that $E = 0$. Since $E$ has finite dimension over $\mathbf{Q}$ (by construction), it will suffice to show that the dual space $E^*$ is 0, i.e., that any linear function $f$: $E \to \mathbf{Q}$ is zero.

Given $f$, let $y$ be that element of $\mathfrak{gl}(V)$ whose matrix relative to our given basis is diag $(f(a_1), \ldots, f(a_m))$. If $\{e_{ij}\}$ is the corresponding basis of $\mathfrak{gl}(V)$, we saw in (4.2) that: ad $s(e_{ij}) = (a_i - a_j)e_{ij}$, ad $y(e_{ij}) = (f(a_i) - f(a_j))e_{ij}$. Now let $r(T) \in \mathsf{F}[T]$ be a polynomial without constant term satisfying $r(a_i - a_j) = f(a_i) - f(a_j)$ for all pairs $i, j$. The existence of such $r(T)$ follows from Lagrange interpolation; there is no ambiguity in the assigned values, since $a_i - a_j = a_k - a_l$ implies (by linearity of $f$) that $f(a_i) - f(a_j) = f(a_k) - f(a_l)$. Evidently ad $y = r$ (ad $s$).

Now ad $s$ is the semisimple part of ad $x$, by Lemma A of (4.2), so it can be written as a polynomial in ad $x$ without constant term (Proposition 4.2). Therefore, ad $y$ is also a polynomial in ad $x$ without constant term. By hypothesis, ad $x$ maps $B$ into $A$, so we also have ad $y$ $(B) \subset A$, i.e., $y \in M$. Using the hypothesis of the lemma, $Tr(xy) = 0$, we get $\Sigma a_i f(a_i) = 0$. The left side is a $\mathbf{Q}$-linear combination of elements of $E$; applying $f$, we obtain $\Sigma f(a_i)^2 = 0$. But the numbers $f(a_i)$ are rational, so this forces all of them to be 0. Finally, $f$ must be identically 0, because the $a_i$ span $E$. □

Before stating our solvability criterion, we record a useful identity: If $x$, $y$, $z$ are endomorphisms of a finite dimensional vector space, then (*) $Tr([x, y]z) = Tr(x[y, z])$. To verify this, write $[x, y]z = xyz - yxz$, $x[y, z] = xyz - xzy$, and use the fact that $Tr(y(xz)) = Tr((xz)y)$.

**Theorem (Cartan's Criterion).** *Let $L$ be a subalgebra of $\mathfrak{gl}(V)$, $V$ finite dimensional. Suppose that $Tr(xy) = 0$ for all $x \in [LL]$, $y \in L$. Then $L$ is solvable.*

*Proof.* As remarked at the beginning of (4.3), it will suffice to prove that $[LL]$ is nilpotent, or just that all $x$ in $[LL]$ are nilpotent endomorphisms (Lemma 3.2 and Engel's Theorem). For this we apply the above lemma to the situation: $V$ as given, $A = [LL]$, $B = L$, so $M = \{x \in \mathfrak{gl}(V) | [x, L] \subset [LL]\}$. Obviously $L \subset M$. Our hypothesis is that $Tr(xy) = 0$ for $x \in [LL]$, $y \in L$, whereas to conclude from the lemma that each $x \in [LL]$ is nilpotent we need the stronger statement: $Tr(xy) = 0$ for $x \in [LL]$, $y \in M$.

Now if $[x, y]$ is a typical generator of $[LL]$, and if $z \in M$, then identity (*) above shows that $Tr([x, y]z) = Tr(x[y, z]) = Tr([y, z]x)$. By definition of $M$, $[y, z] \in [LL]$, so the right side is 0 by hypothesis. ▯

**Corollary.** *Let $L$ be a Lie algebra such that $Tr(\text{ad } x \text{ ad } y) = 0$ for all $x \in [LL]$, $y \in L$. Then $L$ is solvable.*

*Proof.* Applying the theorem to the adjoint representation of $L$, we get ad $L$ solvable. Since Ker ad $= Z(L)$ is solvable, $L$ itself is solvable (Proposition 3.1). ▯

## Exercises

1. Let $L = \mathfrak{sl}(V)$. Use Lie's Theorem to prove that Rad $L = Z(L)$; conclude that $L$ is semisimple (cf. Exercise 2.3). [Observe that Rad $L$ lies in each maximal solvable subalgebra $B$ of $L$. Select a basis of $V$ so that $B = L \cap \mathfrak{t}(n, \mathsf{F})$, and notice that the transpose of $B$ is also a maximal solvable subalgebra of $L$. Conclude that Rad $L \subset L \cap \mathfrak{d}(n, \mathsf{F})$, then that Rad $L = Z(L)$.]
2. Show that the proof of Theorem 4.1 still goes through in prime characteristic, provided dim $V$ is less than char $\mathsf{F}$.
3. This exercise illustrates the failure of Lie's Theorem when $\mathsf{F}$ is allowed to have prime characteristic $p$. Consider the $p \times p$ matrices:

$$x = \begin{bmatrix} 0 & 1 & 0 & . & . & 0 \\ 0 & 0 & 1 & 0 & . & 0 \\ . & . & . & . & . & . \\ 0 & . & . & . & . & 1 \\ 1 & . & . & . & . & 0 \end{bmatrix}, \quad y = \text{diag}(0, 1, 2, 3, \ldots, p-1).$$

   Check that $[x, y] = x$, hence that $x$ and $y$ span a two dimensional solvable subalgebra $L$ of $\mathfrak{gl}(p, \mathsf{F})$. Verify that $x$, $y$ have no common eigenvector.
4. When $p = 2$, Exercise 3.3 shows that a solvable Lie algebra of endomorphisms over a field of prime characteristic $p$ need not have derived

algebra consisting of nilpotent endomorphisms (cf. Corollary C of Theorem 4.1). For arbitrary $p$, construct a counterexample to Corollary C as follows: Start with $L \subset \mathfrak{gl}(p, \mathsf{F})$ as in Exercise 3. Form the vector space direct sum $M = L + \mathsf{F}^p$, and make $M$ a Lie algebra by decreeing that $\mathsf{F}^p$ is abelian, while $L$ has its usual product and acts on $\mathsf{F}^p$ in the given way. Verify that $M$ is solvable, but that its derived algebra ($= \mathsf{F}x + \mathsf{F}^p$) fails to be nilpotent.

5. If $x, y \in \text{End } V$ commute, prove that $(x+y)_s = x_s + y_s$, and $(x+y)_n = x_n + y_n$. Show by example that this can fail if $x, y$ fail to commute. [Show first that $x, y$ semisimple (resp. nilpotent) implies $x+y$ semisimple (resp. nilpotent).]
6. Check formula (*) at the end of (4.2).
7. Prove the converse of Theorem 4.3.
8. Note that it suffices to check the hypothesis of Theorem 4.3 (or its corollary) for $x, y$ ranging over a basis of $L$. For the example given in Exercise 1.2, verify solvability by using Cartan's Criterion.

## Notes

The proofs here follow Serre [1]. The systematic use of the Jordan decomposition in linear algebraic groups originates with Chevalley [1]; see also Borel [1], where the additive Jordan decomposition in the Lie algebra is emphasized.

# 5. Killing form

### *5.1. Criterion for semisimplicity*

Let $L$ be any Lie algebra. If $x, y \in L$, define $\kappa(x, y) = Tr(\text{ad } x \text{ ad } y)$. Then $\kappa$ is a symmetric bilinear form on $L$, called the **Killing form**. $\kappa$ is also **associative**, in the sense that $\kappa([xy], z) = \kappa(x, [yz])$. This follows from the identity recorded in (4.3): $Tr([x, y]z) = Tr(x[y, z])$, for endomorphisms $x, y, z$ of a finite dimensional vector space.

The following lemma will be handy later on.

**Lemma.** *Let $I$ be an ideal of $L$. If $\kappa$ is the Killing form of $L$ and $\kappa_I$ the Killing form of $I$ (viewed as Lie algebra), then $\kappa_I = \kappa|_{I \times I}$.*

*Proof.* First, a simple fact from linear algebra: If $W$ is a subspace of a (finite dimensional) vector space $V$, and $\phi$ an endomorphism of $V$ mapping $V$ into $W$, then $Tr\phi = Tr(\phi|_W)$. (To see this, extend a basis of $W$ to a basis of $V$ and look at the resulting matrix of $\phi$.) Now if $x, y \in I$, then (ad $x$) (ad $y$) is an endomorphism of $L$, mapping $L$ into $I$, so its trace $\kappa(x, y)$ coincides with the trace $\kappa_I(x, y)$ of $(\text{ad } x)(\text{ad } y)|_I = (\text{ad}_I\, x)(\text{ad}_I\, y)$. $\square$

In general, a symmetric bilinear form $\beta(x, y)$ is called **nondegenerate** if its **radical** $S$ is 0, where $S = \{x \in L | \beta(x, y) = 0 \text{ for all } y \in L\}$. Because the Killing form is associative, its radical is more than just a subspace: $S$ is an *ideal* of $L$. From linear algebra, a practical way to test nondegeneracy is as follows: Fix a basis $x_1, \ldots, x_n$ of $L$. Then $\kappa$ is nondegenerate if and only if the $n \times n$ matrix whose $i, j$ entry is $\kappa(x_i, x_j)$ has nonzero determinant.

As an example, we compute the Killing form of $\mathfrak{sl}(2, \mathsf{F})$, using the standard basis (Example 2.1), which we write in the order $(x, h, y)$. The matrices become:

$$\text{ad } h = \text{diag}(2, 0, -2), \text{ ad } x = \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \text{ ad } y = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}.$$

Therefore $\kappa$ has matrix $\begin{pmatrix} 0 & 0 & 4 \\ 0 & 8 & 0 \\ 4 & 0 & 0 \end{pmatrix}$, with determinant $-128$, and $\kappa$ is nondegenerate. (This is still true so long as char $\mathsf{F} \neq 2$.)

Recall that a Lie algebra $L$ is called semisimple in case $\text{Rad} L = 0$. This is equivalent to requiring that $L$ have no nonzero abelian ideals: indeed, any such ideal must be in the radical, and conversely, the radical (if nonzero) includes such an ideal of $L$, viz., the last nonzero term in the derived series of $\text{Rad} L$ (cf. exercise 3.1).

**Theorem.** *Let L be a Lie algebra. Then L is semisimple if and only if its Killing form is nondegenerate.*

*Proof.* Suppose first that $\text{Rad } L = 0$. Let $S$ be the radical of $\kappa$. By definition, $Tr(\text{ad } x \text{ ad } y) = 0$ for all $x \in S$, $y \in L$ (in particular, for $y \in [SS]$). According to Cartan's Criterion (4.3), $\text{ad}_L S$ is solvable, hence $S$ is solvable. But we remarked above that $S$ is an ideal of $L$, so $S \subset \text{Rad } L = 0$, and $\kappa$ is nondegenerate.

Conversely, let $S = 0$. To prove that $L$ is semisimple, it will suffice to prove that every abelian ideal $I$ of $L$ is included in $S$. Suppose $x \in I$, $y \in L$. Then ad $x$ ad $y$ maps $L \rightarrow L \rightarrow I$, and $(\text{ad } x \text{ ad } y)^2$ maps $L$ into $[II] = 0$. This means that ad $x$ ad $y$ is nilpotent, hence that $0 = Tr(\text{ad } x \text{ ad } y) = \kappa(x, y)$, so $I \subset S = 0$. (This half of the proof remains valid even in prime characteristic (Exercise 6).) ▯

The proof shows that we always have $S \subset \text{Rad } L$; however, the reverse inclusion need not hold (Exercise 4).

### 5.2. *Simple ideals of L*

First a definition. A Lie algebra $L$ is said to be the **direct sum** of ideals $I_1, \ldots, I_t$ provided $L = I_1 + \ldots + I_t$ (direct sum of subspaces). This condition forces $[I_i I_j] \subset I_i \cap I_j = 0$ if $i \neq j$ (so the algebra $L$ can be viewed as gotten from the Lie algebras $I_i$ by defining Lie products componentwise for the external direct sum of these as vector spaces). We write $L = I_1 \oplus \ldots \oplus I_t$.

**Theorem.** *Let $L$ be semisimple. Then there exist ideals $L_1, \ldots, L_t$ of $L$ which are simple (as Lie algebras), such that $L = L_1 \oplus \ldots \oplus L_t$. Every simple ideal of $L$ coincides with one of the $L_i$. Moreover, the Killing form of $L_i$ is the restriction of $\kappa$ to $L_i \times L_i$.*

*Proof.* As a first step, let $I$ be an arbitrary ideal of $L$. Then $I^\perp = \{x \in L | \kappa(x, y) = 0$ for all $y \in I\}$ is also an ideal, by the associativity of $\kappa$. Cartan's Criterion, applied to the Lie algebra $I$, shows that the ideal $I \cap I^\perp$ of $L$ is solvable (hence 0). Therefore, since $\dim I + \dim I^\perp = \dim L$, we must have $L = I \oplus I^\perp$.

Now proceed by induction on $\dim L$ to obtain the desired decomposition into direct sum of simple ideals. If $L$ has no nonzero proper ideal, then $L$ is simple already and we're done. Otherwise let $L_1$ be a minimal nonzero ideal; by the preceding paragraph, $L = L_1 \oplus L_1^\perp$. In particular, any ideal of $L_1$ is also an ideal of $L$, so $L_1$ is semisimple (hence simple, by minimality). For the same reason, $L_1^\perp$ is semisimple; by induction, it splits into a direct sum of simple ideals, which are also ideals of $L$. The decomposition of $L$ follows.

Next we have to prove that these simple ideals are unique. If $I$ is any simple ideal of $L$, then $[IL]$ is also an ideal of $I$, nonzero because $Z(L) = 0$; this forces $[IL] = I$. On the other hand, $[IL] = [IL_1] \oplus \ldots \oplus [IL_t]$, so all but one summand must be 0. Say $[IL_i] = I$. Then $I \subset L_i$, and $I = L_i$ (because $L_i$ is simple).

The last assertion of the theorem follows from Lemma 5.1. ▯

**Corollary.** *If $L$ is semisimple, then $L = [LL]$, and all ideals and homomorphic images of $L$ are semisimple. Moreover, each ideal of $L$ is a sum of certain simple ideals of $L$.* ▯

## *5.3. Inner derivations*

There is a further important consequence of nondegeneracy of the Killing form. Before stating it we recall explicitly the result of Exercise 2.1: ad $L$ is an ideal in Der $L$ (for any Lie algebra $L$). The proof depends on the simple observation: (*) $[\delta, \text{ad } x] = \text{ad } (\delta x)$, $x \in L$, $\delta \in \text{Der } L$.

**Theorem.** *If $L$ is semisimple, then ad $L$ = Der $L$ (i.e., every derivation of $L$ is inner).*

*Proof.* Since $L$ is semisimple, $Z(L) = 0$. Therefore, $L \to \text{ad } L$ is an isomorphism of Lie algebras. In particular, $M = \text{ad } L$ itself has nondegenerate Killing form (Theorem 5.1). If $D = \text{Der } L$, we just remarked that $[D, M] \subset M$. This implies (by Lemma 5.1) that $\kappa_M$ is the restriction to $M \times M$ of the Killing form $\kappa_D$ of $D$. In particular, if $I = M^\perp$ is the subspace of $D$ orthogonal to $M$ under $\kappa_D$, then the nondegeneracy of $\kappa_M$ forces $I \cap M = 0$. Both $I$ and $M$ are ideals of $D$, so we obtain $[I, M] = 0$. If $\delta \in I$, this forces $\text{ad } (\delta x) = 0$ for all $x \in L$ (by (*)), so in turn $\delta x = 0$ $(x \in L)$ because ad is $1-1$, and $\delta = 0$. Conclusion: $I = 0$, Der $L = M = \text{ad } L$. ▯

### *5.4. Abstract Jordan decomposition*

Theorem 5.3 can be used to introduce an abstract Jordan decomposition in an arbitrary semisimple Lie algebra $L$. Recall (Lemma B of (4.2)) that if $\mathfrak{A}$ is any F-algebra of finite dimension, then Der $\mathfrak{A}$ contains the semisimple and nilpotent parts in End $\mathfrak{A}$ of all its elements. In particular, since Der $L$ coincides with ad $L$ (5.3), while $L \to \text{ad } L$ is $1-1$, each $x \in L$ determines unique elements $s, n \in L$ such that ad $x$ = ad $s$ + ad $n$ is the usual Jordan decomposition of ad $x$ (in End $L$). This means that $x = s+n$, with $[sn] = 0$, $s$ **ad-semisimple** (i.e., ad $s$ semisimple), $n$ ad-nilpotent. We write $s = x_s$, $n = x_n$, and (by abuse of language) call these the **semisimple** and **nilpotent parts** of $x$.

The alert reader will object at this point that the notation $x_s$, $x_n$ is ambiguous in case $L$ happens to be a linear Lie algebra. It will be shown in (6.4) that the abstract decomposition of $x$ just obtained does in fact agree with the usual Jordan decomposition in all such cases. For the moment we shall be content to point out that this is true in the special case $L = \mathfrak{sl}(V)$ ($V$ finite dimensional): Write $x = x_s + x_n$ in End $V$ (usual Jordan decomposition), $x \in L$. Since $x_n$ is a nilpotent endomorphism, its trace is 0 and therefore $x_n \in L$. This forces $x_s$ also to have trace 0, so $x_s \in L$. Moreover, $\text{ad}_{\mathfrak{gl}(V)} x_s$ is semisimple (Lemma A of (4.2)), so $\text{ad}_L\, x_s$ is a fortiori semisimple; similarly $\text{ad}_L\, x_n$ is nilpotent, and $[\text{ad}_L\, x_s, \text{ad}_L\, x_n] = \text{ad}_L[x_s x_n] = 0$. By the uniqueness of the abstract Jordan decomposition in $L$, $x = x_s + x_n$ must be it.

## Exercises

1. Prove that if $L$ is nilpotent, the Killing form of $L$ is identically zero.
2. Prove that $L$ is solvable if and only if $[LL]$ lies in the radical of the Killing form.
3. Let $L$ be the two dimensional nonabelian Lie algebra (1.4), which is solvable. Prove that $L$ has nontrivial Killing form.
4. Let $L$ be the three dimensional solvable Lie algebra of Exercise 1.2. Compute the radical of its Killing form.
5. Let $L = \mathfrak{sl}(2, \mathsf{F})$. Compute the basis of $L$ dual to the standard basis, relative to the Killing form.
6. Let char $\mathsf{F} = p \neq 0$. Prove that $L$ is semisimple if its Killing form is nondegenerate. Show by example that the converse fails. [Look at $\mathfrak{sl}(3, \mathsf{F})$ modulo its center, when char $\mathsf{F} = 3$.]
7. Relative to the standard basis of $\mathfrak{sl}(3, \mathsf{F})$, compute the determinant of $\kappa$. Which primes divide it?
8. Let $L = L_1 \oplus \ldots \oplus L_t$ be the decomposition of a semisimple Lie algebra $L$ into its simple ideals. Show that the semisimple and nilpotent parts of $x \in L$ are the sums of the semisimple and nilpotent parts in the various $L_i$ of the components of $x$.

## Notes

Even in prime characteristic, nondegeneracy of the Killing form has very strong implications for the structure of a Lie algebra. See Seligman [1], Pollack [1], Kaplansky [1].

# 6. Complete reducibility of representations

*In this section all representations are finite dimensional, unless otherwise noted.*

We are going to study a semisimple Lie algebra $L$ by means of its adjoint representation (see §8). It turns out that $L$ is built up from copies of $\mathfrak{sl}(2, \mathsf{F})$; to study the adjoint action of such a three dimensional subalgebra of $L$, we need precise information about the representations of $\mathfrak{sl}(2, \mathsf{F})$, to be given in §7 below. First we prove an important general theorem (due to Weyl) about representations of an arbitrary semisimple Lie algebra.

### *6.1. Modules*

Let $L$ be a Lie algebra. It is often convenient to use the language of modules along with the (equivalent) language of representations. As in other algebraic theories, there is a natural definition. A vector space $V$, endowed with an operation $L \times V \to V$ (denoted $(x, v) \mapsto x.v$ or just $xv$) is called an **L-module** if the following conditions are satisfied:

(M1) $(ax+by).v = a(x.v)+b(y.v)$,

(M2) $x.(av+bw) = a(x.v)+b(x.w)$,

(M3) $[xy].v = x.y.v-y.x.v. \quad (x, y \in L;\ v, w \in V;\ a, b \in \mathsf{F})$.

For example, if $\phi: L \to \mathfrak{gl}(V)$ is a representation of $L$, then $V$ may be viewed as an $L$-module via the action $x.v = \phi(x)(v)$. Conversely, given an $L$-module $V$, this equation defines a representation $\phi: L \to \mathfrak{gl}(V)$.

A **homomorphism of L-modules** is a linear map $\phi: V \to W$ such that $\phi(x.v) = x.\phi(v)$. The kernel of such a homomorphism is then an $L$-submodule of $V$ (and the standard homomorphism theorems all go through without difficulty). When $\phi$ is an isomorphism of vector spaces, we call it an **isomorphism** of $L$-modules; in this case, the two modules are said to afford **equivalent** representations of $L$. An $L$-module $V$ is called **irreducible** if it has precisely two $L$-submodules (itself and 0); in particular, *we do not regard a zero dimensional vector space as an irreducible L-module*. We do, however, allow a one dimensional space on which $L$ acts (perhaps trivially) to be called irreducible. $V$ is called **completely reducible** if $V$ is a direct sum of irreducible $L$-submodules, or equivalently (Exercise 2), if each $L$-submodule $W$ of $V$ has a complement $W'$ (an $L$-submodule such that $V = W \oplus W'$). When

$W$, $W'$ are arbitrary $L$-modules, we can of course make their direct sum an $L$-module in the obvious way, by defining $x.(w, w') = (x.w, x.w')$. These notions are all standard and also make sense when $\dim V = \infty$. Of course, the terminology "irreducible" and "completely reducible" applies equally well to representations of $L$.

Given a representation $\phi: L \to \mathfrak{gl}(V)$, the associative algebra (with 1) generated by $\phi(L)$ in End $V$ leaves invariant precisely the same subspaces as $L$. Therefore, all the usual results (e.g., Jordan-Hölder Theorem) for modules over associative rings hold for $L$ as well. For later use, we recall the well known Schur's Lemma.

**Schur's Lemma.** *Let $\phi: L \to \mathfrak{gl}(V)$ be irreducible. Then the only endomorphisms of $V$ commuting with all $\phi(x)$ $(x \in L)$ are the scalars.* □

$L$ itself is an $L$-module (for the adjoint representation). An $L$-submodule is just an ideal, so it follows that a simple algebra $L$ is irreducible as $L$-module, while a semisimple algebra is completely reducible (Theorem 5.2).

For later use we mention a couple of standard ways in which to fabricate new $L$-modules from old ones. Let $V$ be an $L$-module. Then the dual vector space $V^*$ becomes an $L$-module (called the **dual** or **contragredient**) if we define, for $f \in V^*$, $v \in V$, $x \in L$: $(x.f)(v) = -f(x.v)$. Axioms (M1), (M2) are almost obvious, so we just check (M3):

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

If $V$, $W$ are $L$-modules, let $V \otimes W$ be the tensor product over $\mathsf{F}$ of the underlying vector spaces. Recall that if $V$, $W$ have respective bases $(v_1, \dots, v_m)$ and $(w_1, \dots, w_n)$, then $V \otimes W$ has a basis consisting of the $mn$ vectors $v_i \otimes w_j$. The reader may know how to give a module structure to the tensor product of two modules for a group $G$: on the generators $v \otimes w$, require $g.(v \otimes w) = g.v \otimes g.w$. For Lie algebras the correct definition is gotten by "differentiating" this one: $x.(v \otimes w) = x.v \otimes w + v \otimes x.w$. As before, the crucial axiom to verify is (M3):

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

Expanding $(x.y-y.x).(v\otimes w)$ yields the same result.

Given a vector space $V$ over $\mathsf{F}$, there is a standard (and very useful) isomorphism of vector spaces: $V^* \otimes V \to \text{End } V$, given by sending a typical generator $f \otimes v$ $(f \in V^*, v \in V)$ to the endomorphism whose value at $w \in V$

is $f(w)v$. It is a routine matter (using dual bases) to show that this does set up an epimorphism $V^* \otimes V \to \text{End } V$; since both sides have dimension $n^2$ ($n = \dim V$), this must be an isomorphism.

Now if $V$ (hence $V^*$) is in addition an $L$-module, then $V^* \otimes V$ becomes an $L$-module in the way described above. Therefore, End $V$ can also be viewed as an $L$-module via the isomorphism just exhibited. This action of $L$ on End $V$ can also be described directly: $(x.f)(v) = x.f(v) - f(x.v)$, $x \in L$, $f \in \text{End } V$, $v \in V$ (verify!). More generally, if $V$ and $W$ are two $L$-modules, then $L$ acts naturally on the space Hom $(V, W)$ of linear maps by the rule $(x.f)(v) = x.f(v) - f(x.v)$. (This action arises from the isomorphism between Hom $(V, W)$ and $V^* \otimes W$.)

### *6.2. Casimir element of a representation*

In §5 we used Cartan's trace criterion for solvability to prove that a semisimple Lie algebra $L$ has nondegenerate Killing form. More generally, let $L$ be semisimple and let $\phi: L \to \mathfrak{gl}(V)$ be a **faithful** (i.e., 1–1) representation of $L$. Define a symmetric bilinear form $\beta(x, y) = Tr(\phi(x)\phi(y))$ on $L$. The form $\beta$ is associative, thanks to identity (*) in (4.3), so in particular its radical $S$ is an ideal of $L$. Moreover, $\beta$ is nondegenerate: indeed, Theorem 4.3 shows that $\phi(S) \cong S$ is solvable, so $S = 0$. (The Killing form is just $\beta$ in the special case $\phi = \text{ad}$.)

Now let $L$ be semisimple, $\beta$ any nondegenerate symmetric associative bilinear form on $L$. If $(x_1, \ldots, x_n)$ is a basis of $L$, there is a uniquely determined dual basis $(y_1, \ldots, y_n)$ relative to $\beta$, satisfying $\beta(x_i, y_j) = \delta_{ij}$. If $x \in L$, we can write $[xx_i] = \sum_j a_{ij}x_j$ and $[xy_i] = \sum_j b_{ij}y_j$. Using the associativity of $\beta$, we compute: $a_{ik} = \sum_j a_{ij}\beta(x_j, y_k) = \beta([xx_i], y_k) = \beta(-[x_ix], y_k) = \beta(x_i, -[xy_k]) = -\sum_j b_{kj}\beta(x_i, y_j) = -b_{ki}$.

If $\phi: L \to \mathfrak{gl}(V)$ is any representation of $L$, write $c_\phi(\beta) = \sum_i \phi(x_i)\phi(y_i) \in$ End $V$ ($x_i$, $y_i$ running over dual bases relative to $\beta$, as above). Using the identity (in End $V$) $[x, yz] = [x, y]z + y[x, z]$ and the fact that $a_{ik} = -b_{ki}$ (for $x \in L$ as above), we obtain: $[\phi(x), c_\phi(\beta)] = \sum_i [\phi(x), \phi(x_i)]\phi(y_i) + \sum_i \phi(x_i)[\phi(x), \phi(y_i)] = \sum_{i,j} a_{ij}\phi(x_j)\phi(y_i) + \sum_{i,j} b_{ij}\phi(x_i)\phi(y_j) = 0$. In other words, $c_\phi(\beta)$ *is an endomorphism of $V$ commuting with $\phi(L)$.*

To bring together the preceding remarks, let $\phi: L \to \mathfrak{gl}(V)$ be a faithful representation, with (nondegenerate!) trace form $\beta(x, y) = Tr(\phi(x)\phi(y))$. In this case, having fixed a basis $(x_1, \ldots, x_n)$ of $L$, we write simply $c_\phi$ for $c_\phi(\beta)$ and call this the **Casimir element of $\boldsymbol{\phi}$**. Its trace is $\sum_i Tr(\phi(x_i)\phi(y_i)) = \sum_i \beta(x_i, y_i) = \dim L$. In case $\phi$ is also irreducible, Schur's Lemma (6.1) implies that $c_\phi$ is a scalar (equal to $\dim L/\dim V$, in view of the preceding sentence); in this case we see that $c_\phi$ is independent of the basis of $L$ which we chose.

*Example.* $L = \mathfrak{sl}(2, \mathsf{F})$, $V = \mathsf{F}^2$, $\phi$ the identity map $L \to \mathfrak{gl}(V)$. Let $(x, h, y)$ be the standard basis of $L$ (2.1). It is quickly seen that the dual basis relative to the trace form is $(y, h/2, x)$, so $c_\phi = xy+(1/2)h^2+yx = \begin{pmatrix} 3/2 & 0 \\ 0 & 3/2 \end{pmatrix}$. Notice that $3/2 = \dim L/\dim V$.

When $\phi$ is no longer faithful, a slight modification is needed. Ker $\phi$ is an ideal of $L$, hence a sum of certain simple ideals (Corollary 5.2). Let $L'$ denote the sum of the remaining simple ideals (Theorem 5.2). Then the restriction of $\phi$ to $L'$ is a faithful representation of $L'$, and we make the preceding construction (using dual bases of $L'$); the resulting element of End $V$ is again called the Casimir element of $\phi$ and denoted $c_\phi$. Evidently it commutes with $\phi(L) = \phi(L')$, etc.

One last remark: It is often convenient to assume that we are dealing with a faithful representation of $L$, which amounts to studying the representations of certain (semisimple) ideals of $L$. If $L$ is simple, only the one dimensional module (on which $L$ acts trivially) or the module 0 will fail to be faithful.

### *6.3. Weyl's Theorem*

**Lemma.** *Let $\phi$: $L \to \mathfrak{gl}(V)$ be a representation of a semisimple Lie algebra $L$. Then $\phi(L) \subset \mathfrak{sl}(V)$. In particular, $L$ acts trivially on any one dimensional $L$-module.*

*Proof.* Use the fact that $L = [LL]$ (5.2) along with the fact that $\mathfrak{sl}(V)$ is the derived algebra of $\mathfrak{gl}(V)$. □

**Theorem (Weyl).** *Let $\phi$: $L \to \mathfrak{gl}(V)$ be a (finite dimensional) representation of a semisimple Lie algebra. Then $\phi$ is completely reducible.*

*Proof.* We start with the *special case* in which $V$ has an $L$-submodule $W$ of codimension one. Since $L$ acts trivially on $V/W$, by the lemma, we may denote this module $\mathsf{F}$ without misleading the reader: $0 \to W \to V \to \mathsf{F} \to 0$ is therefore exact. Using induction on dim $W$, we can reduce to the case where $W$ is an *irreducible* $L$-module, as follows. Let $W'$ be a proper nonzero submodule of $W$. This yields an exact sequence: $0 \to W/W' \to V/W' \to \mathsf{F} \to 0$. By induction, this sequence "splits", i.e., there exists a one dimensional $L$-submodule of $V/W'$ (say $\tilde{W}/W'$) complementary to $W/W'$. So we get another exact sequence: $0 \to W' \to \tilde{W} \to \mathsf{F} \to 0$. This is like the original situation, except that $\dim W' < \dim W$, so induction provides a (one dimensional) submodule $X$ complementary to $W'$ in $\tilde{W}$: $\tilde{W} = W' \oplus X$. But $V/W' = W/W' \oplus \tilde{W}/W'$. It follows that $V = W \oplus X$, since the dimensions add up to dim $V$ and since $W \cap X = 0$.

Now we may assume that $W$ is irreducible. (We may also assume without loss of generality that $L$ acts faithfully on $V$.) Let $c = c_\phi$ be the Casimir element of $\phi$ (6.2). Since $c$ commutes with $\phi(L)$, *$c$ is actually an $L$-module endomorphism of $V$*; in particular, $c(W) \subset W$ and Ker $c$ is an $L$-submodule of $V$. Because $L$ acts trivially on $V/W$ (i.e., $\phi(L)$ sends $V$ into $W$), $c$ must do likewise (as a linear combination of products of elements $\phi(x)$). So $c$ has trace 0 on $V/W$. On the other hand, $c$ acts as a scalar on the irreducible $L$-submodule $W$ (Schur's Lemma); this scalar cannot be 0, because that would force $Tr_V(c) = 0$, contrary to the conclusion of (6.2). It follows that Ker $c$ is a one dimensional $L$-submodule of $V$ which intersects $W$ trivially. This is the desired complement to $W$.

Now we can attack the *general case*. Let $W$ be a nonzero submodule of $V$: $0 \to W \to V \to V/W \to 0$. Let Hom$(V, W)$ be the space of linear maps $V \to W$, viewed as $L$-module (6.1). Let $\mathscr{V}$ be the subspace of Hom $(V, W)$ consisting of those maps whose restriction to $W$ is a scalar multiplication. $\mathscr{V}$ is actually an $L$-submodule: Say $f|_W = a.1_W$; then for $x \in L$, $w \in W$, $(x.f)(w) = x.f(w) - f(x.w) = a(x.w) - a(x.w) = 0$, so $x.f|_W = 0$. Let $\mathscr{W}$ be the subspace of $\mathscr{V}$ consisting of those $f$ whose restriction to $W$ is zero. The preceding calculation shows that $\mathscr{W}$ is also an $L$-submodule and that $L$ maps $\mathscr{V}$ into $\mathscr{W}$. Moreover, $\mathscr{V}/\mathscr{W}$ has dimension one, because each $f \in \mathscr{V}$ is determined (modulo $\mathscr{W}$) by the scalar $f|_W$. This places us precisely in the situation $0 \to \mathscr{W} \to \mathscr{V} \to \mathsf{F} \to 0$ already treated above.

According to the first part of the proof, $\mathscr{V}$ has a one dimensional submodule complementary to $\mathscr{W}$. Let $f$: $V \to W$ span it, so after multiplying by a nonzero scalar we may assume that $f|_W = 1_W$. To say that $L$ kills $f$ is just to say that $0 = (x.f)(v) = x.f(v) - f(x.v)$, i.e., that *$f$ is an $L$-homomorphism.* Therefore Ker $f$ is an $L$-submodule of $V$. Since $f$ maps $V$ into $W$ and acts as $1_W$ on $W$, we conclude that $V = W \oplus \text{Ker } f$, as desired. ▯

### *6.4. Preservation of Jordan decomposition*

Weyl's Theorem is of course fundamental for the study of representations of a semisimple Lie algebra $L$. We offer here a more immediate application, to the problem of showing that the abstract Jordan decomposition (5.4) is compatible with the various linear representations of $L$.

**Theorem.** *Let $L \subset \mathfrak{gl}(V)$ be a semisimple linear Lie algebra ($V$ finite dimensional). Then $L$ contains the semisimple and nilpotent parts in $\mathfrak{gl}(V)$ of all its elements. In particular, the abstract and usual Jordan decompositions in $L$ coincide.*

*Proof.* The last assertion follows from the first, because each type of Jordan decomposition is unique (4.2, 5.4).

Let $x \in L$ be arbitrary, with Jordan decomposition $x = x_s + x_n$ in $\mathfrak{gl}(V)$. The problem is just to show that $x_s$, $x_n$ lie in $L$. Since ad $x(L) \subset L$, it follows from Proposition 4.2(c) that ad $x_s(L) \subset L$ and ad $x_n(L) \subset L$, where ad $=$ ad$_{\mathfrak{gl}(V)}$. In other words, $x_s, x_n \in N_{\mathfrak{gl}(V)}(L) = N$, which is a Lie subalgebra of $\mathfrak{gl}(V)$ including $L$ as an ideal. If we could show that $N = L$ we'd be done, but unfortunately this is false: e.g., since $L \subset \mathfrak{sl}(V)$ (Lemma 6.3), the scalars lie in $N$ but not in $L$. Therefore we need to get $x_s$, $x_n$ into a smaller subalgebra than $N$, which can be shown to equal $L$. If $W$ is any $L$-submodule of $V$, define $L_W = \{y \in \mathfrak{gl}(V) | y(W) \subset W$ and $Tr(y|_W) = 0\}$. For example, $L_V = \mathfrak{sl}(V)$. Since $L = [LL]$, it is clear that $L$ lies in all such $L_W$. Set $L'$ = intersection of $N$ with all spaces $L_W$. Clearly, $L'$ is a subalgebra of $N$ including $L$ as an ideal (but notice that $L'$ does exclude the scalars). Even more is true: If $x \in L$, then $x_s$, $x_n$ also lie in $L_W$, and therefore in $L'$.

It remains to prove that $L = L'$. $L'$ being a finite dimensional $L$-module, Weyl's Theorem (6.3) permits us to write $L' = L + M$ for some $L$-submodule $M$, where the sum is direct. But $[L, L'] \subset L$ (since $L' \subset N$), so the action of $L$ on $M$ is trivial. Let $W$ be any irreducible $L$-submodule of $V$. If $y \in M$, then $[L, y] = 0$, so Schur's Lemma implies that $y$ acts on $W$ as a scalar. On the other hand, $Tr(y|_W) = 0$ because $y \in L_W$. Therefore $y$ acts on $W$ as zero. $V$ can be written as a direct sum of irreducible $L$-submodules (by Weyl's Theorem), so in fact $y = 0$. This means $M = 0$, $L = L'$. ▯

**Corollary.** *Let L be a semisimple Lie algebra,* $\phi: L \to \mathfrak{gl}(V)$ *a (finite dimensional) representation of L. If* $x = s+n$ *is the abstract Jordan decomposition of* $x \in L$, *then* $\phi(x) = \phi(s)+\phi(n)$ *is the usual Jordan decomposition of* $\phi(x)$.

*Proof.* The algebra $\phi(L)$ is spanned by the eigenvectors of $\text{ad}_{\phi(L)}\, \phi(s)$, since $L$ has this property relative to ad $s$; therefore, $\text{ad}_{\phi(L)}\, \phi(s)$ is semisimple. Similarly, $\text{ad}_{\phi(L)}\, \phi(n)$ is nilpotent, and it commutes with $\text{ad}_{\phi(L)}\, \phi(s)$. Accordingly, $\phi(x) = \phi(s)+\phi(n)$ is the abstract Jordan decomposition of $\phi(x)$ in the semisimple Lie algebra $\phi(L)$ (5.4). Applying the theorem, we get the desired conclusion. ▯

## Exercises

1. Using the standard basis for $L = \mathfrak{sl}(2, \mathsf{F})$, write down the Casimir element of the adjoint representation of $L$ (cf. Exercise 5.5). Do the same thing for the usual (3-dimensional) representation of $\mathfrak{sl}(3, \mathsf{F})$, first computing dual bases relative to the trace form.
2. Let $V$ be an $L$-module. Prove that $V$ is a direct sum of irreducible submodules if and only if each $L$-submodule of $V$ possesses a complement.
3. If $L$ is solvable, every irreducible representation of $L$ is one dimensional.
4. Use Weyl's Theorem to give another proof that for $L$ semisimple, ad $L$ = Der $L$ (Theorem 5.3). [If $\delta \in$ Der $L$, make the direct sum $\mathsf{F}+L$ into an $L$-module via the rule $x.(a, y) = (0, a\delta(x)+[xy])$. Then consider a complement to the submodule $L$.]
5. A Lie algebra $L$ for which Rad $L = Z(L)$ is called **reductive**. (Examples: $L$ abelian, $L$ semisimple, $L = \mathfrak{gl}(n, \mathsf{F})$.)
   (a) If $L$ is reductive, then $L$ is a completely reducible ad $L$-module. [If ad $L \neq 0$, use Weyl's Theorem.] In particular, $L$ is the direct sum of $Z(L)$ and $[LL]$, with $[LL]$ semisimple.
   (b) If $L$ is a classical linear Lie algebra (1.2), then $L$ is semisimple. [Cf. Exercise 1.9.]

(c) If $L$ is a completely reducible ad $L$-module, then $L$ is reductive.
(d) If $L$ is reductive, then all finite dimensional representations of $L$ in which $Z(L)$ is represented by semisimple endomorphisms are completely reducible.

6. Let $L$ be a simple Lie algebra. Let $\beta(x,y)$ and $\gamma(x,y)$ be two symmetric associative bilinear forms on $L$. If $\beta,\gamma$ are nondegenerate, prove that $\beta$ and $\gamma$ are proportional. [Use Schur's Lemma.]
7. It will be seen later on that $\mathfrak{sl}(n,\mathsf{F})$ is actually *simple*. Assuming this and using Exercise 6, prove that the Killing form $\kappa$ on $\mathfrak{sl}(n,\mathsf{F})$ is related to the ordinary trace form by $\kappa(x,y)=2n\,Tr(xy)$.
8. If $L$ is a Lie algebra, then $L$ acts (via ad) on $(L\otimes L)^*$, which may be identified with the space of all bilinear forms $\beta$ on $L$. Prove that $\beta$ is associative if and only if $L.\beta=0$.
9. Let $L'$ be a semisimple subalgebra of a semisimple Lie algebra $L$. If $x\in L'$, its Jordan decomposition in $L'$ is also its Jordan decomposition in $L$.

## Notes

The proof of Weyl's Theorem is based on Brauer [1]. The original proof was quite different, using integration on compact Lie groups, cf. Freudenthal, de Vries [1]. For Theorem 6.4 we have followed Bourbaki [1].

# 7. Representations of $\mathfrak{sl}(2,\mathsf{F})$

*In this section (as in §6) all modules will be assumed to be finite dimensional over* F. $L$ denotes $\mathfrak{sl}(2,\mathsf{F})$, whose standard basis consists of

$$x=\begin{pmatrix}0&1\\0&0\end{pmatrix},\quad y=\begin{pmatrix}0&0\\1&0\end{pmatrix},\quad h=\begin{pmatrix}1&0\\0&-1\end{pmatrix}$$

(Example 2.1). Then $[hx]=2x$, $[hy]=-2y$, $[xy]=h$.

## *7.1. Weights and maximal vectors*

Let $V$ be an arbitrary $L$-module. Since $h$ is semisimple, Corollary 6.4 implies that $h$ acts diagonally on $V$. (The assumption that F is algebraically closed insures that all the required eigenvalues already lie in F.) This yields a decomposition of $V$ as direct sum of eigenspaces $V_\lambda=\{v\in V\,|\,h.v=\lambda v\}$, $\lambda\in\mathsf{F}$. Of course, the subspace $V_\lambda$ still makes sense (and is 0) when $\lambda$ is not an eigenvalue for the endomorphism of $V$ which represents $h$. Whenever $V_\lambda\neq 0$, we call $\lambda$ a **weight** of $h$ in $V$ and we call $V_\lambda$ a **weight space**.

**Lemma.** *If $v\in V_\lambda$, then $x.v\in V_{\lambda+2}$ and $y.v\in V_{\lambda-2}$.*

*Proof.* $h.(x.v) = [h, x].v + x.h.v = 2x.v + \lambda x.v = (\lambda+2)x.v$, and similarly for $y$. □

*Remark.* The lemma implies that $x$, $y$ are represented by nilpotent endomorphisms of $V$; but this already follows from Theorem 6.4.

Since $\dim V < \infty$, and the sum $V = \coprod_{\lambda \in F} V_\lambda$ is direct, there must exist $V_\lambda \neq 0$ such that $V_{\lambda+2} = 0$. (Thanks to the lemma, $x.v = 0$ for any $v \in V_\lambda$.) For such $\lambda$, any nonzero vector in $V_\lambda$ will be called a **maximal vector** of weight $\lambda$.

### *7.2. Classification of irreducible modules*

Assume now that $V$ is an irreducible $L$-module. Choose a maximal vector, say $v_0 \in V_\lambda$; set $v_{-1} = 0$, $v_i = (1/i!)y^i.v_0$ $(i \geq 0)$.

**Lemma.** (*a*) $h.v_i = (\lambda - 2i)v_i$,

(*b*) $y.v_i = (i+1)v_{i+1}$,

(*c*) $x.v_i = (\lambda - i + 1)v_{i-1}$ $(i \geq 0)$.

*Proof.* (a) follows from repeated application of Lemma 7.1, while (b) is just the definition. To prove (c), use induction on $i$, the case $i = 0$ being clear (since $v_{-1} = 0$, by convention). Observe that

$$
\begin{aligned}
ix.v_i &= x.y.v_{i-1} && \text{(by definition)} \\
&= [x, y].v_{i-1} + y.x.v_{i-1} \\
&= h.v_{i-1} + y.x.v_{i-1} \\
&= (\lambda - 2(i-1))v_{i-1} + (\lambda - i + 2)y.v_{i-2} \\
& && \text{(by (a) and induction)} \\
&= (\lambda - 2i + 2)v_{i-1} + (i-1)(\lambda - i + 2)v_{i-1} && \text{(by (b))} \\
&= i(\lambda - i + 1)v_{i-1}.
\end{aligned}
$$

Then divide both sides by $i$. □

Thanks to formula (a), the nonzero $v_i$ are all linearly independent. But $\dim V < \infty$. Let $m$ be the smallest integer for which $v_m \neq 0$, $v_{m+1} = 0$; evidently $v_{m+i} = 0$ for all $i > 0$. Taken together, formulas (a)–(c) show that the subspace of $V$ with basis $(v_0, v_1, \ldots, v_m)$ is an $L$-submodule, different from 0. Because $V$ is irreducible, this subspace must be all of $V$. Moreover, relative to the ordered basis $(v_0, v_1, \ldots, v_m)$, the matrices of the endomorphisms representing $x$, $y$, $h$ can be written down explicitly; notice that $h$ yields a diagonal matrix, while $x$ and $y$ yield (respectively) upper and lower triangular nilpotent matrices.

A closer look at formula (c) reveals a striking fact: for $i = m+1$, the left side is 0, whereas the right side is $(\lambda - m)v_m$. Since $v_m \neq 0$, we conclude that $\lambda = m$. In other words, *the weight of a maximal vector is a nonnegative integer* (one less than $\dim V$). We call it the **highest weight** of $V$. Moreover, each weight $\mu$ occurs with multiplicity one (i.e., $\dim V_\mu = 1$ if $V_\mu \neq 0$), by formula (a); in particular, since $V$ determines $\lambda$ uniquely ($\lambda = \dim V - 1$), the maximal vector $v_0$ is the only possible one in $V$ (apart from nonzero scalar multiples). To summarize:

**Theorem.** *Let $V$ be an irreducible module for $L = \mathfrak{sl}(2, \mathsf{F})$.*

(*a*) *Relative to $h$, $V$ is the direct sum of weight spaces $V_\mu$, $\mu = m, m-2, \ldots, -(m-2), -m$, where $m+1 = \dim V$ and $\dim V_\mu = 1$ for each $\mu$.*

(*b*) *$V$ has (up to nonzero scalar multiples) a unique maximal vector, whose weight (called the highest weight of $V$) is $m$.*

(*c*) *The action of $L$ on $V$ is given explicitly by the above formulas, if the basis is chosen in the prescribed fashion. In particular, there exists at most one irreducible $L$-module (up to isomorphism) of each possible dimension $m+1$, $m \geq 0$.* □

**Corollary.** *Let $V$ be any (finite dimensional) $L$-module, $L = \mathfrak{sl}(2, \mathsf{F})$. Then the eigenvalues of $h$ on $V$ are all integers, and each occurs along with its negative (an equal number of times). Moreover, in any decomposition of $V$ into direct sum of irreducible submodules, the number of summands is precisely $\dim V_0 + \dim V_1$.*

*Proof.* If $V = 0$, there is nothing to prove. Otherwise use Weyl's Theorem (6.3) to write $V$ as direct sum of irreducible submodules. The latter are described by the theorem, so the first assertion of the corollary is obvious. For the second, just observe that each irreducible $L$-module has a unique occurrence of either the weight 0 or else the weight 1 (but not both). □

For the purposes of this chapter, the theorem and corollary just proved are quite adequate. However, it is unreasonable to leave the subject before investigating whether or not $\mathfrak{sl}(2, \mathsf{F})$ does have an irreducible module of each possible highest weight $m = 0, 1, 2, \ldots$. Of course, we already know how to construct suitable modules in low dimensions: the trivial module (dimension 1), the natural representation (dimension 2), the adjoint representation (dimension 3). For arbitrary $m \geq 0$, formulas (a)–(c) of Lemma 7.2 can actually be used to define an irreducible representation of $L$ on an $m+1$-dimensional vector space over $\mathsf{F}$ with basis $(v_0, v_1, \ldots, v_m)$, called $V(m)$. As is customary, the (easy) verification will be left for the reader (Exercise 3). (For a general existence theorem, see (20.3) below.)

One further observation: The symmetry in the structure of $V(m)$ can be made more obvious if we exploit the discussion of *exponentials* in (2.3). Let $\phi: L \to \mathfrak{gl}(V(m))$ be the irreducible representation of highest weight $m$. Then $\phi(x)$, $\phi(y)$ are nilpotent endomorphisms, in view of the formulas above, so we can define an automorphism of $V(m)$ by $\tau = \exp \phi(x) \exp \phi(-y) \exp \phi(x)$. We may as well assume $m > 0$, so the representation is faithful ($L$ being simple). The discussion in (2.3) shows that conjugating $\phi(h)$ by $\tau$ has precisely the same effect as applying $\exp(\operatorname{ad} \phi(x)) \exp(\operatorname{ad} \phi(-y)) \exp(\operatorname{ad} \phi(x))$ to $\phi(h)$. But $\phi(L)$ is isomorphic to $L$, so this can be calculated just as in (2.3). Conclusion: $\tau\phi(h)\tau^{-1} = -\phi(h)$, or $\tau\phi(h) = -\phi(h)\tau$. From this we see at once that $\tau$ sends the basis vector $v_i$ of weight $m-2i$ to the basis vector $v_{m-i}$ of weight $-(m-2i)$. (The discussion in (2.3) was limited to the special case $m = 1$.) More generally, if $V$ is any finite dimensional $L$-module, then $\tau$ interchanges positive and negative weight spaces.

## Exercises

(In these exercises, $L = \mathfrak{sl}(2, \mathsf{F})$.)

1. Use Lie's Theorem to prove the existence of a maximal vector in an arbitrary finite dimensional $L$-module. [Look at the subalgebra $B$ spanned by $h$ and $x$.]
2. $M = \mathfrak{sl}(3, \mathsf{F})$ contains a copy of $L$ in its upper left-hand $2\times 2$ position. Write $M$ as direct sum of irreducible $L$-submodules ($M$ viewed as $L$-module via the adjoint representation): $V(0) \oplus V(1) \oplus V(1) \oplus V(2)$.
3. Verify that formulas (a)–(c) of Lemma 7.2 do define an irreducible representation of $L$. [To show that they define a representation, it suffices to show that the matrices corresponding to $x$, $y$, $h$ satisfy the same structural equations as $x$, $y$, $h$.]
4. The irreducible representation of $L$ of highest weight $m$ can also be realized "naturally", as follows. Let $X$, $Y$ be a basis for the two dimensional vector space $\mathsf{F}^2$, on which $L$ acts as usual. Let $\mathscr{R} = \mathsf{F}[X, Y]$ be the polynomial algebra in two variables, and extend the action of $L$ to $\mathscr{R}$ by the derivation rule: $z.fg = (z.f)g+f(z.g)$, for $z \in L, f, g \in \mathscr{R}$. Show that this extension is well defined and that $\mathscr{R}$ becomes an $L$-module. Then show that the subspace of homogeneous polynomials of degree $m$, with basis $X^m$, $X^{m-1}Y, \ldots, XY^{m-1}$, $Y^m$, is invariant under $L$ and irreducible of highest weight $m$.
5. Suppose char $\mathsf{F} = p > 0$, $L = \mathfrak{sl}(2, \mathsf{F})$. Prove that the representation $V(m)$ of $L$ constructed as in Exercise 3 or 4 is irreducible so long as the highest weight $m$ is strictly less than $p$, but reducible when $m = p$.
6. Decompose the tensor product of the two $L$-modules $V(3)$, $V(7)$ into the sum of irreducible submodules: $V(4) \oplus V(6) \oplus V(8) \oplus V(10)$. Try to develop a general formula for the decomposition of $V(m) \otimes V(n)$.
7. In this exercise we construct certain *infinite dimensional* $L$-modules. Let $\lambda \in \mathsf{F}$ be an arbitrary scalar. Let $Z(\lambda)$ be a vector space over $\mathsf{F}$ with countably infinite basis $(v_0, v_1, v_2, \ldots)$.
   (a) Prove that formulas (a)–(c) of Lemma 7.2 define an $L$-module structure on $Z(\lambda)$, and that every nonzero $L$-submodule of $Z(\lambda)$ contains at least one maximal vector.
   (b) Suppose $\lambda+1=i$ is a nonnegative integer. Prove that $v_i$ is a maximal vector (e.g., $\lambda=-1$, $i=0$). This induces an $L$-module homomorphism $Z(\mu)\xrightarrow{\phi} Z(\lambda)$, $\mu=\lambda-2i$, sending $v_0$ to $v_i$. Show that $\phi$ is a monomorphism, and that Im $\phi$, $Z(\lambda)/\text{Im } \phi$ are both irreducible $L$-modules (but $Z(\lambda)$ fails to be completely reducible when $i>0$).
   (c) Suppose $\lambda+1$ is not a nonnegative integer. Prove that $Z(\lambda)$ is irreducible.

## 8. Root space decomposition

Throughout this section $L$ denotes a (nonzero) semisimple Lie algebra. We are going to study in detail the structure of $L$, via its adjoint representation. Our main tools will be the Killing form, and Theorems 6.4, 7.2 (which rely heavily on Weyl's Theorem). The reader should bear in mind the special case $L = \mathfrak{sl}(2, \mathsf{F})$ (or more generally, $\mathfrak{sl}(n, \mathsf{F})$) as a guide to what is going on.

### 8.1. Maximal toral subalgebras and roots

If $L$ consisted entirely of nilpotent (i.e., ad-nilpotent) elements, then $L$ would be nilpotent (Engel's Theorem). This not being the case, we can find $x \in L$ whose semisimple part $x_s$ in the abstract Jordan decomposition (5.4) is nonzero. This shows that $L$ possesses nonzero subalgebras (e.g., the span of such $x_s$) consisting of semisimple elements. Call such a subalgebra **toral**. The following lemma is roughly analogous to Engel's Theorem.

**Lemma.** *A toral subalgebra of $L$ is abelian.*

*Proof.* Let $T$ be toral. We have to show that $\mathrm{ad}_T\, x = 0$ for all $x$ in $T$. Since $\mathrm{ad}\, x$ is diagonalizable ($\mathrm{ad}\, x$ being semisimple and $\mathsf{F}$ being algebraically closed), this amounts to showing that $\mathrm{ad}_T\, x$ has no nonzero eigenvalues. Suppose, on the contrary, that $[xy] = ay$ ($a \neq 0$) for some nonzero $y$ in $T$. Then $\mathrm{ad}_T\, y(x) = -ay$ is itself an eigenvector of $\mathrm{ad}_T\, y$, of eigenvalue 0. On the other hand, we can write $x$ as a linear combination of eigenvectors of $\mathrm{ad}_T\, y$ ($y$ being semisimple also); after applying $\mathrm{ad}_T\, y$ to $x$, all that is left is a combination of eigenvectors which belong to nonzero eigenvalues, if any. This contradicts the preceding conclusion. $\square$

Now fix a **maximal toral subalgebra** $H$ of $L$, i.e., a toral subalgebra not properly included in any other. (The notation $H$ is less natural than $T$, but more traditional.) For example, if $L = \mathfrak{sl}(n, \mathsf{F})$, it is easy to verify (Exercise 1) that $H$ can be taken to be the set of diagonal matrices (of trace 0).

Since $H$ is abelian (by the above lemma), $\mathrm{ad}_L\, H$ is a commuting family of semisimple endomorphisms of $L$. According to a standard result in linear algebra, $\mathrm{ad}_L\, H$ is *simultaneously diagonalizable*. In other words, $L$ is the direct sum of the subspaces $L_\alpha = \{x \in L \mid [hx] = \alpha(h)x \text{ for all } h \in H\}$, where $\alpha$ ranges over $H^*$. Notice that $L_0$ is simply $C_L(H)$, the centralizer of $H$; it includes $H$, thanks to the lemma. The set of all nonzero $\alpha \in H^*$ for which $L_\alpha \neq 0$ is denoted by $\Phi$; the elements of $\Phi$ are called the **roots** of $L$ relative to $H$ (and are finite in number). With this notation we have a **root space decomposition** (or **Cartan decomposition**): (*) $L = C_L(H) \oplus \coprod_{\alpha \in \Phi} L_\alpha$. When $L = \mathfrak{sl}(n, \mathsf{F})$, for example, the reader will observe that (*) corresponds to the decomposition of $L$ given by the standard basis (1.2). Our aim in what follows is first to prove that $H = C_L(H)$, then to describe the set of roots in more detail, and ultimately to show that $\Phi$ characterizes $L$ completely.

We begin with a few simple observations about the root space decomposition.

**Proposition.** *For all* $\alpha, \beta \in H^*$, $[L_\alpha L_\beta] \subset L_{\alpha+\beta}$. *If* $x \in L_\alpha$, $\alpha \neq 0$, *then* ad $x$ *is nilpotent. If* $\alpha, \beta \in H^*$, *and* $\alpha+\beta \neq 0$, *then* $L_\alpha$ *is orthogonal to* $L_\beta$, *relative to the Killing form* $\kappa$ *of* $L$.

*Proof.* The first assertion follows from the Jacobi identity: $x \in L_\alpha, y \in L_\beta$, $h \in H$ imply that ad $h([xy]) = [[hx]y]+[x[hy]] = \alpha(h)[xy]+\beta(h)[xy] = (\alpha+\beta)(h)[xy]$. The second assertion is an immediate consequence of the first.

For the remaining assertion, find $h \in H$ for which $(\alpha+\beta)(h) \neq 0$. Then if $x \in L_\alpha$, $y \in L_\beta$, associativity of the form allows us to write $\kappa([hx], y) = -\kappa([xh], y) = -\kappa(x, [hy])$, or $\alpha(h)\kappa(x, y) = -\beta(h)\kappa(x, y)$, or $(\alpha+\beta)(h)\kappa(x, y) = 0$. This forces $\kappa(x, y) = 0$. □

**Corollary.** *The restriction of the Killing form to* $L_0 = C_L(H)$ *is nondegenerate.*

*Proof.* We know from Theorem 5.1 that $\kappa$ is nondegenerate. On the other hand, $L_0$ is orthogonal to all $L_\alpha$ $(\alpha \in \Phi)$, according to the proposition. If $z \in L_0$ is orthogonal to $L_0$ as well, then $\kappa(z, L) = 0$, forcing $z = 0$. □

## *8.2. Centralizer of H*

We shall need a fact from linear algebra, whose proof is trivial:

**Lemma.** *If* $x$, $y$ *are commuting endomorphisms of a finite dimensional vector space, with* $y$ *nilpotent, then* $xy$ *is nilpotent; in particular,* $Tr(xy) = 0$. □

**Proposition.** *Let* $H$ *be a maximal toral subalgebra of* $L$. *Then* $H = C_L(H)$.

*Proof.* We proceed in steps. Write $C = C_L(H)$.

(1) *C contains the semisimple and nilpotent parts of its elements.* To say that $x$ belongs to $C_L(H)$ is to say that ad $x$ maps the subspace $H$ of $L$ into the subspace 0. By Proposition 4.2, $(\text{ad } x)_s$ and $(\text{ad } x)_n$ have the same property. But by (5.4), $(\text{ad } x)_s = \text{ad } x_s$ and $(\text{ad } x)_n = \text{ad } x_n$.

(2) *All semisimple elements of C lie in H.* If $x$ is semisimple and centralizes $H$, then $H+\mathsf{F}x$ (which is obviously an abelian subalgebra of $L$) is toral: the sum of commuting semisimple elements is again semisimple (4.2). By maximality of $H$, $H+\mathsf{F}x = H$, so $x \in H$.

(3) *The restriction of* $\kappa$ *to H is nondegenerate.* Let $\kappa(h, H) = 0$ for some $h \in H$; we must show that $h = 0$. If $x \in C$ is nilpotent, then the fact that $[xH] = 0$ and the fact that ad $x$ is nilpotent together imply (by the above lemma) that $Tr(\text{ad } x \text{ ad } y) = 0$ for all $y \in H$, or $\kappa(x, H) = 0$. But then (1) and (2) imply that $\kappa(h, C) = 0$, whence $h = 0$ (the restriction of $\kappa$ to $C$ being nondegenerate by the Corollary to Proposition 8.1).

(4) *C is nilpotent.* If $x \in C$ is semisimple, then $x \in H$ by (2), and $\text{ad}_C\, x(=0)$ is certainly nilpotent. On the other hand, if $x \in C$ is nilpotent, then $\text{ad}_C\, x$ is a fortiori nilpotent. Now let $x \in C$ be arbitrary, $x = x_s+x_n$. Since both $x_s$, $x_n$

lie in $C$ by (1), $\mathrm{ad}_C\, x$ is the sum of commuting nilpotents and is therefore itself nilpotent. By Engel's Theorem, $C$ is nilpotent.

(5) $H \cap [CC] = 0$. Since $\kappa$ is associative and $[HC] = 0$, $\kappa(H, [CC]) = 0$. Now use (3).

(6) *$C$ is abelian.* Otherwise $[CC] \neq 0$. $C$ being nilpotent, by (4), $Z(C) \cap [CC] \neq 0$ (Lemma 3.3). Let $z \neq 0$ lie in this intersection. By (2) and (5), $z$ cannot be semisimple. Its nilpotent part $n$ is therefore nonzero and lies in $C$, by (1), hence also lies in $Z(C)$ by Proposition 4.2. But then our lemma implies that $\kappa(n, C) = 0$, contrary to Corollary 8.1.

(7) $C = H$. Otherwise $C$ contains a nonzero nilpotent element, $x$, by (1), (2). According to the lemma and (6), $\kappa(x, y) = Tr(\mathrm{ad}\, x\ \mathrm{ad}\, y) = 0$ for all $y \in C$, contradicting Corollary 8.1. ▯

**Corollary.** *The restriction of $\kappa$ to $H$ is nondegenerate.* ▯

The corollary allows us to identify $H$ with $H^*$: to $\phi \in H^*$ corresponds the (unique) element $t_\phi \in H$ satisfying $\phi(h) = \kappa(t_\phi, h)$ for all $h \in H$. In particular, $\Phi$ corresponds to the subset $\{t_\alpha; \alpha \in \Phi\}$ of $H$.

### *8.3. Orthogonality properties*

In this subsection we shall obtain more precise information about the root space decomposition, using the Killing form. We already saw (Proposition 8.1) that $\kappa(L_\alpha, L_\beta) = 0$ if $\alpha, \beta \in H^*$, $\alpha+\beta \neq 0$; in particular, $\kappa(H, L_\alpha) = 0$ for all $\alpha \in \Phi$, so that (Proposition 8.2) the restriction of $\kappa$ to $H$ is nondegenerate.

**Proposition.** (*a*) *$\Phi$ spans $H^*$.*

(*b*) *If $\alpha \in \Phi$, then $-\alpha \in \Phi$.*

(*c*) *Let $\alpha \in \Phi$, $x \in L_\alpha$, $y \in L_{-\alpha}$. Then $[xy] = \kappa(x, y)t_\alpha$ ($t_\alpha$ as in* (8.2)).

(*d*) *If $\alpha \in \Phi$, then $[L_\alpha L_{-\alpha}]$ is one dimensional, with basis $t_\alpha$.*

(*e*) $\alpha(t_\alpha) = \kappa(t_\alpha, t_\alpha) \neq 0$, *for $\alpha \in \Phi$.*

(*f*) *If $\alpha \in \Phi$ and $x_\alpha$ is any nonzero element of $L_\alpha$, then there exists $y_\alpha \in L_{-\alpha}$ such that $x_\alpha$, $y_\alpha$, $h_\alpha = [x_\alpha y_\alpha]$ span a three dimensional simple subalgebra of $L$ isomorphic to $\mathfrak{sl}(2, \mathsf{F})$ via* $x_\alpha \mapsto \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $y_\alpha \mapsto \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $h_\alpha \mapsto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

(*g*) $h_\alpha = \dfrac{2t_\alpha}{\kappa(t_\alpha, t_\alpha)}$; $h_\alpha = -h_{-\alpha}$.

*Proof.* (a) If $\Phi$ fails to span $H^*$, then (by duality) there exists nonzero $h \in H$ such that $\alpha(h) = 0$ for all $\alpha \in \Phi$. But this means that $[h, L_\alpha] = 0$ for all $\alpha \in \Phi$. Since $[hH] = 0$, this in turn forces $[hL] = 0$, or $h \in Z(L) = 0$, which is absurd.

(b) Let $\alpha \in \Phi$. If $-\alpha \notin \Phi$ (i.e., $L_{-\alpha} = 0$), then $\kappa(L_\alpha, L_\beta) = 0$ for all $\beta \in H^*$ (Proposition 8.1). Therefore $\kappa(L_\alpha, L) = 0$, contradicting the nondegeneracy of $\kappa$.

(c) Let $\alpha \in \Phi$, $x \in L_\alpha$, $y \in L_{-\alpha}$. Let $h \in H$ be arbitrary. The associativity of $\kappa$ implies: $\kappa(h, [xy]) = \kappa([hx], y) = \alpha(h)\kappa(x, y) = \kappa(t_\alpha, h)\kappa(x, y) =$

$\kappa(\kappa(x,y)t_\alpha, h) = \kappa(h, \kappa(x,y)t_\alpha)$. This says that $H$ is orthogonal to $[xy]-\kappa(x,y)t_\alpha$, forcing $[xy] = \kappa(x, y)t_\alpha$ (Corollary 8.2).

(d) Part (c) shows that $t_\alpha$ spans $[L_\alpha L_{-\alpha}]$, provided $[L_\alpha L_{-\alpha}] \neq 0$. Let $0 \neq x \in L_\alpha$. If $\kappa(x, L_{-\alpha}) = 0$, then $\kappa(x, L) = 0$ (cf. proof of (b)), which is absurd since $\kappa$ is nondegenerate. Therefore we can find $0 \neq y \in L_{-\alpha}$ for which $\kappa(x, y) \neq 0$. By (c), $[xy] \neq 0$.

(e) Suppose $\alpha(t_\alpha) = 0$, so that $[t_\alpha x] = 0 = [t_\alpha y]$ for all $x \in L_\alpha$, $y \in L_{-\alpha}$. As in (d), we can find such $x$, $y$ satisfying $\kappa(x, y) \neq 0$. Modifying one or the other by a scalar, we may as well assume that $\kappa(x, y) = 1$. Then $[xy] = t_\alpha$, by (c). It follows that the subspace $S$ of $L$ spanned by $x$, $y$, $t_\alpha$ is a three dimensional solvable algebra, $S \cong \text{ad}_L\, S \subset \mathfrak{gl}(L)$. In particular, $\text{ad}_L\, s$ is *nilpotent* for all $s \in [SS]$ (Corollary 4.1A), so $\text{ad}_L\, t_\alpha$ is both semisimple and nilpotent, i.e., $\text{ad}_L\, t_\alpha = 0$. This says that $t_\alpha \in Z(L) = 0$, contrary to choice of $t_\alpha$.

(f) Given $0 \neq x_\alpha \in L_\alpha$, find $y_\alpha \in L_{-\alpha}$ such that $\kappa(x_\alpha, y_\alpha) = \dfrac{2}{\kappa(t_\alpha, t_\alpha)}$. This is possible in view of (e) and the fact that $\kappa(x_\alpha, L_{-\alpha}) \neq 0$. Set $h_\alpha = 2t_\alpha/\kappa(t_\alpha, t_\alpha)$. Then $[x_\alpha y_\alpha] = h_\alpha$, by (c). Moreover, $[h_\alpha x_\alpha] = \dfrac{2}{\alpha(t_\alpha)}[t_\alpha x_\alpha] = \dfrac{2\alpha(t_\alpha)}{\alpha(t_\alpha)} x_\alpha = 2x_\alpha$, and similarly, $[h_\alpha y_\alpha] = -2y_\alpha$. So $x_\alpha$, $y_\alpha$, $h_\alpha$ span a three dimensional subalgebra of $L$ with the same multiplication table as $\mathfrak{sl}(2, \mathsf{F})$ (Example 2.1).

(g) Recall that $t_\alpha$ is defined by $\kappa(t_\alpha, h) = \alpha(h)$ $(h \in H)$. This shows that $t_\alpha = -t_{-\alpha}$, and in view of the way $h_\alpha$ is defined, the assertion follows. ▯

### *8.4. Integrality properties*

For each pair of roots $\alpha$, $-\alpha$ (Proposition 8.3(b)), let $S_\alpha \cong \mathfrak{sl}(2, \mathsf{F})$ be a subalgebra of $L$ constructed as in Proposition 8.3(f). Thanks to Weyl's Theorem and Theorem 7.2, we have a complete description of all (finite dimensional) $S_\alpha$-modules; in particular, we can describe $\text{ad}_L\, S_\alpha$.

Fix $\alpha \in \Phi$. Consider first the subspace $M$ of $L$ spanned by $H$ along with all root spaces of the form $L_{c\alpha}$ $(c \in \mathsf{F}^*)$. This is an $S_\alpha$-submodule of $L$, thanks to Proposition 8.1. The weights of $h_\alpha$ on $M$ are the integers 0 and $2c = c\alpha(h_\alpha)$ (for nonzero $c$ such that $L_{c\alpha} \neq 0$), in view of Theorem 7.2. In particular, all $c$ occurring here must be integral multiples of $1/2$. Now $S_\alpha$ acts trivially on Ker $\alpha$, a subspace of codimension one in $H$ complementary to $\mathsf{F}h_\alpha$, while on the other hand $S_\alpha$ is itself an irreducible $S_\alpha$-submodule of $M$. Taken together, Ker $\alpha$ and $S_\alpha$ exhaust the occurrences of the weight 0 for $h_\alpha$. Therefore, the only even weights occurring in $M$ are 0, $\pm 2$. This proves that $2\alpha$ is not a root, i.e., that *twice a root is never a root.* But then $(1/2)\alpha$ cannot be a root either, so 1 cannot occur as a weight of $h_\alpha$ in $M$. The Corollary of Theorem 7.2 implies that $M = H + S_\alpha$. In particular, *dim* $L_\alpha = 1$ (so $S_\alpha$ is uniquely determined as the subalgebra of $L$ generated by $L_\alpha$ and $L_{-\alpha}$), and *the only multiples of a root* $\alpha$ *which are roots are* $\pm\alpha$.

Next we examine how $S_\alpha$ acts on root spaces $L_\beta$, $\beta \neq \pm\alpha$. Set $K = \sum_{i \in \mathbf{Z}}$

$L_{\beta+i\alpha}$. According to the preceding paragraph, each root space is one dimensional and no $\beta+i\alpha$ can equal 0; so $K$ is an $S_\alpha$-submodule of $L$, with one dimensional weight spaces for the distinct integral weights $\beta(h_\alpha)+2i$ ($i \in \mathbf{Z}$ such that $\beta+i\alpha \in \Phi$). Obviously, not both 0 and 1 can occur as weights of this form, so the Corollary of Theorem 7.2 implies that $K$ is irreducible. The highest (resp. lowest) weight must be $\beta(h_\alpha)+2q$ (resp. $\beta(h_\alpha)-2r$) if $q$ (resp. $r$) is the largest integer for which $\beta+q\alpha$ (resp. $\beta-r\alpha$) is a root. Moreover, the weights on $K$ form an arithmetic progression with difference 2 (Theorem 7.2), which implies that the roots $\beta+i\alpha$ form a string (the **$\alpha$-string through $\beta$**) $\beta-r\alpha, \ldots, \beta, \ldots, \beta+q\alpha$. Notice too that $(\beta-r\alpha)(h_\alpha) = -(\beta+q\alpha)(h_\alpha)$, or $\beta(h_\alpha) = r-q$. Finally, observe that if $\alpha, \beta, \alpha+\beta \in \Phi$, then ad $L_\alpha$ maps $L_\beta$ onto $L_{\alpha+\beta}$ (Lemma 7.2), i.e., $[L_\alpha L_\beta] = L_{\alpha+\beta}$.

To summarize:

**Proposition.** (*a*) *$\alpha \in \Phi$ implies dim $L_\alpha = 1$. In particular, $S_\alpha = L_\alpha + L_{-\alpha} + H_\alpha$ ($H_\alpha = [L_\alpha L_{-\alpha}]$), and for given nonzero $x_\alpha \in L_\alpha$, there exists a unique $y_\alpha \in L_{-\alpha}$ satisfying $[x_\alpha y_\alpha] = h_\alpha$.*

(*b*) *If $\alpha \in \Phi$, the only scalar multiples of $\alpha$ which are roots are $\alpha$ and $-\alpha$.*

(*c*) *If $\alpha, \beta \in \Phi$, then $\beta(h_\alpha) \in \mathbf{Z}$, and $\beta-\beta(h_\alpha)\alpha \in \Phi$. (The numbers $\beta(h_\alpha)$ are called* **Cartan integers**.)

(*d*) *If $\alpha, \beta, \alpha+\beta \in \Phi$, then $[L_\alpha L_\beta] = L_{\alpha+\beta}$.*

(*e*) *Let $\alpha, \beta \in \Phi$, $\beta \neq \pm\alpha$. Let $r, q$ be (respectively) the largest integers for which $\beta-r\alpha$, $\beta+q\alpha$ are roots. Then all $\beta+i\alpha \in \Phi$ ($-r \leq i \leq q$), and $\beta(h_\alpha) = r-q$.*

(*f*) *$L$ is generated (as Lie algebra) by the root spaces $L_\alpha$.* □

## *8.5. Rationality properties. Summary*

$L$ is a semisimple Lie algebra (over the algebraically closed field F of characteristic 0), $H$ a maximal toral subalgebra, $\Phi \subset H^*$ the set of roots of $L$ (relative to $H$), $L = H + \coprod_{\alpha\in\Phi} L_\alpha$ the root space decomposition.

Since the restriction to $H$ of the Killing form is nondegenerate (Corollary 8.2), we may transfer the form to $H^*$, letting $(\gamma, \delta) = \kappa(t_\gamma, t_\delta)$ for all $\gamma, \delta \in H^*$. We know that $\Phi$ spans $H^*$ (Proposition 8.3(a)), so choose a basis $\alpha_1, \ldots, \alpha_\ell$ of $H^*$ consisting of roots. If $\beta \in \Phi$, we can then write $\beta$ uniquely as $\beta = \sum_{i=1}^{\ell} c_i\alpha_i$, where $c_i \in$ F. We claim that in fact $c_i \in \mathbf{Q}$. To see this, we use a little linear algebra. For each $j = 1, \ldots, \ell$, $(\beta, \alpha_j) = \sum_{i=1}^{\ell} c_i(\alpha_i, \alpha_j)$, so multiplying both sides by $2/(\alpha_j, \alpha_j)$ yields: $2(\beta, \alpha_j)/(\alpha_j, \alpha_j) = \sum_{i=1}^{\ell} \frac{2(\alpha_i, \alpha_j)}{(\alpha_j, \alpha_j)} c_i$. This may be viewed as a system of $\ell$ equations in $\ell$ unknowns $c_i$, with integral (in particular, rational) coefficients, thanks to Proposition 8.4(c). Since $(\alpha_1, \ldots, \alpha_\ell)$ is a basis of $H^*$, and the form is nondegenerate, the matrix $((\alpha_i, \alpha_j))_{1 \leq i,j \leq \ell}$ is nonsingular; so the same is true of the coefficient matrix of this system of equations. We conclude that the equations already possess a unique solution over $\mathbf{Q}$, thereby proving our claim.

We have just shown that the $\mathbf{Q}$-subspace $E_{\mathbf{Q}}$ of $H^*$ spanned by all the roots has $\mathbf{Q}$-dimension $\ell = \dim_F H^*$. Even more is true: Recall that for $\lambda, \mu \in H^*$, $(\lambda, \mu) = \kappa(t_\lambda, t_\mu) = \Sigma\alpha(t_\lambda)\alpha(t_\mu) = \Sigma(\alpha, \lambda)(\alpha, \mu)$, where the sum is over $\alpha \in \Phi$. In particular, for $\beta \in \Phi$, $(\beta, \beta) = \Sigma(\alpha, \beta)^2$. Dividing by $(\beta, \beta)^2$, we get $1/(\beta, \beta) = \Sigma(\alpha, \beta)^2/(\beta, \beta)^2$, the latter in $\mathbf{Q}$ because $2(\alpha, \beta)/(\beta, \beta) \in \mathbf{Z}$ by Proposition 8.4(c). Therefore $(\beta, \beta) \in \mathbf{Q}$, and in turn, $(\alpha, \beta) \in \mathbf{Q}$. It follows that all inner products of vectors in $E_{\mathbf{Q}}$ are rational, so we obtain a nondegenerate form on $E_{\mathbf{Q}}$. As above, $(\lambda, \lambda) = \Sigma(\alpha, \lambda)^2$, so that for $\lambda \in E_{\mathbf{Q}}$, $(\lambda, \lambda)$ is a sum of squares of rational numbers and hence is positive (unless $\lambda = 0$). Therefore, the form on $E_{\mathbf{Q}}$ is *positive definite*.

Now let E be the real vector space obtained by extending the base field from $\mathbf{Q}$ to $\mathbf{R}$: $E = \mathbf{R} \otimes_{\mathbf{Q}} E_{\mathbf{Q}}$. The form extends canonically to E and is positive definite, by the preceding remarks, i.e., E is a euclidean space. $\Phi$ contains a basis of E, and $\dim_{\mathbf{R}} E = \ell$. The following theorem summarizes the basic facts about $\Phi$: cf. Propositions 8.3(a) (b) and 8.4(b) (c).

**Theorem.** *$L$, $H$, $\Phi$, E as above. Then*:

(*a*) *$\Phi$ spans E, and 0 does not belong to $\Phi$.*

(*b*) *If $\alpha \in \Phi$ then $-\alpha \in \Phi$, but no other scalar multiple of $\alpha$ is a root.*

(*c*) *If $\alpha, \beta \in \Phi$, then $\beta - \dfrac{2(\beta, \alpha)}{(\alpha, \alpha)}\alpha \in \Phi$.*

(*d*) *If $\alpha, \beta \in \Phi$, then $\dfrac{2(\beta, \alpha)}{(\alpha, \alpha)} \in \mathbf{Z}$.* □

In the language of Chapter III, the theorem asserts that $\Phi$ is a **root system** in the real euclidean space E. We have therefore set up a correspondence $(L, H) \mapsto (\Phi, E)$. Pairs $(\Phi, E)$ will be completely classified in Chapter III. Later (Chapters IV and V) it will be seen that the correspondence here is actually 1–1, and that the apparent dependence of $\Phi$ on the choice of $H$ is not essential.

## Exercises

1. If $L$ is a classical linear Lie algebra of type $A_\ell$, $B_\ell$, $C_\ell$, or $D_\ell$ (see (1.2)), prove that the set of all diagonal matrices in $L$ is a maximal toral subalgebra, of dimension $\ell$. (Cf. Exercise 2.8.)
2. For each algebra in Exercise 1, determine the roots and root spaces. How are the various $h_\alpha$ expressed in terms of the basis for $H$ given in (1.2)?
3. If $L$ is of classical type, compute explicitly the restriction of the Killing form to the maximal toral subalgebra described in Exercise 1.
4. If $L = \mathfrak{sl}(2, F)$, prove that each maximal toral subalgebra is one dimensional.
5. If $L$ is semisimple, $H$ a maximal toral subalgebra, prove that $H$ is self-normalizing (i.e., $H = N_L(H)$).

6. Compute the basis of $\mathfrak{sl}(n, \mathsf{F})$ which is dual (via the Killing form) to the standard basis. (Cf. Exercise 5.5.)
7. Let $L$ be semisimple, $H$ a maximal toral subalgebra. If $h \in H$, prove that $C_L(h)$ is *reductive* (in the sense of Exercise 6.5). Prove that $H$ contains elements $h$ for which $C_L(h) = H$; for which $h$ in $\mathfrak{sl}(n, \mathsf{F})$ is this true?
8. For $\mathfrak{sl}(n, \mathsf{F})$ (and other classical algebras), calculate explicitly the root strings and Cartan integers. In particular, prove that all Cartan integers $2(\alpha, \beta)/(\beta, \beta)$, $\alpha \neq \pm\beta$, for $\mathfrak{sl}(n, \mathsf{F})$ are $0, \pm 1$.
9. Prove that every three dimensional semisimple Lie algebra has the same root system as $\mathfrak{sl}(2, \mathsf{F})$, hence is isomorphic to $\mathfrak{sl}(2, \mathsf{F})$.
10. Prove that no four, five or seven dimensional semisimple Lie algebras exist.
11. If $(\alpha, \beta) > 0$, prove that $\alpha - \beta \in \Phi$ ($\alpha, \beta \in \Phi$). Is the converse true?

## Notes

The use of maximal toral subalgebras rather than the more traditional (but equivalent) Cartan subalgebras is suggested by the parallel theory of semisimple algebraic groups: cf. Borel [1], Seligman [2], Winter [1].