Chapter I

Basic Concepts

In this chapter F *denotes an arbitrary (commutative) field.*

1. Definitions and first examples

1.1. The notion of Lie algebra

Lie algebras arise "in nature" as vector spaces of linear transformations endowed with a new operation which is in general neither commutative nor associative: $[x, y] = xy - yx$ (where the operations on the right side are the usual ones). It is possible to describe this kind of system abstractly in a few axioms.

Definition. A vector space L over a field F, with an operation $L \times L \to L$, denoted $(x, y) \mapsto [xy]$ and called the **bracket** or **commutator** of x and y, is called a **Lie algebra** over F if the following axioms are satisfied:

($L1$) The bracket operation is bilinear.
($L2$) $[xx] = 0$ for all x in L.
($L3$) $[x[yz]]+[y[zx]]+[z[xy]] = 0 \quad (x, y, z \in L)$.

Axiom ($L3$) is called the **Jacobi identity.** Notice that ($L1$) and ($L2$), applied to $[x+y, x+y]$, imply anticommutativity: ($L2'$) $[xy] = -[yx]$. (Conversely, if char $\mathsf{F} \neq 2$, it is clear that ($L2'$) will imply ($L2$).)

We say that two Lie algebras L, L' over F are **isomorphic** if there exists a vector space isomorphism $\phi\colon L \to L'$ satisfying $\phi([xy]) = [\phi(x)\phi(y)]$ for all x, y in L (and then ϕ is called an **isomorphism** of Lie algebras). Similarly, it is obvious how to define the notion of (Lie) **subalgebra** of L: A subspace K of L is called a subalgebra if $[xy] \in K$ whenever $x, y \in K$; in particular, K is a Lie algebra in its own right relative to the inherited operations. Note that any nonzero element $x \in L$ defines a one dimensional subalgebra $\mathsf{F}x$, with trivial multiplication, because of ($L2$).

In this book we shall be concerned almost exclusively with Lie algebras L whose underlying vector space is *finite dimensional* over F. *This will always be assumed, unless otherwise stated.* We hasten to point out, however, that certain infinite dimensional vector spaces and associative algebras over F will play a vital role in the study of representations (Chapters V–VII). We also mention, before looking at some concrete examples, that the axioms for a Lie algebra make perfectly good sense if L is only assumed to be a module over a commutative ring, but we shall not pursue this point of view here.

1.2. Linear Lie algebras

If V is a finite dimensional vector space over F, denote by End V the set of linear transformations $V \to V$. As a vector space over F, End V has dimension n^2 ($n = \dim V$), and End V is a ring relative to the usual product operation. Define a new operation $[x, y] = xy - yx$, called the **bracket** of x and y. With this operation End V becomes a Lie algebra over F: axioms ($L1$) and ($L2$) are immediate, while ($L3$) requires a brief calculation (which the reader is urged to carry out at this point). In order to distinguish this new algebra structure from the old associative one, we write $\mathfrak{gl}(V)$ for End V viewed as Lie algebra and call it the **general linear algebra** (because it is closely associated with the **general linear group** $GL(V)$ consisting of all invertible endomorphisms of V). When V is infinite dimensional, we shall also use the notation $\mathfrak{gl}(V)$ without further comment.

Any subalgebra of a Lie algebra $\mathfrak{gl}(V)$ is called a **linear Lie algebra**. The reader who finds matrices more congenial than linear transformations may prefer to fix a basis for V, thereby identifying $\mathfrak{gl}(V)$ with the set of all $n \times n$ matrices over F, denoted $\mathfrak{gl}(n, \mathsf{F})$. This procedure is harmless, and very convenient for making explicit calculations. For reference, we write down the multiplication table for $\mathfrak{gl}(n, \mathsf{F})$ relative to the standard basis consisting of the matrices e_{ij} (having 1 in the (i, j) position and 0 elsewhere). Since $e_{ij}e_{kl} = \delta_{jk}e_{il}$, it follows that:

$$(*) \qquad [e_{ij}, e_{kl}] = \delta_{jk}\, e_{il} - \delta_{li}\, e_{kj}.$$

Notice that the coefficients are all ± 1 or 0; in particular, all of them lie in the prime field of F.

Now for some further examples, which are central to the theory we are going to develop in this book. They fall into four families A_ℓ, B_ℓ, C_ℓ, D_ℓ ($\ell \geq 1$) and are called the **classical algebras** (because they correspond to certain of the classical linear Lie groups).

A_ℓ: Let $\dim V = \ell+1$. Denote by $\mathfrak{sl}(V)$, or $\mathfrak{sl}(\ell+1, \mathsf{F})$, the set of endomorphisms of V having trace zero. (Recall that the **trace** of a matrix is the sum of its diagonal entries; this is independent of choice of basis for V, hence makes sense for an endomorphism of V.) Since $Tr(xy) = Tr(yx)$, and $Tr(x+y) = Tr(x)+Tr(y)$, $\mathfrak{sl}(V)$ is a subalgebra of $\mathfrak{gl}(V)$, called the **special linear algebra** because of its connection with the **special linear group** $SL(V)$ of endomorphisms of det 1. What is its dimension? On the one hand $\mathfrak{sl}(V)$ is a proper subalgebra of $\mathfrak{gl}(V)$, hence of dimension at most $(\ell+1)^2-1$. On the other hand, we can exhibit this number of linearly independent matrices of trace zero: Take all e_{ij} ($i \neq j$), along with all $h_i = e_{ii} - e_{i+1,i+1}$ ($1 \leq i \leq \ell$), for a total of $\ell+(\ell+1)^2-(\ell+1)$ matrices. We shall always view this as the standard basis for $\mathfrak{sl}(\ell+1, \mathsf{F})$.

C_ℓ: Let $\dim V = 2\ell$, with basis $(v_1, \ldots, v_{2\ell})$. Define a nondegenerate skew-symmetric form f on V by the matrix $s = \begin{pmatrix} 0 & I_\ell \\ -I_\ell & 0 \end{pmatrix}$. (It can be shown that even dimensionality is a necessary condition for existence of a nondegenerate bilinear form satisfying $f(v, w) = -f(w, v)$.) Denote by $\mathfrak{sp}(V)$, or $\mathfrak{sp}(2\ell, \mathsf{F})$, the **symplectic algebra**, which by definition consists of all endomorphisms x of V satisfying $f(x(v), w) = -f(v, x(w))$. The reader can easily verify that $\mathfrak{sp}(V)$ is closed under the bracket operation. In matrix terms, the condition for $x = \begin{pmatrix} m & n \\ p & q \end{pmatrix}$ $(m, n, p, q \in \mathfrak{gl}(\ell, \mathsf{F}))$ to be symplectic is that $sx = -x^t s$ (x^t = transpose of x), i.e., that $n^t = n$, $p^t = p$, and $m^t = -q$. (This last condition forces $Tr(x) = 0$.) It is easy now to compute a basis for $\mathfrak{sp}(2\ell, \mathsf{F})$. Take the diagonal matrices $e_{ii}-e_{\ell+i,\ell+i}$ $(1 \leq i \leq \ell)$, ℓ in all. Add to these all $e_{ij}-e_{\ell+j,\ell+i}$ $(1 \leq i \neq j \leq \ell)$, $\ell^2-\ell$ in number. For n we use the matrices $e_{i,\ell+i}$ $(1 \leq i \leq \ell)$ and $e_{i,\ell+j}+e_{j,\ell+i}$ $(1 \leq i < j \leq \ell)$, a total of $\ell+\frac{1}{2}\ell(\ell-1)$, and similarly for the positions in p. Adding up, we find $\dim \mathfrak{sp}(2\ell, \mathsf{F}) = 2\ell^2+\ell$.

B_ℓ: Let $\dim V = 2\ell+1$ be odd, and take f to be the nondegenerate symmetric bilinear form on V whose matrix is $s = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & I_\ell \\ 0 & I_\ell & 0 \end{pmatrix}$. The **orthogonal algebra** $\mathfrak{o}(V)$, or $\mathfrak{o}(2\ell+1, \mathsf{F})$, consists of all endomorphisms of V satisfying $f(x(v), w) = -f(v, x(w))$ (the same requirement as for C_ℓ). If we partition x in the same form as s, say $x = \begin{pmatrix} a & b_1 & b_2 \\ c_1 & m & n \\ c_2 & p & q \end{pmatrix}$, then the condition $sx = -x^t s$ translates into the following set of conditions: $a = 0$, $c_1 = -b_2^t$, $c_2 = -b_1^t$, $q = -m^t$, $n^t = -n$, $p^t = -p$. (As in the case of C_ℓ, this shows that $Tr(x) = 0$.) For a basis, take first the ℓ diagonal matrices $e_{ii}-e_{\ell+i,\ell+i}$ $(2 \leq i \leq \ell+1)$. Add the 2ℓ matrices involving only row one and column one: $e_{1,\ell+i+1}-e_{i+1,1}$ and $e_{1,i+1}-e_{\ell+i+1,1}$ $(1 \leq i \leq \ell)$. Corresponding to $q = -m^t$, take (as for C_ℓ) $e_{i+1,j+1}-e_{\ell+j+1,\ell+i+1}$ $(1 \leq i \neq j \leq \ell)$. For n take $e_{i+1,\ell+j+1}-e_{j+1,\ell+i+1}$ $(1 \leq i < j \leq \ell)$, and for p, $e_{i+\ell+1,j+1}-e_{j+\ell+1,i+1}$ $(1 \leq j < i \leq \ell)$. The total number of basis elements is $2\ell^2+\ell$ (notice that this was also the dimension of C_ℓ).

D_ℓ: Here we obtain another **orthogonal algebra**. The construction is identical to that for B_ℓ, except that $\dim V = 2\ell$ is even and s has the simpler form $\begin{pmatrix} 0 & I_\ell \\ I_\ell & 0 \end{pmatrix}$. We leave it as an exercise for the reader to construct a basis and to verify that $\dim \mathfrak{o}(2\ell, \mathsf{F}) = 2\ell^2-\ell$ (Exercise 8).

We conclude this subsection by mentioning several other subalgebras of $\mathfrak{gl}(n, \mathsf{F})$ which play an important subsidiary role for us. Let $\mathfrak{t}(n, \mathsf{F})$ be the set of **upper triangular matrices** (a_{ij}), $a_{ij} = 0$ if $i > j$. Let $\mathfrak{n}(n, \mathsf{F})$ be the **strictly upper triangular matrices** ($a_{ij} = 0$ if $i \geq j$). Finally, let $\mathfrak{d}(n, \mathsf{F})$ be the set of all **diagonal matrices.** It is trivial to check that each of these is closed under the bracket. Notice also that $\mathfrak{t}(n, \mathsf{F}) = \mathfrak{d}(n, \mathsf{F})+\mathfrak{n}(n, \mathsf{F})$ (vector space direct sum), with $[\mathfrak{d}(n, \mathsf{F}), \mathfrak{n}(n, \mathsf{F})] = \mathfrak{n}(n, \mathsf{F})$, hence $[\mathfrak{t}(n, \mathsf{F}), \mathfrak{t}(n, \mathsf{F})] = \mathfrak{n}(n, \mathsf{F})$, cf. Exercise 5. (If H, K are subalgebras of L, $[H\,K]$ denotes the subspace of L spanned by commutators $[xy]$, $x \in H$, $y \in K$.)

1.3. Lie algebras of derivations

Some Lie algebras of linear transformations arise most naturally as derivations of algebras. By an **F-algebra** (not necessarily associative) we simply mean a vector space $\mathfrak{A}$ over F endowed with a bilinear operation $\mathfrak{A} \times \mathfrak{A} \to \mathfrak{A}$, usually denoted by juxtaposition (unless $\mathfrak{A}$ is a Lie algebra, in which case we always use the bracket). By a **derivation** of $\mathfrak{A}$ we mean a linear map $\delta: \mathfrak{A} \to \mathfrak{A}$ satisfying the familiar product rule $\delta(ab) = a\delta(b) + \delta(a)b$. It is easily checked that the collection Der $\mathfrak{A}$ of all derivations of $\mathfrak{A}$ is a vector subspace of End $\mathfrak{A}$. The reader should also verify that the commutator $[\delta, \delta']$ of two derivations is again a derivation (though the ordinary product need not be, cf. Exercise 11). So Der $\mathfrak{A}$ is a subalgebra of $\mathfrak{gl}(\mathfrak{A})$.

Since a Lie algebra L is an F-algebra in the above sense, Der L is defined. Certain derivations arise quite naturally, as follows. If $x \in L$, $y \mapsto [xy]$ is an endomorphism of L, which we denote ad x. In fact, ad $x \in$ Der L, because we can rewrite the Jacobi identity (using ($L2'$)) in the form: $[x[yz]] = [[xy]z] + [y[xz]]$. Derivations of this form are called **inner**, all others **outer**. It is of course perfectly possible to have ad $x = 0$ even when $x \neq 0$: this occurs in any one dimensional Lie algebra, for example. The map $L \to$ Der L sending x to ad x is called the **adjoint representation** of L; it plays a decisive role in all that follows.

Sometimes we have occasion to view x simultaneously as an element of L and of a subalgebra K of L. To avoid ambiguity, the notation $\text{ad}_L x$ or $\text{ad}_K x$ will be used to indicate that x is acting on L (respectively, K). For example, if x is a diagonal matrix, then $\text{ad}_{\mathfrak{d}(n,F)}(x) = 0$, whereas $\text{ad}_{\mathfrak{gl}(n,F)}(x)$ need not be zero.

1.4. Abstract Lie algebras

We have looked at some natural examples of linear Lie algebras. It is known that, in fact, every (finite dimensional) Lie algebra is isomorphic to some linear Lie algebra (theorems of Ado, Iwasawa). This will not be proved here (cf. Jacobson [1] Chapter VI, or Bourbaki [1]); however, it will be obvious at an early stage of the theory that the result is true for all cases we are interested in.

Sometimes it is desirable, however, to contemplate Lie algebras abstractly. For example, if L is an arbitrary finite dimensional vector space over F, we can view L as a Lie algebra by setting $[xy] = 0$ for all $x, y \in L$. Such an algebra, having trivial Lie multiplication, is called **abelian** (because in the linear case $[x, y] = 0$ just means that x and y commute). If L is any Lie algebra, with basis $x_1, \ldots, x_n$ it is clear that the entire multiplication table of L can be recovered from the **structure constants** a_{ij}^k which occur in the expressions $[x_i x_j] = \sum_{k=1}^{n} a_{ij}^k x_k$. Those for which $i \geq j$ can even be deduced from the others, thanks to ($L2$), ($L2'$). Turning this remark around, it is possible to define an abstract Lie algebra from scratch simply by specifying a set of structure constants. Naturally, not just any set of scalars $\{a_{ij}^k\}$ will do, but a moment's thought shows that it is enough to require the "obvious" identities, those implied by $(L2)$ and $(L3)$:

$$a_{ii}^k = 0 = a_{ij}^k + a_{ji}^k;$$

$$\sum_k (a_{ij}^k a_{k\ell}^m + a_{j\ell}^k a_{ki}^m + a_{\ell i}^k a_{kj}^m) = 0.$$

In practice, we shall have no occasion to construct Lie algebras in this artificial way. But, as an application of the abstract point of view, we can determine (up to isomorphism) all Lie algebras of dimension ≤ 2. In dimension 1 there is a single basis vector x, with multiplication table $[xx] = 0$ $(L2)$. In dimension 2, start with a basis x, y of L. Clearly, all products in L yield scalar multiples of $[xy]$. If these are all 0, then L is abelian. Otherwise, we can replace x in the basis by a vector spanning the one dimensional space of multiples of the original $[xy]$, and take y to be any other vector independent of the new x. Then $[xy] = ax$ $(a \neq 0)$. Replacing y by $a^{-1}y$, we finally get $[xy] = x$. Abstractly, therefore, at most one nonabelian L exists (the reader should check that $[xy] = x$ actually defines a Lie algebra).

Exercises

1. Let L be the real vector space $\mathbf{R}^3$. Define $[xy] = x \times y$ (cross product of vectors) for $x, y \in L$, and verify that L is a Lie algebra. Write down the structure constants relative to the usual basis of $\mathbf{R}^3$.
2. Verify that the following equations and those implied by $(L1)$ $(L2)$ define a Lie algebra structure on a three dimensional vector space with basis (x, y, z): $[xy] = z$, $[xz] = y$, $[yz] = 0$.
3. Let $x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ be an ordered basis for $\mathfrak{sl}(2, \mathsf{F})$. Compute the matrices of ad x, ad h, ad y relative to this basis.
4. Find a linear Lie algebra isomorphic to the nonabelian two dimensional algebra constructed in (1.4). [Hint: Look at the adjoint representation.]
5. Verify the assertions made in (1.2) about $\mathfrak{t}(n, \mathsf{F})$, $\mathfrak{d}(n, \mathsf{F})$, $\mathfrak{n}(n, \mathsf{F})$, and compute the dimension of each algebra, by exhibiting bases.
6. Let $x \in \mathfrak{gl}(n, \mathsf{F})$ have n *distinct* eigenvalues $a_1, \ldots, a_n$ in F. Prove that the eigenvalues of ad x are precisely the n^2 scalars $a_i - a_j$ $(1 \leq i, j \leq n)$, which of course need not be distinct.
7. Let $\mathfrak{s}(n, \mathsf{F})$ denote the **scalar matrices** (=scalar multiples of the identity) in $\mathfrak{gl}(n, \mathsf{F})$. If char F is 0 or else a prime not dividing n, prove that $\mathfrak{gl}(n, \mathsf{F}) = \mathfrak{sl}(n, \mathsf{F}) + \mathfrak{s}(n, \mathsf{F})$ (direct sum of vector spaces), with $[\mathfrak{s}(n, \mathsf{F}), \mathfrak{gl}(n, \mathsf{F})] = 0$.
8. Verify the stated dimension of D_ℓ.
9. When char $\mathsf{F} = 0$, show that each classical algebra $L = \mathsf{A}_\ell, \mathsf{B}_\ell, \mathsf{C}_\ell$, or D_ℓ is equal to $[LL]$. (This shows again that each algebra consists of trace 0 matrices.)

10. For small values of ℓ, isomorphisms occur among certain of the classical algebras. Show that A_1, B_1, C_1 are all isomorphic, while D_1 is the one dimensional Lie algebra. Show that B_2 is isomorphic to C_2, D_3 to A_3. What can you say about D_2?
11. Verify that the commutator of two derivations of an F-algebra is again a derivation, whereas the ordinary product need not be.
12. Let L be a Lie algebra over an algebraically closed field and let $x \in L$. Prove that the subspace of L spanned by the eigenvectors of ad x is a subalgebra.

2. Ideals and homomorphisms

2.1. Ideals

A subspace I of a Lie algebra L is called an **ideal** of L if $x \in L$, $y \in I$ together imply $[xy] \in I$. (Since $[xy] = -[yx]$, the condition could just as well be written: $[yx] \in I$.) Ideals play the role in Lie algebra theory which is played by normal subgroups in group theory and by two sided ideals in ring theory: they arise as kernels of homomorphisms (2.2).

Obviously 0 (the subspace consisting only of the zero vector) and L itself are ideals of L. A less trivial example is the **center** $Z(L) = \{z \in L | [xz] = 0 \text{ for all } x \in L\}$. Notice that L is abelian if and only if $Z(L) = L$. Another important example is the **derived algebra** of L, denoted $[LL]$, which is analogous to the commutator subgroup of a group. It consists of all linear combinations of commutators $[xy]$, and is clearly an ideal.

Evidently L is abelian if and only if $[LL] = 0$. At the other extreme, a study of the multiplication table for $L = \mathfrak{sl}(n, \mathsf{F})$ in (1.2) ($n \neq 2$ if char $\mathsf{F} = 2$) shows that $L = [LL]$ in this case, and similarly for other classical linear Lie algebras (Exercise 1.9).

If I, J are two ideals of a Lie algebra L, then $I+J = \{x+y | x \in I, y \in J\}$ is also an ideal. Similarly, $[IJ] = \{\Sigma x_i y_i | x_i \in I, y_i \in J\}$ is an ideal; the derived algebra $[LL]$ is just a special case of this construction.

It is natural to analyze the structure of a Lie algebra by looking at its ideals. If L has no ideals except itself and 0, and if moreover $[LL] \neq 0$, we call L **simple**. The condition $[LL] \neq 0$ (i.e., L nonabelian) is imposed in order to avoid giving undue prominence to the one dimensional algebra. Clearly, L simple implies $Z(L) = 0$ and $L = [LL]$.

Example. Let $L = \mathfrak{sl}(2, \mathsf{F})$, char $\mathsf{F} \neq 2$. Take as standard basis for L the three matrices (cf. (1.2)): $x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. The multiplication table is then completely determined by the equations: $[xy] = h$, $[hx] = 2x$, $[hy] = -2y$. (Notice that x, y, h are eigenvectors for ad h, corresponding to the eigenvalues $2, -2, 0$. Since char $\mathsf{F} \neq 2$, these eigenvalues are distinct.) If $I \neq 0$ is an ideal of L, let $ax+by+ch$ be an arbitrary nonzero element of I. Applying ad x twice, we get $-2bx \in I$, and applying ad y twice, we get $-2ay \in I$. Therefore, if a or b is nonzero, I contains either y or x (char $\mathsf{F} \neq 2$), and then, clearly, $I = L$ follows. On the other hand, if $a = b = 0$, then $0 \neq ch \in I$, so $h \in I$, and again $I = L$ follows. We conclude that *L is simple*.

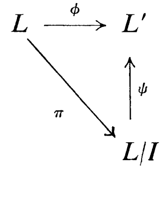
In case a Lie algebra L is not simple (and not one dimensional) it is possible to "factor out" a nonzero proper ideal I and thereby obtain a Lie algebra of smaller dimension. The construction of a **quotient algebra** L/I (I an ideal of L) is formally the same as the construction of a quotient ring: as vector space L/I is just the quotient space, while its Lie multiplication is defined by $[x+I, y+I] = [xy]+I$. This is unambiguous, since if $x+I = x'+I$, $y+I = y'+I$, then we have $x' = x+u$ ($u \in I$), $y' = y+v$ ($v \in I$), whence $[x'y'] = [xy]+([uy]+[xv]+[uv])$, and therefore $[x'y']+I = [xy]+I$, since the terms in parentheses all lie in I.

For later use we mention a couple of related notions, analogous to those which arise in group theory. The **normalizer** of a subalgebra (or just subspace) K of L is defined by $N_L(K) = \{x \in L | [xK] \subset K\}$. By the Jacobi identity, $N_L(K)$ is a subalgebra of L; it may be described verbally as the largest subalgebra of L which includes K as an ideal (in case K is a subalgebra to begin with). If $K = N_L(K)$, we call K **self-normalizing**; some important examples of this behavior will emerge later. The **centralizer** of a subset X of L is $C_L(X) = \{x \in L | [xX] = 0\}$. Again by the Jacobi identity, $C_L(X)$ is a subalgebra of L. For example, $C_L(L) = Z(L)$.

2.2. Homomorphisms and representations

The definition should come as no surprise. A linear transformation $\phi: L \to L'$ (L, L' Lie algebras over F) is called a **homomorphism** if $\phi([xy]) = [\phi(x)\phi(y)]$, for all $x, y \in L$. ϕ is called a **monomorphism** if Ker $\phi = 0$, an **epimorphism** if Im $\phi = L'$, an **isomorphism** (as in (1.1)) if it is both mono- and epi-. The first interesting observation to make is that *Ker ϕ is an ideal of L*: indeed, if $\phi(x) = 0$, and if $y \in L$ is arbitrary, then $\phi([xy]) = [\phi(x)\phi(y)] = 0$. It is also apparent that Im ϕ is a subalgebra of L'. As in other algebraic theories, there is a natural 1–1 correspondence between homomorphisms and ideals: to ϕ is associated Ker ϕ, and to an ideal I is associated the **canonical map** $x \mapsto x+I$ of L onto L/I. It is left as an easy exercise for the reader to verify the standard homomorphism theorems:

Proposition. (*a*) *If ϕ: $L \to L'$ is a homomorphism of Lie algebras, then $L/Ker\ \phi \cong Im\ \phi$. If I is any ideal of L included in Ker ϕ, there exists a unique homomorphism ψ: $L/I \to L'$ making the following diagram commute (π = canonical map)*:



(*b*) *If I and J are ideals of L such that $I \subset J$, then J/I is an ideal of L/I and $(L/I)/(J/I)$ is naturally isomorphic to L/J.*

(*c*) *If I, J are ideals of L, there is a natural isomorphism between $(I+J)/J$ and $I/(I \cap J)$.* ☐

A **representation** of a Lie algebra L is a homomorphism $\phi: L \to \mathfrak{gl}(V)$ (V = vector space over F). Although we require L to be finite dimensional, it is useful to allow V to be of arbitrary dimension: $\mathfrak{gl}(V)$ makes sense in any case. However, for the time being the only important example to keep in mind is the adjoint representation ad: $L \to \mathfrak{gl}(L)$ introduced in (1.3), which sends x to ad x, where ad $x(y) = [xy]$. (The image of ad is in Der $L \subset \mathfrak{gl}(L)$, but this does not concern us at the moment.) It is clear that ad is a linear transformation. To see that it preserves the bracket, we calculate:

$$
\begin{aligned}
[\text{ad}\, x, \text{ad}\, y]\,(z) &= \text{ad}\, x\, \text{ad}\, y(z) - \text{ad}\, y\, \text{ad}\, x(z) \\
&= \text{ad}\, x([yz]) - \text{ad}\, y([xz]) \\
&= [x[yz]] - [y[xz]] \\
&= [x[yz]] + [[xz]y] && (L2') \\
&= [[xy]z] && (L3) \\
&= \text{ad}\, [xy]\,(z).
\end{aligned}
$$

What is the kernel of ad? It consists of all $x \in L$ for which ad $x = 0$, i.e., for which $[xy] = 0$ (all $y \in L$). So Ker ad $= Z(L)$. This already has an interesting consequence: If L is simple, then $Z(L) = 0$, so that ad: $L \to \mathfrak{gl}(L)$ is a monomorphism. This means that *any simple Lie algebra is isomorphic to a linear Lie algebra.*

2.3. *Automorphisms*

An **automorphism** of L is an isomorphism of L onto itself. Aut L denotes the group of all such. Important examples occur when L is a linear Lie algebra $\subset \mathfrak{gl}(V)$. If $g \in GL(V)$ is any invertible endomorphism of V, and if moreover $gLg^{-1} = L$, then it is immediate that the map $x \mapsto gxg^{-1}$ is an automorphism of L. For instance, if $L = \mathfrak{gl}(V)$ or even $\mathfrak{sl}(V)$, the second condition is automatic, so we obtain in this way a large collection of automorphisms. (Cf. Exercise 12.)

Now specialize to the case: char F = 0. Suppose $x \in L$ is an element for which ad x is **nilpotent**, i.e., $(\text{ad}\, x)^k = 0$ for some $k > 0$. Then the usual exponential power series for a linear transformation over **C** makes sense over F, because it has only finitely many terms: $\exp(\text{ad}\, x) = 1 + \text{ad}\, x + (\text{ad}\, x)^2/2! + (\text{ad}\, x)^3/3! + \ldots + (\text{ad}\, x)^{k-1}/(k-1)!$. We claim that $\exp(\text{ad}\, x) \in \text{Aut}\, L$. More generally, this is true if ad x is replaced by an arbitrary nilpotent derivation δ of L. For this, use the familiar *Leibniz rule*:

$$\frac{\delta^n}{n!}(xy) = \sum_{i=0}^{n} (1/i!)\,(\delta^i x)\,(1/(n-i)!)\,(\delta^{n-i} y).$$

Then calculate as follows: (say $\delta^k = 0$)

$$\begin{aligned}
\exp \delta(x) \exp \delta(y) &= \left(\sum_{i=0}^{k-1}\left(\frac{\delta^i x}{i!}\right)\right)\left(\sum_{j=0}^{k-1}\left(\frac{\delta^j y}{j!}\right)\right) \\
&= \sum_{n=0}^{2k-2}\left(\sum_{i=0}^{n}\left(\frac{\delta^i x}{i!}\right)\left(\frac{\delta^{n-i} y}{(n-i)!}\right)\right) \\
&= \sum_{n=0}^{2k-2}\frac{\delta^n(xy)}{n!} && \text{(Leibniz)} \\
&= \sum_{n=0}^{k-1}\frac{\delta^n(xy)}{n!} && (\delta^k = 0) \\
&= \exp \delta(xy).
\end{aligned}$$

The fact that $\exp \delta$ is invertible follows (in the usual way) by exhibiting the explicit inverse $1-\eta+\eta^2-\eta^3+\ldots\pm\eta^{k-1}$, $\exp \delta = 1+\eta$.

An automorphism of the form $\exp(\operatorname{ad} x)$, $\operatorname{ad} x$ nilpotent, is called **inner**; more generally, the subgroup of Aut L generated by these is denoted Int L and its elements called inner automorphisms. It is a *normal* subgroup: If $\phi \in \operatorname{Aut} L$, $x \in L$, then $\phi(\operatorname{ad} x)\phi^{-1} = \operatorname{ad} \phi(x)$, whence $\phi \exp(\operatorname{ad} x)\phi^{-1} = \exp(\operatorname{ad} \phi(x))$.

For example, let $L = \mathfrak{sl}(2, \mathsf{F})$, with standard basis (x, y, h). Define $\sigma = \exp \operatorname{ad} x \cdot \exp \operatorname{ad}(-y) \cdot \exp \operatorname{ad} x$ (so $\sigma \in \operatorname{Int} L$). It is easy to compute the effect of σ on the basis (Exercise 10): $\sigma(x) = -y$, $\sigma(y) = -x$, $\sigma(h) = -h$. In particular, σ has order 2. Notice that $\exp x$, $\exp(-y)$ are well defined elements of $SL(2, \mathsf{F})$, the group of 2×2 matrices of det 1, conjugation by which leaves L invariant (as noted at the start of this subsection), so the product $s = (\exp x)(\exp -y)(\exp x)$ induces an automorphism $z \mapsto szs^{-1}$ of L. A quick calculation shows that $s = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and that conjugating by s has precisely the same effect on L as applying σ.

The phenomenon just observed is not accidental: If $L \subset \mathfrak{gl}(V)$ is an arbitrary linear Lie algebra (char $\mathsf{F} = 0$), and $x \in L$ is nilpotent, then we claim that

(*) $\qquad (\exp x)\, y\, (\exp x)^{-1} = \exp \operatorname{ad} x\,(y)$ for all $y \in L$.

To prove this, notice that $\operatorname{ad} x = \lambda_x + \rho_{-x}$, where λ_x, ρ_x denote left and right multiplication by x in the ring End V (these commute, of course, and are nilpotent). Then the usual rules of exponentiation show that $\exp \operatorname{ad} x = \exp(\lambda_x + \rho_{-x}) = \exp \lambda_x . \exp \rho_{-x} = \lambda_{\exp x} \cdot \rho_{\exp(-x)}$, which implies (*).

Exercises

1. Prove that the set of all inner derivations $\operatorname{ad} x$, $x \in L$, is an ideal of Der L.
2. Show that $\mathfrak{sl}(n, \mathsf{F})$ is precisely the derived algebra of $\mathfrak{gl}(n, \mathsf{F})$ (cf. Exercise 1.9).

3. Prove that the center of $\mathfrak{gl}(n, \mathsf{F})$ equals $\mathfrak{s}(n, \mathsf{F})$ (the scalar matrices). Prove that $\mathfrak{sl}(n, \mathsf{F})$ has center 0, unless char F divides n, in which case the center is $\mathfrak{s}(n, \mathsf{F})$.
4. Show that (up to isomorphism) there is a unique Lie algebra over F of dimension 3 whose derived algebra has dimension 1 and lies in $Z(L)$.
5. Suppose dim $L = 3$, $L = [LL]$. Prove that L must be simple. [Observe first that any homomorphic image of L also equals its derived algebra.] Recover the simplicity of $\mathfrak{sl}(2, \mathsf{F})$, char $\mathsf{F} \neq 2$.
6. Prove that $\mathfrak{sl}(3, \mathsf{F})$ is simple, unless char $\mathsf{F} = 3$ (cf. Exercise 3). [Use the standard basis h_1, h_2, e_{ij} $(i \neq j)$. If $I \neq 0$ is an ideal, then I is the direct sum of eigenspaces for ad h_1 or ad h_2; compare the eigenvalues of ad h_1, ad h_2 acting on the e_{ij}.]
7. Prove that $\mathfrak{t}(n, \mathsf{F})$ and $\mathfrak{d}(n, \mathsf{F})$ are self-normalizing subalgebras of $\mathfrak{gl}(n, \mathsf{F})$, whereas $\mathfrak{n}(n, \mathsf{F})$ has normalizer $\mathfrak{t}(n, \mathsf{F})$.
8. Prove that in each classical linear Lie algebra (1.2), the set of diagonal matrices is a self-normalizing subalgebra, when char $\mathsf{F} = 0$.
9. Prove Proposition 2.2.
10. Let σ be the automorphism of $\mathfrak{sl}(2, \mathsf{F})$ defined in (2.3). Verify that $\sigma(x) = -y$, $\sigma(y) = -x$, $\sigma(h) = -h$.
11. If $L = \mathfrak{sl}(n, \mathsf{F})$, $g \in GL(n, \mathsf{F})$, prove that the map of L to itself defined by $x \mapsto -gx^tg^{-1}$ (x^t = transpose of x) belongs to Aut L. When $n = 2$, g = identity matrix, prove that this automorphism is inner.
12. Let L be an orthogonal Lie algebra (type B_ℓ or D_ℓ). If g is an **orthogonal** matrix, in the sense that g is invertible and $g^tsg = s$, prove that $x \mapsto gxg^{-1}$ defines an automorphism of L.

3. Solvable and nilpotent Lie algebras

3.1. Solvability

It is natural to study a Lie algebra L via its ideals. In this section we exploit the formation of derived algebras. First, define a sequence of ideals of L (the **derived series**) by $L^{(0)} = L$, $L^{(1)} = [LL]$, $L^{(2)} = [L^{(1)}L^{(1)}], \ldots, L^{(i)} = [L^{(i-1)}L^{(i-1)}]$. Call L **solvable** if $L^{(n)} = 0$ for some n. For example, abelian implies solvable, whereas simple algebras are definitely nonsolvable.

An example which turns out to be rather general is the algebra $\mathfrak{t}(n, \mathsf{F})$ of upper triangular matrices, which was introduced in (1.2). The obvious basis for $\mathfrak{t}(n, \mathsf{F})$ consists of the matrix units e_{ij} for which $i \leq j$; the dimension is $1+2+\ldots+n = n(n+1)/2$. To show that $L = \mathfrak{t}(n, \mathsf{F})$ is solvable we compute explicitly its derived series, using the formula for commutators in (1.2). In particular, we have $[e_{ii}, e_{il}] = e_{il}$ for $i < l$, which shows that $\mathfrak{n}(n, \mathsf{F}) \subset [LL]$, where $\mathfrak{n}(n, \mathsf{F})$ is the subalgebra of upper triangular nilpotent matrices. Since $\mathfrak{t}(n, \mathsf{F}) = \mathfrak{d}(n, \mathsf{F}) + \mathfrak{n}(n, \mathsf{F})$, and since $\mathfrak{d}(n, \mathsf{F})$ is abelian, we conclude that $\mathfrak{n}(n, \mathsf{F})$ is equal to the derived algebra of L (cf. Exercise 1.5). Working next inside the algebra $\mathfrak{n}(n, \mathsf{F})$, we have a natural notion of "level" for e_{ij}, namely

$j-i$. In the formula for commutators, assume that $i < j$, $k < l$. Without losing any products we may also require $i \neq l$. Then $[e_{ij}, e_{kl}] = e_{il}$ (if $j = k$) or 0 (otherwise). In particular, each e_{il} is commutator of two matrices whose levels add up to that of e_{il}. We conclude that $L^{(2)}$ is spanned by those e_{ij} of level ≥ 2, $L^{(i)}$ by those of level $\geq 2^{i-1}$. Finally, it is clear that $L^{(i)} = 0$ whenever $2^{i-1} > n-1$.

Next we assemble a few simple observations about solvability.

Proposition. *Let L be a Lie algebra.*

(*a*) *If L is solvable, then so are all subalgebras and homomorphic images of L.*

(*b*) *If I is a solvable ideal of L such that L/I is solvable, then L itself is solvable.*

(*c*) *If I, J are solvable ideals of L, then so is I+J.*

Proof. (a) From the definition, if K is a subalgebra of L, then $K^{(i)} \subset L^{(i)}$. Similarly, if $\phi: L \to M$ is an epimorphism, an easy induction on i shows that $\phi(L^{(i)}) = M^{(i)}$.

(b) Say $(L/I)^{(n)} = 0$. Applying part (a) to the canonical homomorphism $\pi: L \to L/I$, we get $\pi(L^{(n)}) = 0$, or $L^{(n)} \subset I = \operatorname{Ker} \pi$. Now if $I^{(m)} = 0$, the obvious fact that $(L^{(i)})^{(j)} = L^{(i+j)}$ implies that $L^{(n+m)} = 0$ (apply proof of part (a) to the situation $L^{(n)} \subset I$).

(c) One of the standard homomorphism theorems (Proposition 2.2 (c)) yields an isomorphism between $(I+J)/J$ and $I/(I \cap J)$. As a homomorphic image of I, the right side is solvable, so $(I+J)/J$ is solvable. Then so is $I+J$, by part (b) applied to the pair $I+J$, J. □

As a first application, let L be an arbitrary Lie algebra and let S be a maximal solvable ideal (i.e., one included in no larger solvable ideal). If I is any other solvable ideal of L, then part (c) of the Proposition forces $S+I=S$ (by maximality), or $I \subset S$. This proves the existence of a unique maximal solvable ideal, called the **radical** of L and denoted $\operatorname{Rad} L$. In case $\operatorname{Rad} L = 0$, L is called **semisimple**. For example, *a simple algebra is semisimple:* L has no ideals except itself and 0, and L is nonsolvable. Also, $L = 0$ is semisimple. Notice that for arbitrary L, $L/\operatorname{Rad} L$ is semisimple (use part (b) of the proposition). The study of semisimple Lie algebras (char $\mathsf{F} = 0$) will occupy most of this book. (But certain solvable subalgebras will also be needed along the way.)

3.2. Nilpotency

The definition of solvability imitates the corresponding notion in group theory, which goes back to Abel and Galois. By contrast, the notion of nilpotent group is more recent, and is modeled on the corresponding notion for Lie algebras. Define a sequence of ideals of L (the **descending central series**, also called the **lower central series**) by $L^0 = L$, $L^1 = [LL]$ $(=L^{(1)})$, $L^2 = [LL^1], \ldots, L^i = [L\,L^{i-1}]$. L is called **nilpotent** if $L^n = 0$ for some n. For example, any abelian algebra is nilpotent. Clearly, $L^{(i)} \subset L^i$ for all i, so nilpotent algebras are solvable. The converse is false, however. Consider again $L = \mathfrak{t}(n, \mathsf{F})$. Our discussion in (3.1) showed that $L^{(1)} = L^1$ is $\mathfrak{n}(n, \mathsf{F})$, and also that $L^2 = [L\,L^1] = L^1$, so $L^i = L^1$ for all $i \geq 1$. On the other hand, it is easy to see that $M = \mathfrak{n}(n, \mathsf{F})$ is nilpotent: M^1 is spanned by those e_{ij} of level ≥ 2, M^2 by those of level $\geq 3, \ldots, M^i$ by those of level $\geq i+1$.

Proposition. *Let L be a Lie algebra.*

(*a*) *If L is nilpotent, then so are all subalgebras and homomorphic images of L.*

(*b*) *If $L/Z(L)$ is nilpotent, then so is L.*

(c) *If L is nilpotent and nonzero, then $Z(L) \neq 0$.*

Proof. (a) Imitate the proof of Proposition 3.1 (a).

(b) Say $L^n \subset Z(L)$, then $L^{n+1} = [LL^n] \subset [LZ(L)] = 0$.

(c) The last nonzero term of the descending central series is central. ▯

The condition for L to be nilpotent can be rephrased as follows: For some n (depending only on L), $\operatorname{ad} x_1 \operatorname{ad} x_2 \ldots \operatorname{ad} x_n(y) = 0$ for all $x_i, y \in L$. In particular, $(\operatorname{ad} x)^n = 0$ for all $x \in L$. Now if L is any Lie algebra, and $x \in L$, we call x **ad-nilpotent** if ad x is a nilpotent endomorphism. Using this language, our conclusion can be stated: If L is nilpotent, then all elements of L are ad-nilpotent. It is a pleasant surprise to find that the converse is also true.

Theorem (Engel). *If all elements of L are ad-nilpotent, then L is nilpotent.*

The proof will be given in the next subsection. Using Engel's Theorem, it is easy to prove that $\mathfrak{n}(n, \mathsf{F})$ is nilpotent, without actually calculating the descending central series. We need only apply the following simple lemma.

Lemma. *Let $x \in \mathfrak{gl}(V)$ be a nilpotent endomorphism. Then ad x is also nilpotent.*

Proof. As in (2.3), we may associate to x two endomorphisms of End V, left and right translation: $\lambda_x(y) = xy$, $\rho_x(y) = yx$, which are nilpotent because x is. Moreover λ_x and ρ_x obviously commute. In any ring (here End (End V)) the sum or difference of two commuting nilpotents is again nilpotent (why?), so $\operatorname{ad} x = \lambda_x - \rho_x$ is nilpotent. ▯

A word of warning: It is easy for a matrix to be ad-nilpotent in $\mathfrak{gl}(n, \mathsf{F})$ without being nilpotent. (The identity matrix is an example.) The reader should keep in mind two contrasting types of nilpotent linear Lie algebras: $\mathfrak{d}(n, \mathsf{F})$ and $\mathfrak{n}(n, \mathsf{F})$.

3.3. Proof of Engel's Theorem

Engel's Theorem (3.2) will be deduced from the following result, which is of interest in its own right. Recall that a single nilpotent linear transformation always has at least one eigenvector, corresponding to its unique eigenvalue 0. This is just the case $\dim L = 1$ of the following theorem.

Theorem. *Let L be a subalgebra of $\mathfrak{gl}(V)$, V finite dimensional. If L consists of nilpotent endomorphisms and $V \neq 0$, then there exists nonzero $v \in V$ for which $L.v = 0$.*

Proof. Use induction on dim L, the case dim $L = 0$ (or dim $L = 1$) being obvious. Suppose $K \neq L$ is any subalgebra of L. According to Lemma 3.2, K acts (via ad) as a Lie algebra of nilpotent linear transformations on the vector space L, hence also on the vector space L/K. Because dim $K <$ dim L, the induction hypothesis guarantees existence of a vector $x+K \neq K$ in L/K killed by the image of K in $\mathfrak{gl}(L/K)$. This just means that $[yx] \in K$ for all $y \in K$, whereas $x \notin K$. In other words, K is properly included in $N_L(K)$ (the normalizer of K in L, see (2.1)).

Now take K to be a maximal proper subalgebra of L. The preceding argument forces $N_L(K) = L$, i.e., K is an *ideal* of L. If dim L/K were greater than one, then the inverse image in L of a one dimensional subalgebra of L/K (which always exists) would be a proper subalgebra properly containing K, which is absurd; therefore, K has codimension one. This allows us to write $L = K+\mathsf{F}z$ for any $z \in L-K$.

By induction, $W = \{v \in V | K.v = 0\}$ is nonzero. Since K is an ideal, W is stable under L: $x \in L$, $y \in K$, $w \in W$ imply $yx.w = xy.w-[x, y].w = 0$. Choose $z \in L-K$ as above, so the nilpotent endomorphism z (acting now on the subspace W) has an eigenvector, i.e., there exists nonzero $v \in W$ for which $z.v = 0$. Finally, $L.v = 0$, as desired. ▯

Proof of Engel's Theorem. We are given a Lie algebra L all of whose elements are ad-nilpotent; therefore, the algebra ad $L \subset \mathfrak{gl}(L)$ satisfies the hypothesis of Theorem 3.3. (We can assume $L \neq 0$.) Conclusion: There exists $x \neq 0$ in L for which $[Lx] = 0$, i.e., $Z(L) \neq 0$. Now $L/Z(L)$ evidently consists of ad-nilpotent elements and has smaller dimension than L. Using induction on dim L, we find that $L/Z(L)$ is nilpotent. Part (b) of Proposition 3.2 then implies that L itself is nilpotent. ▯

There is a useful corollary (actually, an equivalent version) of Theorem 3.3, which shows how "typical" $\mathfrak{n}(n, \mathsf{F})$ is. First a definition: If V is a finite dimensional vector space (say dim $V = n$), a **flag** in V is a chain of subspaces $0 = V_0 \subset V_1 \subset \ldots \subset V_n = V$, dim $V_i = i$. If $x \in$ End V, we say x stabilizes (or leaves invariant) this flag provided $x.V_i \subset V_i$ for all i.

Corollary. *Under the hypotheses of the theorem there exists a flag (V_i) in V stable under L, with $x.V_i \subset V_{i-1}$ for all i. In other words, there exists a basis of V relative to which the matrices of L are all in $\mathfrak{n}(n, \mathsf{F})$.*

Proof. Begin with any nonzero $v \in V$ killed by L, the existence of which is assured by the theorem . Set $V_1 = \mathsf{F}v$. Let $W = V/V_1$, and observe that the induced action of L on W is also by nilpotent endomorphisms. By induction on dim V, W has a flag stabilized by L, whose inverse image in V does the trick. ▯

To conclude this section, we mention a typical application of Theorem 3.3, which will be needed later on.

Lemma. *Let L be nilpotent, K an ideal of L. Then if $K \neq 0$, $K \cap Z(L) \neq 0$. (In particular, $Z(L) \neq 0$; cf. Proposition* 3.2(*c*).)

Proof. L acts on K via the adjoint representation, so Theorem 3.3 yields nonzero $x \in K$ killed by L, i.e., $[Lx] = 0$, so $x \in K \cap Z(L)$. ▯

Exercises

1. Let I be an ideal of L. Then each member of the derived series or descending central series of I is also an ideal of L.
2. Prove that L is solvable if and only if there exists a chain of subalgebras $L = L_0 \supset L_1 \supset L_2 \supset \ldots \supset L_k = 0$ such that L_{i+1} is an ideal of L_i and such that each quotient L_i/L_{i+1} is abelian.
3. Let char $\mathsf{F} = 2$. Prove that $\mathfrak{sl}(2, \mathsf{F})$ is nilpotent.
4. Prove that L is solvable (resp. nilpotent) if and only if ad L is solvable (resp. nilpotent).
5. Prove that the nonabelian two dimensional algebra constructed in (1.4) is solvable but not nilpotent. Do the same for the algebra in Exercise 1.2.
6. Prove that the sum of two nilpotent ideals of a Lie algebra L is again a nilpotent ideal. Therefore, L possesses a unique maximal nilpotent ideal. Determine this ideal for each algebra in Exercise 5.
7. Let L be nilpotent, K a proper subalgebra of L. Prove that $N_L(K)$ includes K properly.
8. Let L be nilpotent. Prove that L has an ideal of codimension 1.
9. Prove that every nilpotent Lie algebra L has an outer derivation (see (1.3)), as follows: Write $L = K + \mathsf{F}x$ for some ideal K of codimension one (Exercise 8). Then $C_L(K) \neq 0$ (why?). Choose n so that $C_L(K) \subset L^n$, $C_L(K) \not\subset L^{n+1}$, and let $z \in C_L(K) - L^{n+1}$. Then the linear map δ sending K to 0, x to z, is an outer derivation.
10. Let L be a Lie algebra, K an ideal of L such that L/K is nilpotent and such that ad $x|_K$ is nilpotent for all $x \in L$. Prove that L is nilpotent.