4

Lie Algebras and Lie Groups

Summary. In this chapter we will discuss topics on differential geometry. In the spirit of the book, the proofs will be restricted to the basic ideas. Our goal is twofold: to explain the meaning of polarization operators as part of Lie algebra theory and to introduce the basic facts of this theory. The theory of Lie algebras is presented extensively in various books, as well as the theory of Lie groups (cf. [J1], [J2], [Ho], [Kn], [Se1], [Se2], [B1], [B2], [B3], [Wa]).

We assume that the reader is familiar with the basic definitions of differential geometry.

1 Lie Algebras and Lie Groups

1.1 Lie Algebras

Polarization operators (Chapter 3, §2) are special types of *derivations*. Let us recall the general definitions. Given an associative algebra A, we define the *Lie product*

$$[a, b] := ab - ba,$$

and verify immediately that it satisfies the *Lie axioms*:

$[a, b] = -[b, a]$ (*antisymmetry*), and $[a, [b, c]] + [b, [c, a]] + [c, [a, b]] = 0$ (*Jacobi identity*).

Recall that a general *algebra* is a vector space with a bilinear product.

Definition 1. An algebra with a product $[a, b]$ satisfying the antisymmetry axiom and the Jacobi identity is called a *Lie algebra*. $[a, b]$ is called a *Lie bracket*.

Exercise. The algebra $L, [\ ,\ ]$ with the new product $\{a, b\} := [b, a]$ is also a Lie algebra isomorphic to L.

The first class of Lie algebras to be considered are the algebras $gl(U)$, the Lie algebra associated to the associative algebra $\mathrm{End}(U)$ of linear operators on a vector space U.

Given any algebra A (not necessarily associative), with product denoted ab, we define:

Definition 2. A derivation of A is a linear mapping $D : A \to A$, satisfying $D(ab) = D(a)b + aD(b)$, for every $a, b \in A$.

The main remarks are:

Proposition.

(i) In a Lie algebra L the Jacobi identity expresses the fact that the map

$$\mathrm{ad}(a) := b \mapsto [a, b]$$

is a derivation.[15]

(ii) The derivations of any algebra A form a Lie subalgebra of the space of linear operators.

Proof. The proof is by direct verification. □

In an associative algebra or a Lie algebra, the derivations given by $b \mapsto [a, b]$ are called *inner*.

The reason why Lie algebras and derivations are important is that they express infinitesimal analogues of groups of symmetries, as explained in the following sections. The main geometric example is:

Definition 3. A derivation X of the algebra $C^\infty(M)$ of C^∞ functions on a manifold M is called a *vector field*. We will denote by $\mathcal{L}(M)$ the Lie algebra of all vector fields on M.

Our guiding principle is (cf. 1.4):

$\mathcal{L}(M)$ is the infinitesimal form of the group of all diffeomorphisms of M.

The first formal property of vector fields is that they are *local*. This means that, given a function $f \in C^\infty(M)$ and an open set U, the value of $X(f)$ on U depends only on the value of f on U. In other words:

Lemma. *If $f = 0$ on U, then also $X(f) = 0$ on U.*

Proof. Let $p \in U$ and let V be a small neighborhood of p in U. We can find a C^∞ function u on M which is 1 outside of U and 0 on V.

Hence $f = uf$ and $X(f) = X(u)f + uX(f)$ which is manifestly 0 at p. □

Recall that given an n-dimensional manifold M, a *tangent vector* v at a point $p \in M$ is a linear map $v : C^\infty(M) \to \mathbb{R}$ satisfying $v(fg) = v(f)g(p) + f(p)v(g)$. The tangent vectors in p are an n-dimensional vector space, denoted $T_p(M)$, with basis the operators $\frac{\partial}{\partial x_i}$ if $x_1, \dots, x_n$ are local coordinates.

The union of the tangent spaces forms the *tangent bundle* to M, itself a manifold.

Finally, if $F : M \to N$ is a C^∞ map of manifolds and $p \in M$, we have the *differential* $dF_p : T_p(M) \to T_{F(p)}M$ which is implicitly defined by the formula

[15] ad stands for *adjoint* action.

$$dF_p(v)(f) := v(f \circ F). \tag{1.1.1}$$

For a composition of maps $M \xrightarrow{F} N \xrightarrow{G} P$ we clearly have $d(G \circ F) = dG \circ dF$ (the chain rule).

The previous property can be easily interpreted (cf. [He]) by saying that we can consider a derivation of the algebra of C^∞ functions as a section of the tangent bundle. In local coordinates x_i, we have $X = \sum_{i=1}^n f_i(x_1, \dots, x_n)\frac{\partial}{\partial x_i}$, a *linear differential operator.*

Remark. Composing two linear differential operators gives a quadratic operator. The remarkable fact, which we have proved using the notion of derivation, is that the Lie bracket $[X, Y]$ of two linear differential operators is still linear. The quadratic terms cancel.

As for any type of algebraic structure we have the notion of *Lie subalgebra* A of a Lie algebra L: a subspace $A \subset L$ closed under bracket.

The homomorphism $f : L_1 \to L_2$ of Lie algebras, i.e., a linear map preserving the Lie product.

Given a Lie algebra L and $x \in L$ we have defined the linear operator $\mathrm{ad}(x) : L \to L$ by $\mathrm{ad}(x)(y) := [x, y]$. The operator $\mathrm{ad}(x)$ is called the *adjoint* of x.

As for the notion of ideal, which is the kernel of a homomorphism, the reader will easily verify this:

Definition 4. An ideal I of L is a linear subspace stable under all of the operators $\mathrm{ad}(x)$.

The quotient L/I is naturally a Lie algebra and the usual homomorphism theorems hold. Conversely, the kernel of a homomorphism is an ideal.

1.2 Exponential Map

The main step in passing from infinitesimal to global transformations is done by integrating a system of differential equations. Formally (but often concretely) this consists of taking the exponential of the linear differential operator.

Consider the finite-dimensional vector space F^n where F is either $\mathbb{C}$ or $\mathbb{R}$ (complex or real field), with its standard Hilbert norm. Given a matrix A we define its norm:

$$|A| := \max\left\{\frac{|A(v)|}{|v|},\ v \neq 0\right\}, \text{ or equivalently } \quad |A| = \max\{|A(v)|,\ |v| = 1\}.$$

Of course this extends to infinite-dimensional Hilbert spaces and bounded operators, i.e., linear operators A with $\sup_{|v|=1} |A(v)| := |A| < \infty$. $|A|$ is a *norm* in the sense that:

(1) $|A| \geq 0$, $|A| = 0$ if and only if $A = 0$.
(2) $|\alpha A| = |\alpha||A|$, $\forall \alpha \in F, \forall A$.
(3) $|A + B| \leq |A| + |B|$.

With respect to the multiplicative structure, the following facts can be easily verified:

Proposition 1.

(i) Given two operators A, B *we have* $|AB| \leq |A||B|$.
(ii) The series $e^A := \sum_{k=0}^{\infty} \frac{A^k}{k!}$ *is totally convergent in any bounded set of operators.*
(iii) $\log(1 + A) := \sum_{k=1}^{\infty}(-1)^{k+1}\frac{A^k}{k}$ *is totally convergent for* $|A| \leq 1 - \epsilon$, *for any* $1 \geq \varepsilon > 0$.
(iv) The functions e^A, $\log A$ *are inverse of each other in suitable neighborhoods of* 0 *and* 1.

Remark. For matrices (a_{ij}) we can also take the equivalent norm $\max(|a_{ij}|)$.

The following properties of the *exponential map*, $A \mapsto e^A$, are easily verified:

Proposition 2.

(i) If A, B *are two commuting operators (i.e.,* $AB = BA$*) we have* $e^A e^B = e^{A+B}$ *and also* $\log(AB) = \log(A) + \log(B)$ *if* A, B *are sufficiently close to 1.*
(ii) $e^{-A}e^A = 1$.
(iii) $\frac{de^{tA}}{dt} = Ae^{tA}$.
(iv) $Be^AB^{-1} = e^{BAB^{-1}}$.
If A is an $n \times n$ *matrix we also have:*
(v) If $\alpha_1, \alpha_2, \ldots, \alpha_n$ *are the eigenvalues of A, the eigenvalues of* e^A *are* $e^{\alpha_1}, e^{\alpha_2}, \ldots, e^{\alpha_n}$.
(vi) $\det(e^A) = e^{Tr(A)}$.
(vii) $e^{A^t} = (e^A)^t$.

In particular the mapping $t \mapsto e^{tA}$ is a homomorphism from the additive group of real (or complex) numbers to the multiplicative group of matrices (real or complex).

Definition. The map $t \mapsto e^{tA}$ is called the 1*-parameter subgroup* generated by A.
A is called its infinitesimal generator.

Theorem. *Given a vector* v_0 *the function* $v(t) := e^{tA}v_0$ *is the solution to the differential equation* $v'(t) = Av(t)$ *with initial condition* $v(0) = v_0$.

Proof. The proof follows immediately from (iii). □

It is not restrictive to consider such 1-parameter subgroups. In fact we have:

Exercise. A continuous homomorphism $\varphi : (\mathbb{R}, +) \to Gl(n, F)$ is of the form e^{tA} for a unique matrix A, the *infinitesimal generator* of the group φ. We also have $A = \frac{d\varphi(t)}{dt}\big|_{t=0}$.

Hint. Take the logarithm and prove that we obtain a linear map. □

1.3 Fixed Points, Linear Differential Operators

Let us develop a few basic properties of 1-parameter groups. First:

Proposition 1. *A vector v is fixed under a group $\{e^{tA},\ t \in \mathbb{R}\}$ if and only if $Av = 0$.*

Proof. If $Av = 0$, then $t^k A^k v = 0$ for $k > 0$. Hence $e^{tA}v = v$. Conversely, if $e^{tA}v$ is constant, then its derivative is 0. But its derivative at 0 is in fact Av. □

Remark. Suppose that e^{tA} leaves a subspace U stable. Then, if $v \in U$, we have that $Av \in U$ since this is the derivative of $e^{tA}v$ at 0. Conversely, if A leaves U stable, it is clear from the definition that e^{tA} leaves U stable. Its action on U agrees with the 1-parameter subgroup generated by the restriction of A.

In dual bases e_i, e^j the vector $x(t)$ of coordinate functions $x_i(t) := (e^i, e^{tA}v)$ of the evolving vector satisfies the system of ordinary linear differential equations $\dot{x}(t) = A^t x(t)$.[16] Such a system, with the initial condition $x(0)$, has the global solution $x(t) = e^{tA^t}x(0)$.

Let us now consider a function $f(x)$ on an n-dimensional vector space. We can follow its evolution under a 1-parameter group and set $\varphi(t)(f) := F(t, x) := f(e^{-tA}x)$.

We thus have a 1-parameter group of linear transformations on the space of functions, induced from the action of $\varphi(t)$ on this space.

This is not a finite-dimensional space, and we cannot directly apply the results from the previous section 1.2. If we restrict to homogeneous polynomials, we are in the finite-dimensional case. Thus for this group we have $\varphi(t)(f) = e^{tD_A}f$, where D_A is the operator

$$D_A(f) = \left.\frac{dF(t, x)}{dt}\right|_{t=0}.$$

We have

$$\frac{dF(t, x)}{dt} = \sum_{i=1}^{n} \frac{\partial f(e^{-tA}x)}{\partial x_i}\frac{dx_i(t)}{dt}$$

and, since $\frac{dx(t)}{dt} = -Ax(t)$, at $t = 0$ we have

$$\left.\frac{dx_i(t)}{dt}\right|_{t=0} = -\sum_{j=1}^{n} a_{ij}x_j.$$

Hence

$$\left.\frac{dF(t, x)}{dt}\right|_{t=0} = -\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_j \frac{\partial f}{\partial x_i}.$$

Thus we have found that D_A is the differential operator:

[16] $\dot{f}(t)$ is the time derivative.

(1.3.1)
$$D_A := -\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_j\frac{\partial}{\partial x_i}.$$

We deduce that the formula $\varphi(t)(f) = e^{tD_A}f$ is just the Taylor series:

$$F(t,x) = \sum_{k=0}^{\infty}\frac{(tD_A)^k}{k!}F(0,x) = \sum_{k=0}^{\infty}\frac{t^k}{k!}D_A^k f(x).$$

In order to better understand the operators D_A, let us compute

$$D_A x_i = -\sum_{j=1}^{n} a_{ij}x_j.$$

We see that on the linear space with basis the functions x_i, this is just the linear operator given by the matrix $-A^t$.

Since a derivation is determined by its action on the variables x_i we have:

Proposition 2. *The differential operators D_A are a Lie algebra and*

(1.3.2)
$$[D_A, D_B] = D_{[A,B]}.$$

Proof. This is true on the space spanned by the x_i since $[-A^t, -B^t] = -[A, B]^t$. □

Example. $Sl(2, \mathbb{C})$. We want to study the case of polynomials in two variables x, y.

The Lie algebra of 2×2 matrices decomposes as the direct sum of the 1-dimensional algebra generated by $D = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}$ and the 3-dimensional algebra $sl(2, \mathbb{C})$ with basis

$$H = -x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y},\quad E = -y\frac{\partial}{\partial x},\quad F = -x\frac{\partial}{\partial y}.$$

These operators correspond to the matrices

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

We can see how these operators act on the space P_n of homogeneous polynomials of degree n. This is an $n+1$-dimensional space spanned by the monomials $u_i := (-1)^i y^{n-i}x^i$, on which D acts by multiplication by n. We have

(1.3.3)
$$Hu_i = (n-2i)u_i\,,\quad Fu_i = (n-i)u_{i+1}\,,\quad Eu_i = iu_{i-1}.$$

The reader who has seen these operators before will recognize the standard irreducible representations of the Lie algebra $sl(2, \mathbb{C})$ (cf. Chapter 10, 1.1).

The action of $Sl(2, \mathbb{C})$ on the polynomials $\sum_i a_i(-1)^i y^{n-i}x^i \in P_n$ extends to an action on the functions $p(a_0, a_1, \dots, a_n)$ of the coordinates a_i. We have that

$$E = -\sum(i+1)a_{i+1}\frac{\partial}{\partial a_i},\quad F = -\sum(n-i+1)a_{i-1}\frac{\partial}{\partial a_i},$$

(1.3.4)
$$H = -\sum(n-2i)a_i\frac{\partial}{\partial a_i}$$

are the induced differential operators. We will return to this point later.

1.4 One-Parameter Groups

Of course these ideas have a more general range of validity. For instance, the main facts about the exponential and the logarithm are sufficiently general to hold for any Banach algebra, i.e., an algebra with a norm under which it is complete. Thus one can also apply these results to bounded operators on a Hilbert space.

The linearity of the transformations A is not essential. If we consider a C^∞ differentiable manifold M, we can discuss dynamical systems (in this case also called *flows*) in 2 ways.

Definition 1. A C^∞ flow on a manifold M, or a 1-parameter group of diffeomorphisms, is a C^∞ map:

$$\phi(t, x) : \mathbb{R} \times M \to M$$

which defines an additive group action of $\mathbb{R}$ on M.

This can also be thought of as a C^∞ family of diffeomorphisms:

$$\phi_s : M \to M, \qquad \phi_s(m) := \phi(s, m), \qquad \phi_0 = 1_M, \quad \phi_{s+t} = \phi_s \circ \phi_t.$$

To a flow is associated a vector field X, called the *infinitesimal generator of the flow*.

The vector field X associated to a flow $\phi(t, x)$ can be defined at each point p as follows. Let us start from a fixed p. Denote by $\phi_p(t) := \phi(t, p)$. This is now a map from $\mathbb{R}$ to M which represents the evolution of p with time. In this notation the group property is $\phi_{\phi_p(s)}(t) = \phi_p(s + t)$. Denote the differential of ϕ_p at a point s, by $d\phi_p(s)$. Let

$$X_p := d\phi_p(0)\Big(\frac{d}{dt}\Big), \quad \text{i.e.,} \quad X_p(f) = \frac{d}{dt} f(\phi(t, p))_{t=0}, \ \forall f \in C^\infty(M).$$

Definition 2. The vector field X which, given a function f on M, produces

$$X_p(f) = X(f)(p) := \frac{d}{dt} f(\phi(t, p))_{t=0}, \tag{1.4.1}$$

is the *infinitesimal generator* of the flow $\phi(t)$.

Given a point p, the map $t \mapsto \phi_p(t) = \phi(t, p)$ describes a curve in M which is the *evolution of* p, i.e., the *orbit* under the flow. X_p is the *velocity* of this evolution of p, which by the property $\phi_{\phi_p(s)}(t) = \phi_p(s + t)$ depends only on the position and not on the time:

$$X_{\phi(s,p)} = d\phi_{\phi(s,p)}(0)\Big(\frac{d}{dt}\Big) = d\phi_p(s)\Big(\frac{d}{dt}\Big).$$

A linear operator T on a vector space V induces two linear flows in V and V^*. Identifying V with its tangent space at each point, the vector field X_A associated to the linear flow at a point v is Tv, while the generator of the action on functions is $-T^t$. In dual bases, if A denotes the matrix of T, the vector field is thus (cf. 1.3.2):

(1.4.2) $$-X_A = D_A = -\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_j \frac{\partial}{\partial x_i} \implies [X_A, X_B] = X_{[B,A]}.$$

A vector field $\sum_i f_i(x_1, \dots, x_n)\frac{\partial}{\partial x_i}$ gives rise, at least locally, to a flow which one obtains by solving the linear system of ordinary differential equations

$$\frac{dx_i(t)}{dt} = f_i\big(x_1(t), \dots, x_n(t)\big).$$

Thus one has local solutions $\varphi(t)(x) = F(t, x)$, depending on the parameters x, with initial conditions $F(0, x) = x$. For a given point x^0 such a solution exists, for small values of t in a neighborhood of a given point x^0.

We claim that the property of a local 1-parameter group is a consequence of the uniqueness of solutions of ordinary differential equations. In fact we have that for given t, letting s be a new parameter,

$$\frac{dx_i(t+s)}{ds} = f_i\big(x_1(t+s), \dots, x_n(t+s)\big).$$

Remark. A point is fixed under the flow if and only if the vector field vanishes on it.

For the vector field $X_A = -D_A$, the flow is the 1-parameter group of linear operators e^{tA}. We can in fact, at least locally, linearize every flow by looking at its induced action on functions. By the general principle of (Chapter 1, 2.5.1) the evolution of a function f is given by the formula $\phi(t)f(x) = f(\phi(-t)x) = f(F(-t, x))$. When we fix x, we are in fact restricting to an orbit. We can now develop $\phi(t)f(x)$ in Taylor series. By definition, the derivative with respect to t at a point of the orbit is the same as the derivative with respect to the vector given by X, hence the Taylor series is $\phi(t)f(x) = \sum_{k=0}^{\infty}(-t)^k \frac{X^k}{k!} f(x)$.

In this sense the flow $\phi(t)$ becomes a linear flow on the space of functions, with infinitesimal generator $-X$. We have $-Xf = \frac{d\phi(t)f}{dt}(0)$, and $\phi(t) = e^{-tX}$.[17]

Of course in order to make the equality $\phi(t)f(x) = \sum_{k=0}^{\infty} t^k \frac{(-X)^k}{k!} f(x)$ valid for all x, t, we need some hypotheses, such as the fact that the flow exists globally, and also that the functions under consideration are analytic.

The special case of linear flows has the characteristic that one can find global coordinates on the manifold so that the evolution of these coordinates is given by a linear group of transformations of the finite-dimensional vector space, spanned by the coordinates!

In general of course the evolution of coordinates develops nonlinear terms. We will use a simple criterion for Lie groups, which ensures that a flow exists globally.

Lemma 1. *Suppose there is an $\epsilon > 0$, independent of p, so that the flow exists for all p and all $t < \epsilon$. Then the flow exists globally for all values of t.*

[17] If we want to be consistent with the definition of action induced on functions by an action of a group on a set, we must define the evolution of f by $f(t, x) := f(F(-t, x))$, so the flow on functions is e^{-tX}.

Proof. We have for small values of t the diffeomorphisms $\phi(t)$ which, for t, s sufficiently small, satisfy $\phi(s+t) = \phi(s)\phi(t)$. Given any t we consider a large integer N and set $\phi(t) := \phi(t/N)^N$. The previous property implies that this definition is independent of N and defines the flow globally. □

We have already seen that the group of diffeomorphisms of a manifold (as well as subgroups, as for instance a flow) acts linearly as algebra automorphisms on the algebra of C^∞ functions, by $(gf)(x) = f(g^{-1}x)$. We can go one step further and deduce a linear action on the space of linear operators T on functions. The formula is $g \circ T \circ g^{-1}$ or gTg^{-1}. In particular we may apply this to vector fields. Recall that:

(i) A vector field on M is just a derivation of the algebra of functions $C^\infty(M)$.
(ii) Given a derivation D and an automorphism g of an algebra A, gDg^{-1} is a derivation.

We see that given a diffeomorphism g, the induced action on operators maps vector fields to vector fields.

We can compute at each point $g \circ X \circ g^{-1}$ by computing on functions, and see that:

Lemma 2.

(i) At each point $p = gq$ we have $(gXg^{-1})_p = dg_q(X_q)$.
(ii) Take a vector field X generating a 1-parameter group $\phi_X(t)$ and a diffeomorphism g. Then, the vector field gXg^{-1} generates the 1-parameter group $g\phi_X(t)g^{-1}$.
(iii) The map $X \mapsto gXg^{-1}$ is a homomorphism of the Lie algebra structure.

Proof.

(i)
$$dg_q(X_q)(f) = X_q(f \circ g) = X(f(gx))(q)$$
(1.4.3)
$$= X(g^{-1}f)(q) = g(X(g^{-1}f))(p).$$

(ii) Let us compute $\frac{d}{dt} f(g\phi(t)g^{-1}p)$. The curve $\phi(t)g^{-1}p = \phi(t)q$ has tangent vector X_q at $t = 0$. The curve $g\phi(t)g^{-1}p$ has tangent vector $dg_q(X_q) = (gXg^{-1})_p$ at $t = 0$.
(iii) The last claim is obvious. □

The main result regarding evolution of vector fields is:

Theorem. *Let X and Y be vector fields, $\phi(t)$ the 1-parameter group generated by X. Consider the time-dependent vector field*

$$Y(t) := \phi(t)Y\phi(t)^{-1}, \quad Y(t) = d\phi(t)_{\phi(t)^{-1}p}(Y_{\phi(t)^{-1}p}).$$

Then $Y(t)$ satisfies the differential equation $\frac{dY(t)}{dt} = [Y(t), X]$.

Proof. Thanks to the group property, it is enough to check the equality at $t = 0$. Let f be a function and let us compute the Taylor series of $Y(t)f$. We have

$$
\begin{aligned}
(1.4.4)\qquad Y(t)f &= \phi(t)Y\phi(t)^{-1}f = \phi(t)[Yf + tYXf + O(t^2)] \\
&= Yf + tYXf + O(t^2) - tX[Yf + tYXf + O(t^2)] + O(t^2) \\
&= Yf + t[Y, X]f + O(t^2).
\end{aligned}
$$

In general by $O(t^k)$ we mean a function $h(t)$ which is infinitesimal of order t^k, i.e., $\lim_{t\to 0} t^{k-1}h(t) = 0$. From this the statement follows. □

In different terminology, consider $d\phi(t)^{-1}_{\phi(t)p}Y_{\phi(t)p}$. The derivative of this field at $t = 0$ is $[X, Y]$ and it is called the *Lie derivative*. In other words $[X, Y]_p$ measures the infinitesimal evolution of Y_p on the orbit of p under the time evolution of $\phi(t)$.

Corollary 1. *The* 1*-parameter groups of diffeomorphisms generated by two vector fields X, Y commute if and only if $[X, Y] = 0$.*

Proof. Clearly $Y(t) = \phi_X(t)Y\phi_X(t)^{-1}$ is the generator of $\phi_X(t)\phi_Y(s)\phi_X(t)^{-1}$ (in the parameter s). If the two groups commute, $Y(t)$ is constant so its derivative $[Y, X]$ is 0. Conversely, if $Y(t)$ is constant, then the two groups commute. □

Remark. If X and Y are commuting vector fields and a, b two numbers, $aX + bY$ is the infinitesimal generator of the flow $\phi_X(at)\phi_Y(bt)$.

There is an important special case to notice, when p is a fixed point of $\phi_X(t)$. In this case, although p is fixed, the tangent vectors at p are not necessarily fixed, but *move* according to the linear 1-parameter group $d\phi_X(t)_p$. Thus the previous formulas imply the following.

Corollary 2. *Let X be a vector field vanishing at p, and Y any vector field. Then the value $[Y, X]_p$ depends only on Y_p. The linear map $Y_p \to [Y, X]_p$ is the infinitesimal generator of the* 1*-parameter group $d\phi_X(t)_p$ on the tangent space at p.*

1.5 Derivations and Automorphisms

Let us go back to *derivations* and *automorphisms.*

Consider an algebra A and a linear operator D on A. Assume that there are sufficient convergence properties to ensure the existence of e^{tD} as a convergent power series:[18]

Proposition. *D is a derivation if and only if e^{tD} is a group of automorphisms.*

[18] as for Banach algebras

Proof. This is again a variation of the fact that a vector v is fixed under e^{tD} if and only if $Dv = 0$. In fact to say that e^{tD} are automorphisms means that

$$\forall a, b \in A, \ e^{tD}(ab) - e^{tD}(a)e^{tD}(b) = 0.$$

Writing in power series and taking the coefficient of the linear term we get

$$D(ab) - D(a)b - aD(b) = 0,$$

the condition for a derivation.

Conversely, given a derivation, we see by easy induction that for any positive integer k,

$$D^k(ab) = \sum_{i=0}^{k} \binom{k}{i} D^{k-i}(a)D^i(b).$$

Hence

$$e^{tD}(ab) = \sum_{k=0}^{\infty} \frac{t^k D^k(ab)}{k!} = \sum_{k=0}^{\infty} t^k \sum_{i=0}^{k} \frac{1}{k!} \binom{k}{i} D^{k-i}(a)D^i(b)$$

$$= \sum_{k=0}^{\infty} \sum_{i=0}^{k} \frac{1}{(k-i)!i!} t^{k-i} D^{k-i}(a) t^i D^i(b) = e^{tD}(a)e^{tD}(b). \qquad \square$$

Our heuristic idea is that, for a differentiable manifold M, its group of diffeomorphisms should be *the group of automorphisms of the algebra of C^∞ functions*. Our task is then to translate this idea into rigorous finite-dimensional statements.

2 Lie Groups

2.1 Lie Groups

As we have already mentioned, it is quite interesting to analyze group actions subject to special structural requirements.

The structure of each group G is described by two basic maps: the multiplication $m : G \times G \to G$, $m(a, b) := ab$ and the inverse $i : G \to G$, $i(g) := g^{-1}$. If G has an extra geometric structure we require the compatibility of these maps with it. Thus we say that:

Definition. A group G is a: (1) topological group, (2) Lie group, (3) complex analytic group, (4) algebraic group, (5) affine group,

if G is also a (1) topological space, (2) differentiable manifold, (3) complex analytic manifold, (4) algebraic variety, (5) affine variety,

and if the two maps m, i are compatible with the given structure, i.e., are continuous, differentiable, complex analytic or regular algebraic.

When speaking of Lie groups we have not discussed the precise differentiability hypotheses. A general theorem (solution of Hilbert's 5^{th} problem (cf. [MZ], [Kap])) ensures that a topological group which is locally homeomorphic to Euclidean space can be naturally given a real analytic structure. So Lie groups are in fact real analytic manifolds.

The group $GL(n, \mathbb{C})$ is an affine algebraic group (cf. Chapter 7), acting on $\mathbb{C}^n$ by linear and hence algebraic transformations. A group G is called a *linear group* if it can be embedded in $GL(n, \mathbb{C})$ (of course one should more generally consider as linear groups the subgroups of $GL(n, F)$ for an arbitrary field F).

For an action of G on a set X we can also have the same type of analysis: continuous action of a topological group on a topological space, differentiable actions of Lie groups on manifolds, etc. We shall meet many very interesting examples of these actions in the course of our treatment.

Before we concentrate on Lie groups, let us collect a few simple general facts about topological groups, and actions (cf. [Ho], [B1]). For our discussion, "connected" will always mean "arc-wise connected." Let G be a topological group.

Proposition.

(1) Two actions of G coinciding on a set of generators of G are equal.[19]
(2) An open subgroup of a topological group is also closed.
(3) A connected group is generated by the elements of any given nonempty open set.
(4) A normal discrete subgroup Z of a connected group is in the center.
(5) A topological group G is discrete if and only if 1 is an isolated point.
(6) Let $f : H \to G$ be a continuous homomorphism of topological groups. Assume G connected and there is a neighborhood U of $1 \in H$ for which $f(U)$ is open and $f : U \to f(U)$ is a homeomorphism. Then f is a covering space.

Proof.

(1) is obvious.
(2) Let H be an open subgroup. G is the disjoint union of all its left cosets gH which are then open sets. This means that the complement of H is also open, hence H is closed.
(3) Let U be a nonempty open set of a group G and H be the subgroup that it generates. If $h \in H$ we have that also hU is in H. Hence H is an open subgroup, and by the previous step, it is also closed. Since G is connected and H nonempty, we have $H = G$.
(4) Let $x \in Z$. Consider the continuous map $g \mapsto gzg^{-1}$ from G to Z. Since G is connected and Z discrete, this map is constant and thus equal to $z = 1z1^{-1}$.
(5) is clear.
(6) By hypothesis and (5), $A := f^{-1}1$ is a discrete group. We have $f^{-1}f(U) = \sqcup_{h \in A} Uh$. The covering property is proved for a neighborhood of 1. From 3) it follows that f is surjective. Then for any $g \in G$, $g = f(b)$ we have

[19] For topological groups and continuous actions, we can take topological generators, i.e., elements which generate a dense subgroup.

$$f^{-1}f(bU) = \sqcup_{h\in A} bUh. \qquad \square$$

There is a converse to point 6) which is quite important. We can apply the theory of covering spaces to a connected, locally connected and locally simply connected topological group G. Let $\tilde{G}$ be the universal covering space of G, with covering map $\pi : \tilde{G} \to G$. Let us first define a transformation group $\overline{G}$ on $\tilde{G}$:

$$\overline{G} := \{T \mid T : \tilde{G} \to \tilde{G}, \text{ continuous} \mid \exists g \in G \text{ with } \pi(T(x)) = g\pi(x)\}.$$

The theory of covering spaces implies immediately:

Theorem. *$\overline{G}$ is a group which acts in a simply transitive way on $\tilde{G}$.*

Proof. It is clear that $\overline{G}$ is a group. Given two points $x, y \in \tilde{G}$ there is a unique $g \in G$ with $g\pi(x) = \pi(y)$, and therefore a unique lift T of the multiplication by g, with $T(x) = y$. $\square$

Therefore, given $x \in \tilde{G}$, we can identify $\overline{G}$ with $\tilde{G}$. One easily verifies that:

Corollary. *With this identification $\tilde{G}$ becomes a simply connected topological group. If $\pi(x) = 1$, the mapping π is a homomorphism.*

3 Correspondence between Lie Algebras and Lie Groups

We review here the basics of Lie theory referring to standard books for a more leisurely discussion.

3.1 Basic Structure Theory

In this section the general theory of vector fields and associated 1-parameter groups will be applied to Lie groups. Given a Lie group G, we will associate to it a Lie algebra $\mathfrak{g}$ and an exponential map $\exp : \mathfrak{g} \to G$. The Lie algebra $\mathfrak{g}$ can be defined as the Lie algebra of left-invariant vector fields on G, under the usual bracket of vector fields. The exponential map is obtained by integrating these vector fields, proving that, in this case, the associated 1-parameter groups are global.[20] A homomorphism $\phi : G_1 \to G_2$ of Lie groups induces a homomorphism $d\phi : \mathfrak{g}_1 \to \mathfrak{g}_2$ of the associated Lie algebras. In particular, this applies to linear representations of G which induce linear representations of the Lie algebra. Conversely, a homomorphism of Lie algebras integrates to a homomorphism of Lie groups, provided that G_1 is simply connected.

Before we enter into the details let us make a fundamental definition:

Definition 1. A 1-parameter subgroup of a topological group G is a continuous homomorphism $\phi : \mathbb{R} \to G$, i.e., $\phi(s + t) = \phi(s)\phi(t)$.

[20] This can also be interpreted in terms of geodesics of Riemannian geometry, for a left-invariant metric.

Remark. For a Lie group we will assume that ϕ is C^∞, although one can easily see that this is a consequence of continuity.

Let G be a Lie group. Consider left and right actions, $L_g(h) := gh$, $R_g(h) := hg^{-1}$.

Lemma. *If a transformation $T : G \to G$ commutes with the left action, then $T = R_{T(1)^{-1}}$.*

Proof. We have $T(g) = T(L_g(1)) = L_g(T(1)) = gT(1) = R_{T(1)^{-1}}(g)$. □

Definition 2. We say that a vector field X is left-invariant if $L_g \circ X \circ L_g^{-1} = X$, $\forall g \in G$.

Proposition 1.

(1) A left-invariant vector field $X := X_a$ is uniquely determined by the value $X(1) := a$. Then its value at $g \in G$ is given by the formula

$$X_a(g) = L_g X_a L_g^{-1}(g) = dL_g(X_a(1)) = dL_g(a). \tag{3.1.1}$$

(2) A tangent vector $a \in T_1(G)$ is the velocity vector of a uniquely determined 1-parameter group $t \mapsto \phi_a(t)$. The corresponding left-invariant vector field X_a is the infinitesimal generator of the 1-parameter group of right translations $\phi_{X_a}(t)(g) := g\phi_a(t)$.

Proof.

(1) 3.1.1 is a special case of formula 1.4.3.

(2) By Lemma 1.4 a left-invariant vector field is the infinitesimal generator of a 1-parameter group of diffeomorphisms commuting with left translations. By the previous lemma these diffeomorphisms must be right translations which are defined globally.

Given a tangent vector a at 1, the corresponding vector field X_a and the 1-parameter group of diffeomorphisms $\Phi_a(t) := \phi_{X_a}(t)$, consider the curve $\phi_a(t) := \Phi_a(t)(1)$. We thus have $\Phi_a(t)(g) = g\phi_a(t)$ and also $\phi_a(t+s) = \phi_{X_a}(t+s)(1) = \phi_{X_a}(t)\phi_{X_a}(s)(1) = \phi_a(s)\phi_a(t)$ is a 1-parameter subgroup of G. □

Remark. The 1-parameter group $\phi_a(t)$ with velocity a is also called the *exponential,* denoted $\phi_a(t) = e^{ta} = \exp(ta)$. The map $\exp : T_1(G) \to G$, $\exp(a) := e^a$ is called the *exponential map.*

The differential of exp at 0 is the identity, so $\exp(T_1(G))$ contains a neighborhood of 1. Therefore, if G is connected, we deduce that G is generated by $\exp(T_1(G))$.

Since applying a diffeomorphism to a vector field preserves the Lie bracket we have:

Theorem 1. *The left-invariant vector fields form a Lie subalgebra of the Lie algebra of vector fields on G, called the Lie algebra of G and denoted by L_G. Evaluation at 1 establishes a linear isomorphism of L_G with the tangent space $T_1(G)$.*

In view of the linear isomorphism $L_G \equiv T_1(G)$, it is usual to endow $T_1(G)$ with the Lie algebra structure given implicitly by $X_{[a,b]} := [X_a, X_b]$.

We could have started from right-invariant vector fields Z_a for which the expression is $Z_a(g) = dR_{g^{-1}}(a)$. It is easy to compare the two approaches from the following:

Remark 1. Consider the diffeomorphism $i : x \mapsto x^{-1}$. We have $iL_g i^{-1} = R_g$, so i transforms right into left-invariant vector fields.

We claim that the differential at 1 is $di(a) = -a$. In fact, for the differential of the multiplication map $m : G \times G \to G$ we must have $dm : (a, b) \mapsto a + b$, since composing with the two inclusions, $i_1 : g \mapsto (g, 1)$, $i_2 : g \mapsto (1, g)$ gives the identity on G. Then $m(g, i(g)) = m(i(g), g) = 1$ implies that $di(a) + a = 0$.

In summary, we have two Lie algebra structures on $T_1(G)$ induced by right and left-invariant vector fields. The map $a \mapsto -a$ is an isomorphism of the two structures. In other words, $[Z_a, Z_b] = Z_{[b,a]}$.

The reason to choose left-invariant rather than right-invariant vector fields in the definition of the Lie algebra is the following:

Proposition. *When $G = GL(n, \mathbb{R})$ is the linear group of invertible matrices, its tangent space at 1 can be identified with the space of all matrices $M_n(\mathbb{R})$. If we describe the Lie algebra of $GL(n, \mathbb{R})$ as left-invariant vector fields, we get the usual Lie algebra structure $[A, B] = AB - BA$ on $M_n(\mathbb{R})$.*[21]

Proof. A matrix A (in the tangent space) generates the 1-parameter group of right translations $X \mapsto e^{tA}$, whose infinitesimal generator is the linear vector field associated to the map $R_A : X \mapsto XA$. We have

$$[R_B, R_A](X) = (R_B R_A - R_A R_B)X$$
$$= XAB - XBA = R_{[A,B]}(X), \quad \text{i.e,} \quad [R_B, R_A] = R_{[A,B]}.$$

From 1.4.3, if X_A is the linear vector field associated to R_A we finally have $[X_A, X_B] = X_{[A,B]}$, as required. □

Remark 2. From 1.4 it follows that if $[a, b] = 0$, then $\exp(a)\exp(b) = \exp(a + b)$. In general the Lie bracket is a second-order correction to this equality.

Since right translations commute with left translations, they must map L_G into itself. We have thus a linear action of G on the Lie algebra, called the *adjoint action,* given by

$$\text{(3.1.2)} \qquad \mathrm{Ad}(g)X_a := R_g \circ X_a \circ R_g^{-1}.$$

Explicitly, since $\mathrm{Ad}(g)X_a \in L_G$, we have $\mathrm{Ad}(g)X_a = X_{\mathrm{Ad}(g)(a)}$ for some unique element $\mathrm{Ad}(g)(a)$. From formulas 1.4.3 and 3.1.1 we have

[21] With right-invariant vector fields, we would get $BA - AB$.

$$\mathrm{Ad}(g)(a) = dR_g(X_a(R_g^{-1}(1))) = dR_g(X_a(g))$$

(3.1.3) $$= dR_g(dL_g(a)) = d(R_g \circ L_g)(a).$$

The diffeomorphism $R_g \circ L_g$ is just $x \mapsto gxg^{-1}$, i.e., it is conjugation by g.

For an element $g \in G$ let us denote by $C_g : x \mapsto gxg^{-1}$ the conjugation action. C_g is an automorphism and $C_g \circ C_h = C_{gh}$. Of course $C_g(1) = 1$, hence C_g induces a linear map $dC_g : T_1(G) \to T_1(G)$. The map $g \mapsto dC_g$ is a linear representation of G on $L_G = T_1(G)$. We have just proved:

Proposition 2. *We have*

(3.1.4) $$\mathrm{Ad}(g) = dC_g, \quad \textit{differential of the adjoint map.}$$

(3.1.5) $$ge^a g^{-1} = e^{\mathrm{Ad}(g)(a)}, \ \forall g \in G, \ a \in L_G.$$

Proof. We have proved the first formula. As for the second we have the composition $ge^{ta}g^{-1} : \mathbb{R} \xrightarrow{e^{ta}} G \xrightarrow{C_g} G$ so $\mathrm{Ad}(g)(a) = dC_g(a) = d(ge^{ta}g^{-1})\left(\frac{d}{dt}\right)$ is the infinitesimal generator of the 1-parameter group $ge^{ta}g^{-1}$. □

At this point we can make explicit the Lie algebra structure on $T_1(G)$.

Let $a, b \in T_1(G)$ and consider the two left-invariant vector fields X_a, X_b so that $X_{[a,b]} := [X_a, X_b]$ by definition. By Theorem 1.4 we have that $[X_a, X_b]$ is the derivative at 0 of the variable vector field

$$R^{-1}_{\phi_a(t)} X_b R_{\phi_a(t)}.$$

At the point 1 this takes the value

$$dR_{\phi_a(-t)}(\phi_a(t))X_b(\phi_a(t)) = dR_{\phi_a(-t)}(\phi_a(t))dL_{\phi_a(t)}b = dC_{\phi_a(t)}(b) = \mathrm{Ad}(\phi_a(t))(b).$$

We deduce:

Theorem 2. *Given $a \in T_1(G)$, the linear map $\mathrm{ad}(a) : b \mapsto [a, b]$ is the infinitesimal generator of the 1-parameter group $\mathrm{Ad}(\phi_a(t)) : L_G \to L_G$.*

(3.1.6) $$\mathrm{Ad}(e^a) = e^{\mathrm{ad}(a)}, \quad e^a e^b e^{-a} = e^{\mathrm{Ad}(e^a)(b)} = e^{e^{\mathrm{ad}(a)}(b)}, \ \forall a, b \in L_G.$$

It may be useful to remark how one computes the differential of the multiplication m at a pair of points $g_0, h_0 \in G$. Recall that we have identified $T_{g_0}(G)$ with L via the linear map $dL_{g_0} : L = T_1(G) \to T_{g_0}(G)$. Therefore given two elements $dL_{g_0}(a), dL_{h_0}(b), \ a, b \in L$, computing $dm_{g_0,h_0}(dL_{g_0}(a), dL_{h_0}(b))$ is equivalent to computing $df_{1,1}(a, b)$ where $f : (x, y) \mapsto g_0xh_0y$. Moreover we want to find the $c \in L$ such that $df_{1,1}(a, b) = dL_{g_0h_0}(c)$, and $c = dh_{1,1}(a, b)$ where $h : (x, y) \mapsto (g_0h_0)^{-1}g_0xh_0y = \mathrm{Ad}(h_0^{-1})(x)y$; thus we get

(3.1.7) $$c = \mathrm{Ad}(h_0^{-1})(a) + b.$$

3.2 Logarithmic Coordinates

We have seen that Lie algebras are an infinitesimal version of Lie groups. In order to understand this correspondence, we need to be able to do some local computations in a Lie group G with Lie algebra L in a neighborhood of 1.

Lemma 1. *The differential at* 0 *of the map* $a \mapsto \exp(a)$ *is the identity. Therefore* exp *is a local diffeomorphism between* L *and* G.

We can thus use the coordinates given by $\exp(a)$ which we call *logarithmic coordinates.* Fixing an arbitrary Euclidean norm on L, they will be valid for $|a| < R$ for some R. In these coordinates, if $|a|, N|a| < R$ and N is an integer, we have that $\exp(Na) = \exp(a)^N$. Moreover, (introducing a small parameter t) the multiplication in these coordinates has the form $m(ta, tb) = ta+tb+t^2R(t, a, b)$ where $R(t, a, b)$ is bounded around 0. An essential remark is:

Lemma 2. *Given two* 1*-parameter groups* $\exp(ta)$, $\exp(tb)$ *we have*

$$\lim_{N\to\infty}[\exp\left(\frac{t}{N}a\right)\exp\left(\frac{t}{N}b\right)]^N = \exp(t(a+b)).$$

Proof. If t is very small we can apply the formulas in logarithmic coordinates and see that the coordinates of $[\exp\left(\frac{t}{N}a\right)\exp\left(\frac{t}{N}b\right)]^N$ are $ta+tb+t^2/N\ R(t, a/N, b/N)$. Since $R(t, a/N, b/N)$ is bounded, $\lim_{N\to\infty} ta + tb + t^2/N\ R(t, a/N, b/N) = t(a + b)$ as required. One reduces immediately to the case where t is small. □

We shall use the previous lemma to show that a closed subgroup H of a Lie group G is a Lie subgroup, i.e., it is also a differentiable submanifold. Moreover we will see that the Lie algebra of H is given by restricting to H the left-invariant vector fields in G which are tangent to H at 1. In other words, the Lie algebra of H is the subspace of tangent vectors to G at 1 which are tangent to H as a Lie subalgebra.

Theorem 1. *Let* G *be a Lie group and* H *a closed subgroup. Then* H *is a Lie subgroup of* G. *Its Lie algebra is* $L_H := \{a \in L_G | e^{ta} \in H,\ \forall t \in \mathbb{R}\}$.

Proof. First we need to see that L_H is a Lie subalgebra. We use the previous lemma.

Clearly L_H is stable under multiplication by scalars. If $a, b \in L_H$, we have that $\lim_{N\to\infty}[\exp\left(\frac{t}{N}a\right)\exp\left(\frac{t}{N}b\right)]^N = \exp(t(a+b)) \in H$ since H is closed; hence $a + b \in L_H$. To see that it is closed under Lie bracket we use Theorem 1.4 and the fact that $e^{ta}e^{sb}e^{-ta} \in H$ if $a, b \in L_H$. Thus, we have that $\mathrm{Ad}(e^{ta})(b) \in L_H$, $\forall t$. Finally $[a, b] \in L_H$ since it is the derivative at 0 of $\mathrm{Ad}(e^{ta})(b)$.

Now consider any linear complement A to L_H in L_G and the map $\psi : L_G = A \oplus L_H \to G$, $\psi(a, b) := e^a e^b$. The differential at 0 is the identity map so ψ is a local homeomorphism. We claim that in a small neighborhood of 0, we have $\psi(a, b) \in H$ if and only if $a = 0$; of course $e^a e^b = \psi(a, b) \in H$ if and only if $e^a \in H$. Otherwise, fix an arbitrary Euclidean norm in A. There is an infinite sequence of nonzero elements $a_i \in A$ tending to 0 and with $e^{a_i} \in H$. By compactness we can extract from this sequence a subsequence for which $a_i/|a_i|$ has as its limit a unit vector $a \in A$. We claim that $a \in L_H$, which is a contradiction. In fact, compute $\exp(ta) = \lim_{i\to\infty} \exp(ta_i/|a_i|)$. Let m_i be the integral part of $t/|a_i|$. Clearly since the a_i tend to 0 we have $\exp(ta) = \lim_{i\to\infty} \exp(ta_i/|a_i|) = \lim_{i\to\infty} \exp(m_i a_i) = \lim_{i\to\infty} \exp(a_i)^{m_i} \in H$.

Logarithmic coordinates thus show that the subgroup H is a submanifold in a neighborhood U of 1. By the group property this is true around any other element as well. Given any $h \in H$ we have that $H \cap hU = h(H \cap U)$ is a submanifold in hU. Thus H is a submanifold of G. The fact that L_H is its Lie algebra follows from the definition of L_H. □

In the correspondence between Lie groups and Lie algebras we have:

Theorem 2.

(i) A homomorphism $\rho : G_1 \to G_2$ of Lie groups induces a homomorphism $d\rho$ of the corresponding Lie algebras.
(ii) The kernel of $d\rho$ is the Lie algebra of the kernel of ρ. The map $d\rho$ is injective if and only if the kernel of ρ is discrete.
(iii) If G_2 is connected, $d\rho$ is surjective if and only if ρ is surjective. The map $d\rho$ is an isomorphism if and only if ρ is a covering.

Proof.

(i) Given $a \in L_{G_1}$, we know that $\mathrm{ad}(a)$ is the generator of the 1-parameter group $\mathrm{Ad}(\phi_a(t))$ acting on the tangent space at 1. Under ρ we have that $\rho(\phi_a(t)) = \phi_{d\rho(a)}(t)$ and $\rho \circ C_g = C_{\rho(g)} \circ \rho$. Thus $d\rho \circ \mathrm{Ad}(g) = d\rho \circ dC_g = dC_{\rho(g)} \circ d\rho = \mathrm{Ad}(\rho(g)) \circ d\rho$.

We deduce $d\rho \circ \mathrm{Ad}(\phi_a(t)) = \mathrm{Ad}(\rho(\phi_a(t))) \circ d\rho = \mathrm{Ad}(\phi_{d\rho(a)}(t)) \circ d\rho$. Taking derivatives we have $d\rho \circ \mathrm{ad}(a) = \mathrm{ad}(d\rho(a)) \circ d\rho$ and the formula follows.

(ii) Now, $d\rho(a) = 0$ if and only if $\rho(e^{ta}) = 1\ \forall t$. This means that a is in the Lie algebra of the kernel K, by Theorem 1. To say that this Lie algebra is 0 means that the group K is discrete.

(iii) If $d\rho$ is surjective, by the implicit function theorem, the image of ρ contains a neighborhood U of 1 in G_2. Hence also the subgroup H generated by U. Since G_2 is connected, $H = G_2$. If $d\rho(a)$ is an isomorphism, ρ is also a local isomorphism. The kernel is thus a discrete subgroup and the map ρ is a covering by (6) of Proposition 2.1. □

Exercise. If G is a connected Lie group, then its universal cover is also a Lie group.

The simplest example is $SU(2, \mathbb{C})$, which is a double covering of the 3-dimensional special orthogonal group $SO(3, \mathbb{R})$, providing the first example of *spin* (cf. Chapter 5, §6). Nevertheless, one can establish a bijective correspondence between Lie algebras and Lie groups by restricting to simply connected groups. This is the topic of the next two sections.

3.3 Frobenius Theorem

There are several steps in the correspondence. We begin by recalling the relevant theory. First, we fix some standard notation.

Definition 1. A C^∞ map $i : N \to M$ of differentiable manifolds is an *immersion* if it is injective and, for each $x \in N$, the differential di_x is also injective.

Definition 2. An n-dimensional distribution on a manifold M is a function that to each point $p \in M$ assigns an n-dimensional subspace $P_p \subset T_p(M)$.

The distribution is smooth if, for every $p \in M$, there is a neighborhood U_p and n-linearly independent smooth vector fields X_i on U_p, such that $X_i(p)$ is a basis of $P_p, \forall p \in U_p$.

An *integral manifold* for the distribution is an immersion $j : N \to M$ of an n-dimensional manifold N, so that for every $x \in N$ we have $dj_x(T_xN) = P_{j(x)}$. In other words, N is a submanifold for which the tangent space at each point x is the prescribed space P_x.

It is quite clear that in general there are no integral manifolds. Formally, finding integral manifolds means solving a system of partial differential equations, and as usual there is a compatibility condition. This condition is easy to understand geometrically since the following is an easy exercise.

Exercise. Given an immersion $j : N \to M$ and two vector fields X and Y on M tangent to N, then $[X, Y]$ is also tangent to N.

This remark suggests the following:

Definition 3. A smooth distribution on M is said to be *involutive* if, given any point p in M and a basis $X_1, \dots, X_n$ of vector fields in a neighborhood U of p for the distribution, there exist C^∞ functions $f_{i,j}^k$ on U, with $[X_i, X_j] = \sum_k f_{i,j}^k X_k$.

The prime and only example which will concern us is the distribution induced on a Lie group G by a Lie subalgebra H. In this case a basis of H gives a global basis of vector fields for the distribution. It is clear that the distribution is involutive (in fact the functions $f_{i,j}^k$ are the multiplication constants of the Lie bracket).

If $Y_1, \dots, Y_n$ is a basis of a distribution and $f_{i,j}(x)$ is an invertible $n \times n$ matrix of functions, then $Z_i = \sum_j f_{j,i} Y_j$ is again a basis. The property of being involutive is independent of a choice of the basis since $[Z_h, Z_k] = \sum_{s,t}(f_{s,h} f_{t,k}[Y_s, Y_t] + f_{s,h} Y_s(f_{t,k}) Y_t - f_{t,k} Y_t(f_{s,h}) Y_s)$.

Proposition. *Given an involutive distribution and a point $p \in M$, there exists a neighborhood U of p and vector fields $X_1, \dots, X_n$ in U such that.*

(1) The vector fields $X_1, \dots, X_n$ are a basis of the distribution in U.
(2) $[X_i, X_j] = 0$.

Proof. Start in some coordinates $x_1, \dots, x_m$ with some basis $Y_i = \sum_j a_{j,i}(x)\frac{\partial}{\partial x_j}$.

Since the Y_i are linearly independent, a maximal $n \times n$ minor of the matrix $(a_{j,i})$ is invertible (in a possibly smaller neighborhood). Changing basis using the inverse of this minor, which we may assume to be the first, we reduce to the case in which $a_{i,j}(x) = \delta_i^j$, $\forall i, j \leq n$. Thus the new basis is $X_i = \frac{\partial}{\partial x_i} + \sum_{h=n+1}^m a_{h,i}(x)\frac{\partial}{\partial x_h}$. The Lie bracket $[X_i, X_j]$ is a linear combination of the derivatives $\frac{\partial}{\partial x_h}$, $h > n$. On the other hand the assumption of being involutive means that this commutator is some linear combination $[X_i, X_j] = \sum_{k=1}^n f_{i,j}^k X_k = \sum_{k=1}^n f_{i,j}^k \frac{\partial}{\partial x_k}$ + other terms. Since the coefficients of $\frac{\partial}{\partial x_k}$, $k \leq n$ in $[X_i, X_j]$ equal 0, we deduce that all $f_{i,j}^k = 0$. Hence $[X_i, X_j] = 0$. □

Theorem (Frobenius). *Given an involutive distribution and $p \in M$, there exists a neighborhood of p and a system of local coordinates $(x_1, \dots, x_m)$, such that the distribution has as basis $\frac{\partial}{\partial x_i}$, $i = 1, \dots, n$. So it is formed by the tangent spaces to the* ***level*** *manifolds $x_i = a_i, i = n+1, \dots m$ (a_i constants). These are integral manifolds for the distribution.*

Proof. First we use the previous proposition to choose a basis of commuting vector fields X_i for the distribution. Integrating the vector fields X_i in a neighborhood of p gives rise to n commuting local 1-parameter groups $\phi_i(t_i)$. Choose a system of coordinates y_i around p so that the coordinates of p are 0 and $\frac{\partial}{\partial y_i}$ equals X_i at $p, i = 1, \dots, n$. Consider the map of a local neighborhood of 0 in $\mathbb{R}^n \times \mathbb{R}^{m-n}$ given by $\pi : (x_1, \dots, x_n, x_{n+1}, \dots, x_m) := \phi_1(x_1)\phi_2(x_2)\dots\phi_n(x_n)(0, 0, \dots, 0, x_{n+1}, \dots, x_m)$. It is clear that the differential $d\pi$ at 0 is the identity matrix, so π is locally a diffeomorphism. Further, since the groups $\phi_i(x_i)$ commute, acting on the source space by the translations $x_i \mapsto x_i + s_i$ corresponds under π to the action by $\phi_i(s_i)$:

$$\begin{aligned} \text{(3.3.1)} \qquad \pi &: (x_1, \dots, x_i + s_i, \dots, x_n, x_{n+1}, \dots, x_m) \\ &= \phi_i(s_i)\phi_1(x_1)\phi_2(x_2)\dots\phi_n(x_n)(x_{n+1}, \dots, x_m). \end{aligned}$$

Thus, in the coordinates $x_1, \dots, x_n, x_{n+1}, \dots, x_m$, we have $X_i = \frac{\partial}{\partial x_i}$. The rest follows. □

The special coordinate charts in which the integral manifolds are the submanifolds in which the last $m - n$ coordinates are fixed will be called *adapted* to the distribution.

By the construction of integral manifolds, it is clear that an integral manifold through a point p is (at least locally) uniquely determined and spanned by the evolution of the 1-parameter groups generated by the vector fields defining the distribution. It follows that if we have two integral manifolds A, B and p is a point in their intersection, then an entire neighborhood of p in A is also a neighborhood of p in B. This allows us to construct *maximal integral manifolds* as follows. Let us define a new topology on M. If U is an open set with an adapted chart we redefine the topology on U by *separating* all the level manifolds; in other words, we declare all level manifolds open, leaving in each level manifold its induced topology. The previous remarks show that if we take two adapted charts U_1, U_2, then the new topology on U_1 induces on $U_1 \cap U_2$ the same topology as the new topology induced by U_2.

Call a maximal integral manifold a connected component M_α of M under the new topology. It is clear that such a component is covered by coordinate charts, the coordinate changes are C^∞, and the inclusion map $M_\alpha \to M$ is an immersion. The only unclear point is the existence of a countable set dense in M_α. For a topological group we can use:

Lemma. *Let G be a connected topological group such that there is a neighborhood U of 1 with a countable dense set. Then G has a countable dense set.*

Proof. We may assume $U = U^{-1}$ and X dense in U and countable. Since a topological group is generated by a neighborhood of the identity, we have $G = \cup_{k=1}^{\infty} U^k$.

Then $Y := \cup_{k=1}^{\infty} X^k$ is dense and countable. □

In the case of a Lie group and a Lie subalgebra M the maximal integral manifold through 1 of the distribution satisfies the previous lemma.

In the general case it is still true that maximal integral manifolds satisfy the countability axioms. We leave this as an exercise. *Hint*: If A is a level manifold in a given adapted chart U and U' is a second adapted chart, prove that $A \cap U'$ is contained in a countable number of level manifolds for U'. Then cover M with countably many adapted charts.

Theorem 2. *The maximal integral manifold H through 1 is a subgroup of G. The other maximal integral manifolds are the left cosets of H in G. With the natural topology and local charts H is a Lie group of Lie algebra M. The inclusion map is an immersion.*

Proof. Given $g \in G$, consider the diffeomorphism $x \mapsto gx$. Since the vector fields of the Lie algebra M are left-invariant this diffeomorphism preserves the distribution, hence it permutes the maximal integral manifolds. Thus it is sufficient to prove that H is a subgroup. If we take $g \in H$, we have $g1 = g$, hence H is sent to itself. Thus H is closed under multiplication. Applying now the diffeomorphism $x \mapsto g^{-1}x$, we see that $1 \in g^{-1}H$, hence $g^{-1}H = H$ and $g^{-1} = g^{-1}1 \in H$. □

As we already remarked, H need not be closed. Nevertheless, $\overline{H}$ is clearly a subgroup. Thus we find the following easy criterion for H to be closed.

Criterion. *If there is a neighborhood A of 1, and a closed set $X \supset H$ such that $X \cap A = H \cap A$, then $H = \overline{H}$.*

Proof. Both H and $\overline{H}$ are connected and $\overline{H} \subset X$. Thus $A \cap \overline{H} = A \cap H$. By 2.1, Proposition (2), a connected group is generated by any open neighborhood of the identity, hence the claim. □

3.4 Simply Connected Groups

A given connected Lie group G has a unique universal covering space which is a simply connected Lie group with the same Lie algebra.

The main existence theorem is:

Theorem.

(i) For every Lie algebra $\mathfrak{g}$, there exists a unique simply connected Lie group G such that $\mathfrak{g} = \mathrm{Lie}(G)$.

(ii) Given a morphism $f : \mathfrak{g}_1 \to \mathfrak{g}_2$ of Lie algebras there exists a unique morphism of the associated simply connected Lie groups, $\phi : G_1 \to G_2$ which induces f, i.e., $f = d\phi_1$.

Proof. (i) We will base this result on Ado's theorem (Chapter 5, §7), stating that a finite-dimensional Lie algebra L can be embedded in matrices. If $L \subset gl(n, \mathbb{R})$, then by Theorem 2, 3.3 we can find a Lie group H with Lie algebra L and an immersion to $GL(n, \mathbb{R})$.

Its universal cover is the required simply connected group. Uniqueness follows from the next part applied to the identity map.

(ii) Given a homomorphism $f : \mathfrak{g}_1 \to \mathfrak{g}_2$ of Lie algebras, its graph $\Gamma_f := \{(a, f(a)) \mid a \in \mathfrak{g}_1\}$ is a Lie subalgebra of $\mathfrak{g}_1 \oplus \mathfrak{g}_2$. Let G_1, G_2 be simply connected groups with Lie algebras $\mathfrak{g}_1, \mathfrak{g}_2$. By the previous theorem we can find a Lie subgroup H of $G_1 \times G_2$ with Lie algebra Γ_f. The projection to G_1 induces a Lie homomorphism $\pi : H \to G_1$ which is the identity at the level of Lie algebras. Hence π is a covering. Since G_1 is simply connected we have that π is an isomorphism. The inverse of π composed with the second projection to G_2 induces a Lie homomorphism whose differential at 1 is the given f. □

One should make some remarks regarding the previous theorem. First, the fact that $f : \mathfrak{g}_1 \to \mathfrak{g}_2$ is an injective map does not imply that G_1 is a subgroup of G_2. Second, if G_1 is not simply connected, the map clearly may not exist.

When $M \subset L$ is a Lie subalgebra, we have found, by the method of Frobenius, a Lie group G_1 of Lie algebra M mapped isomorphically to the subgroup of G_2 generated by the elements $\exp(M)$. In general G_1 is not closed in G_2 and its closure can be a much bigger subgroup. The classical example is when we take the 1-parameter group $t \mapsto (e^{tr_1 i}, \dots, e^{tr_n i})$ inside the *torus* of n-tuples of complex numbers of absolute value 1. It is a well-known observation of Kronecker (and not difficult) that when the numbers r_i are linearly independent over the rational numbers the image of this group is dense.[22] The following is therefore of interest:

Proposition. *If $M \subset L$ is an ideal of the Lie algebra L of a group G, then the subgroup $G_1 \subset G$ generated by $\exp(M)$ is normal, and closed, with Lie algebra M.*

[22] This is a very important example, basic in *ergodic theory*.

Proof. It is enough to prove the proposition when G is simply connected (by a simple argument on coverings). Consider the homomorphism $L \to L/M$ which induces a homomorphism from G to the Lie group K of Lie algebra L/M. Its kernel is a closed subgroup with Lie algebra M, hence the connected component of 1 must coincide with G_1. □

Corollary. *In a connected Lie group G we have a 1-1 correspondence between closed connected normal subgroups of G and ideals of its Lie algebra L.*

Proof. One direction is the previous proposition. Let K be a closed connected normal subgroup of G and let M be its Lie algebra. For every element $a \in L$ we have that $\exp(ta)K\exp(-ta) \in K$. From 3.1, it follows that $\mathrm{ad}(a)M \subset M$, hence M is an ideal. □

These theorems lay the foundations of Lie theory. They have taken some time to prove. In fact, Lie's original approach was mostly infinitesimal, and only the development of topology has allowed us to understand the picture in a more global way.

After having these foundational results (which are quite nontrivial), one can set up a parallel development of the theory of groups and algebras and introduce basic structural notions such as solvable, nilpotent, and semisimple for both cases and show how they correspond.

The proofs are not always simple, and sometimes it is simpler to prove a statement for the group or sometimes for the algebra. For Lie algebras the methods are essentially algebraic, while for groups, more geometric ideas may play a role.

Exercise. Show that the set of Lie groups with a prescribed Lie algebra is in correspondence with the discrete subgroups of the center of the unique simply connected group. Analyze some examples.

3.5 Actions on Manifolds

Let us now analyze group actions on manifolds. If G acts on a manifold M by $\rho : G \times M \to M$, and $\phi_a(t)$ is a 1-parameter subgroup of G, we have the 1-parameter group of diffeomorphisms $\phi_a(t)m = \rho(\phi_a(t), m)$, given by the action. We call Y_a its infinitesimal generator. Its value in m is the velocity of the curve $\phi_a(t)m$ at $t = 0$, or $d\rho_{1,m}(a, 0)$.

Theorem 1. *The map $a \mapsto -Y_a$ from the Lie algebra L of G to the Lie algebra of vector fields on M is a Lie algebra homomorphism.*

Proof. Apply Theorem 1.4. Given $a, b \in L$, the 1-parameter group (in the parameter s) $\phi_a(t)\phi_b(s)\phi_a(-t)m$ (depending on t) is generated by a variable vector field $Y_b(t)$ which satisfies the differential equation $\dot{Y}_b(t) = [Y_b(t), Y_a]$, $Y_b(0) = Y_b$. Now $\phi_a(t)\phi_b(s)\phi_a(-t) = \phi_{\mathrm{Ad}(\phi_a(t))(b)}(s)$, so $[Y_b, Y_a]$ is the derivative at $t = 0$ of $Y_{\mathrm{Ad}(\phi_a(t))(b)}$. This, in any point m, is computed by $\frac{d}{dt}d\rho_{1,m}(\mathrm{Ad}(\phi_a(t))(b), 0)_{t=0}$, which equals $d\rho_{1,m}([a, b], 0)$. Thus $[Y_b, Y_a] = Y_{[a,b]}$, hence the claim. □

Conversely, let us give a homomorphism $a \mapsto Z_a$ of the Lie algebra L_G into $\mathcal{L}(M)$, the vector fields on M. We can then consider a copy of L_G as vector fields on $G \times M$ by adding the vector field X_a on G to the vector field Z_a on M. In this way we have a copy of the Lie algebra L_G, and at each point the vectors are linearly independent. We have thus an integrable distribution and can consider a maximal integral manifold.

Exercise. Use the Frobenius theorem to understand, at least locally, how this distribution gives rise to an action of G on M.

Let us also understand an *orbit map*. Given a point $p \in M$ let G_p be its stabilizer.

Theorem 2. *The Lie algebra $L(G_p)$ of G_p is the set of vectors $v \in L$ for which $Y_v(p) = 0$.*

G/G_p has the structure of a differentiable manifold, so that the orbit map $i : G/G_p \to M$, $i(gG_p) := gp$ is an immersion.

Proof. $v \in L(G_p)$ if and only if $\exp(tv) \in G_p$, which means that the 1-parameter group $\exp(tv)$ fixes p. This happens if and only if Y_v vanishes at p.

Let M be a complementary space to $L(G_p)$ in L. The map $j : M \oplus L(G_p) \to G$, $j(a, b) := \exp(a)\exp(b)$ is a local diffeomorphism from some neighborhood $A \times B$ of 0 to a neighborhood U of 1 in G. Followed by the orbit map we have $\exp(a)\exp(b)p = \exp(a)p$ and the map $a \mapsto \exp(a)p$ is an immersion. This gives the structure of a differentiable manifold to the orbit locally around p. At the other points, we translate the chart by elements of G. $\square$

We want to apply the previous analysis to invariant theory.

Corollary. *If G is connected, acting on M, a function f is invariant under G if and only if it satisfies the differential equations $Y_a f = 0$ for all $a \in L_G$.*

Proof. Since G is connected, it is generated by its 1-parameter subgroups $\exp(ta)$, $a \in L_G$. Hence f is fixed under G if and only if it is fixed under these 1-parameter groups. Now f is constant under $\exp(ta)$ if and only if $Y_a f = 0$. $\square$

For instance, for the invariants of binary forms, the differential equations are the ones obtained using the operators 1.3.4.

3.6 Polarizations

We go back to polarizations. Let us consider, as in Chapter 3, 2.3, m-tuples of vector variables $x_1, x_2, \dots, x_m$, each x_i being a column vector $x_{1i}, x_{2i}, \dots, x_{ni}$. In other words we consider the x_{ij} as the coordinates of the space $M_{n,m}$ of $n \times m$ matrices.

Let $A = F[x_{ij}]$ ($F = \mathbb{R}, \mathbb{C}$) be the polynomial ring in the variables x_{ij}, which we also think of as polynomials in the vector variables x_i given by the columns. We want to consider some special 1-parameter subgroups on $M_{n,m}$ (induced by left or right multiplications).

For any $m \times m$ matrix A we consider the 1-parameter group $X \to Xe^{-tA}$.

In particular for the elementary matrix e_{ij}, $i \neq j$ (with 1 in the ij position and 0 elsewhere), we have $e_{ij}^2 = 0$, $e^{-te_{ij}} = 1 - te_{ij}$ and the matrix $Xe^{-te_{ij}}$ is obtained from X adding to its j^{th} column its i^{th} column multiplied by $-t$.

For e_{ii} we have that $Xe^{-te_{ii}}$ is obtained from X multiplying its i^{th} column by e^{-t}. We act dually on the functions in A and the 1-parameter group acts substituting x_j with $x_j + tx_i$, $i \neq j$, resp. x_i with $e^t x_i$. By the previous sections and Chapter 3, Theorem 2.1 we see:

Proposition. *The infinitesimal generator of the transformation of functions induced by* $X \to Xe^{-te_{ij}}$ *is the polarization operator* D_{ij}.

We should summarize these ideas. The group $GL(m, F)$ (resp. $GL(n, F)$) acts on the space of $n \times m$ matrices by the rule $(A, X) \mapsto XA^{-1}$ (resp. $(B, X) \mapsto BX$).

The infinitesimal action is then $X \mapsto -XA := r_A(X)$ (resp. $X \mapsto BX$).

If we denote this operator by r_A, we have $[r_A, r_B] = r_{[A,B]}$. In other words, the map $A \mapsto r_A$ is a *Lie algebra homomorphism* associated to the given action.

The derivation operators induced on polynomials (by the right multiplication action) are the linear span of the polarization operators which correspond to elementary matrices.

We state the next theorem for complex numbers although this is not really necessary.

Recall that an $n \times m$ matrix X can be viewed either as the list of its column vectors $x_1, \ldots, x_m$ or of its row vectors which we will call $x^1, x^2, \ldots, x^n$.

Theorem. *A space of functions* $f(x_1, \ldots, x_m) = f(X)$ *in* m *vector variables, is stable under polarization if and only if it is stable under the action of* $GL(m, \mathbb{C})$ *given, for* $A = (a_{ji}) \in GL(m, \mathbb{C})$, *by*

$$f^A(X) = f(XA), \quad f^A(x_1, \ldots, x_m) := f\Big(\sum_j a_{j1}x_j, \sum_j a_{j2}x_j, \ldots \sum_j a_{jn}x_j\Big)$$

or

$$f^A(x^1, \ldots, x^n) := f(x^1A, x^2A, \ldots, x^nA).$$

Proof. $GL(m, \mathbb{C})$ is connected, so it is generated by the elements e^{tA}. A subspace of a representation of $GL(m, \mathbb{C})$ is stable under e^{tA}, if and only if it is stable under A. In our case the infinitesimal generators are the polarizations D_{ij}. $\square$

3.7 Homogeneous Spaces

We want to discuss a complementary idea, which is important in itself, but for us it is useful in order to understand which groups are simply connected. Let us explain with an example:

Example. The simplest noncommutative example of a simply connected Lie group is $SL(n, \mathbb{C})$ ($SL(n, \mathbb{R})$ is not simply connected).

One way to compute $\pi_1(G)$ and hence check that a Lie group G is simply connected is to work by induction, using the long exact sequence of homotopy for a fibration $H \to G \to G/H$ where H is some closed subgroup. In algebraic topology there are rather general definitions of fibrations which are special maps $f : X \to B$ of spaces with base point $x_0 \in X, b_0 \in B, \ f(x_0) = b_0$ and for which one considers the *fiber* $F := f^{-1}(b_0)$. One has the long exact sequence of homotopy groups (cf. [Sp]):

$$\begin{aligned}\dots\pi_i(F) \to \pi_i(X) \to \pi_i(B) \to \dots \to \pi_1(F)\\ \to \pi_1(X) \to \pi_1(B) \to \pi_0(F) \to \pi_0(X) \to \pi_0(B).\end{aligned}$$

In our case the situation is particularly simple. We will deal with a *locally trivial fibration*, a very special type of fibration for which the long exact sequence of homotopy holds. This means that we can cover B with open sets U_i and we can identify $\pi^{-1}(U_i)$ with $U_i \times F$, so that under this identification, the map π becomes the first projection.

Theorem. *Given a Lie group G and a closed subgroup H, the coset space G/H naturally has the structure of a differentiable manifold (on which G acts in a C^∞ way).*

The orbit map $g \mapsto gH$ is a locally trivial fibration with fiber H.

Once we have established this fact we can define:

Definition. G/H with its C^∞ structure is called a *homogeneous space.*

In other words, a homogeneous space M for a Lie group G is a manifold M with a C^∞ action of G which is transitive.

The structure of a differentable manifold on G/H is quite canonical in the following sense. Whenever we are given an action of G on a manifold M and H stabilizes a point p, we have that the orbit map $\rho : G \to M, \rho(g) = gp$ factors through a C^∞ map $i : G/H \to M$. If H is the full stabilizer of p, then i is an immersion of manifolds. Thus in practice, rather than describing G/H, we describe an action for which H is a stabilizer.

Let us prove the previous statements. The proof is based on the existence of a *tubular neighborhood* of H in G. Let $H \subset G$ be a closed subgroup, let A and B be the Lie algebras of H and G. Consider a linear complement C to A in B. Consider the map $f : C \times H \to G$ given by $f(c, h) := \exp(c)h$. The differential of f at $(0, 1)$ is the identity. Hence, since the map is equivariant with respect to right multiplication by H, there is an open neighborhood U of 0 in C such that df is bijective at all points of $U \times H$. We want to see that:

Tubular neighborhood lemma. *If U is sufficiently small, we have that $f : U \times H \to G$ is a diffeomorphism to an open subset containing H (a union of cosets $\exp(a)H$, $a \in U$).*

Proof. Since df is bijective on $U \times H$, it is sufficient to prove that we can choose U so that f is injective.

Since H is a closed submanifold, there are neighborhoods U of 0 in C and V of 1 in H so that the map $i : (a, b) \mapsto \exp(a)b,\ U \times V \to G$ is a diffeomorphism to a neighborhood W of 1 and $\exp(a)b \in H$ if and only if $a = 0$.

We can consider a smaller neighborhood A of 0 in C so that, if $a_1, a_2 \in A$, we have $\exp(-a_2)\exp(a_1) \in W$. We claim that A satisfies the property that the map $f : A \times H \to G$ is injective. In fact if $\exp(a_1)b_1 = \exp(a_2)b_2$ we have $b := b_2 b_1^{-1} = \exp(-a_2)\exp(a_1) \in W \cap H = V$. Therefore $i(a_1, 1) = \exp(a_1) = \exp(a_2)b = i(a_2, b), b \in V$. Since the map i on $A \times V \to G$ is injective, this implies $a_1 = a_2, b = 1$. Therefore $f(A \times H) = \exp(A)H$ is the required tubular neighborhood, and we identify it (using f) to $A \times H$. □

Thus A naturally parameterizes a set in G/H. We can now give G/H the structure of a differentiable manifold as follows. First we give G/H the quotient topology induced by the map $\pi : G \to G/H$. By the previous construction the map π restricted to the tubular neighborhood $A \times H$ can be identified with the projection $(a, h) \mapsto a$. Its image in G/H is an open set isomorphic to A and will be identified with A. It remains to prove that the topology is Hausdorff and that one can cover G/H with charts gA translating A. Since G acts continuously on G/H, in order to verify the Hausdorff property it suffices to see that if $g \notin H$, we can separate the two cosets gH and H by open neighborhoods. Clearly we can find a neighborhood A' of 0 in A such that $\exp(-A')g\exp(A') \cap H = \emptyset$. Thus we see that $\exp(A')H \cap g\exp(A')H = \emptyset$. The image of $g\exp(A')H$ is a neighborhood of gH which does not intersect the image of $\exp(A')H$. Next one easily verifies that the coordinate changes are C^∞ (and even analytic), giving a manifold structure to G/H. The explicit description given also shows easily that G acts in a C^∞ way and that π is a locally trivial fibration. We leave the details to the reader.

Let us return to showing that $SL(n, \mathbb{C})$ is simply connected. We start from $SL(1, \mathbb{C}) = \{1\}$. In the case of $SL(n, \mathbb{C})$, $n > 1$ we have that $SL(n, \mathbb{C})$ acts transitively on the nonzero vectors in $\mathbb{C}^n$, which are homotopic to the sphere S^{2n-1}, so that $\pi_1(\mathbb{C}^n - \{0\}) = \pi_2(\mathbb{C}^n - \{0\}) = 0$ (cf. [Sp]). By the long exact sequence in homotopy for the fibration $H \to SL(n, \mathbb{C}) \to \mathbb{C}^n - \{0\}$, where H is the stabilizer of e_1, we have that $\pi_1(SL(n, \mathbb{C})) = \pi_1(H)$. H is the group of block matrices $\begin{vmatrix} 1 & a \\ 0 & B \end{vmatrix}$ where a is an arbitrary vector and $B \in SL(n-1, \mathbb{C})$. Thus H is homeomorphic to $SL(n-1, \mathbb{C}) \times \mathbb{C}^{n-1}$ which is homotopic to $SL(n-1, \mathbb{C})$ and we finish by induction. Remark that the same proof shows also that $SL(n, \mathbb{C})$ is connected.

Remark. There is an important, almost immediate generalization of the fibration $H \to G \to G/H$. If $H \subset K \subset G$ are closed subgroups we have a locally trivial fibration of homogeneous manifolds:

$$K/H \to G/H \to G/K. \tag{3.7.1}$$

4 Basic Definitions

4.1 Modules

An important notion is that of *module* or *representation*.

Definition 1. A module over a Lie algebra L consists of a vector space and a homomorphism (of Lie algebras) $\rho : L \to gl(V)$.

As usual one can also give the definition of module by the axioms of an *action*. This is a bilinear map $L \times V \to V$ denoted by $(a, v) \mapsto av$, satisfying the Lie homomorphism condition $[a, b]v = a(bv) - b(av)$.

Remark. Another interpretation of the Jacobi identity is that L is an L-module under the action $[a, b] = \mathrm{ad}(a)(b)$.

Definition 2. The Lie homomorphism $\mathrm{ad} : L \to gl(L)$, $x \mapsto \mathrm{ad}(x)$ is called the *adjoint representation.*

$\mathrm{ad}(x)$ is a derivation, and $x \mapsto \exp(t\,\mathrm{ad}(x))$ is a 1-parameter group of automorphisms of L (1.6).

Definition 3. The group of automorphisms of L generated by the elements $\exp(\mathrm{ad}(x))$ is called the *adjoint group*, $\mathrm{Ad}(L)$, of L, and its action on L the adjoint action.

If g is in the adjoint group, we indicate by $\mathrm{Ad}(g)$ the linear operator it induces on L.

Remark. From §3.1, if L is the Lie algebra of a connected Lie group G, the adjoint action is induced as the differential at 1 of the conjugation action of G on itself.

The kernel of the adjoint representation of L is

(4.1.1) $$Z(L) := \{x \in L \mid [x, L] = 0\}, \quad \text{the center of } L.$$

From 3.1.2 the kernel of the adjoint representation of G is made of the elements which commute with the 1-parameter groups $\exp(ta)$. If G is connected, these groups generate G, hence the kernel of the adjoint representation is the center of G.

As usual one can speak of homomorphisms of modules, or L-linear maps, of submodules and quotient modules, direct sums, and so on.

Exercise. Given two L-modules, M, N, we have an L-module structure on the vector space $\hom(M, N)$ of all linear maps, given by $(af)(m) := a(f(m)) - f(am)$. A linear map f is L-linear if and only if $Lf = 0$.

4.2 Abelian Groups and Algebras

The basic structural definitions for Lie algebras are similar to those given for groups.

If A, B are two subspaces of a Lie algebra, $[A, B]$ denotes the linear span of the elements $[a, b],\ a \in A,\ b \in B$, called the *commutator* of A and B.

Definition. A Lie algebra L is *abelian* if $[L, L] = 0$.

The first remark of this comparison is the:

Proposition. *A connected Lie group is abelian if and only if its Lie algebra is abelian. In this case the map exp : $L \to G$ is a surjective homomorphism with discrete kernel.*

Proof. From Corollary 1 of 1.4, G is abelian if and only if L is abelian. From Remark 2 of §3.1, then exp is a homomorphism. From point (3) of the Proposition in §2.1, exp is surjective. Finally, Lemma 1 and Theorem 2 of §3.2 imply that the kernel is a discrete subgroup. □

As a consequence we have the description of abelian Lie groups. We have the two basic abelian Lie groups: $\mathbb{R}$, the additive group of real numbers, and $S^1 = U(1, \mathbb{C}) = \mathbb{R}/\mathbb{Z}$, the multiplicative group of complex numbers of absolute value 1. This group is compact.

Theorem. *A connected abelian Lie group G is isomorphic to a product $\mathbb{R}^k \times (S^1)^h$.*

Proof. By the previous proposition $G = \mathbb{R}^n/\Lambda$ where Λ is a discrete subgroup of $\mathbb{R}^n$. Thus it suffices to show that there is a basis e_i of $\mathbb{R}^n$ and an $h \le n$ such that Λ is the set of integral linear combinations of the first h vectors e_i. This is easily proved by induction. If $n = 1$ the argument is quite simple. If $\Lambda \neq 0$, since Λ is a discrete subgroup of $\mathbb{R}$, there is a minimum $a \in \Lambda,\ a > 0$. If $x \in \Lambda$ write $x = ma + r$ where $m \in \mathbb{Z},\ |r| < a$. We see that $\pm r \in \Lambda$ which implies $r = 0$ and $\Lambda = \mathbb{Z}a$. Taking a as basis element, $\Lambda = \mathbb{Z}$.

In general take a vector $e_1 \in \Lambda$ such that e_1 generates the subgroup $\Lambda \cap \mathbb{R}e_1$. We claim that the image of Λ in $\mathbb{R}^n/\mathbb{R}e_1$ is still discrete. If we can prove this we find the required basis by induction. Otherwise we can find a sequence of elements $a_i \in \Lambda,\ a_i \notin \mathbb{Z}e_1$ whose images in $\mathbb{R}^n/\mathbb{R}e_1$ tend to 0. Completing e_1 to a basis we write $a_i = \lambda_i e_1 + b_i$ where the b_i are linear combinations of the remaining basis elements. By hypothesis $b_i \neq 0,\ \lim_{i\to\infty} b_i = 0$. We can modify each a_i by subtracting an integral multiple of e_1 so to assume that $|\lambda_i| \le 1$. By compactness we can extract from this sequence another one, converging to some vector λe_1. Since the group is discrete, this means that $a_i = \lambda e_1$ for large i and this contradicts the hypothesis $b_i \neq 0$. □

A compact connected abelian group is isomorphic to $(S^1)^h$ and is called a *compact h-dimensional torus*.

4.3 Nilpotent and Solvable Algebras

Definition 1. Let L be a Lie algebra. The *derived series* is defined inductively:

$$L^{(1)} = L, \ldots, L^{(i+1)} := [L^{(i)}, L^{(i)}].$$

The *lower central series* is defined inductively:

$$L^1 = L, \ldots, L^{i+1} := [L, L^i].$$

A Lie algebra is *solvable* (resp. *nilpotent*) if $L^{(i)} = 0$ (resp. $L^i = 0$) for some i.

Clearly $L^{(i)} \subset L^i$ so nilpotent implies solvable. The opposite is not true as we see by the following:

Basic example. Let B_n resp. U_n be the algebra of upper triangular $n \times n$ matrices over a field F (i.e., $a_{i,j} = 0$ if $i > j$), resp. of strictly upper triangular $n \times n$ matrices (i.e., $a_{i,j} = 0$ if $i \geq j$).

These are two subalgebras for the ordinary product of matrices, hence also Lie subalgebras. Prove that $B_n^{(n)} = 0,\ B_n^i = U_n, \forall i \geq 1, U_n^n = 0$.

B_n is solvable but not nilpotent; U_n is nilpotent.

Remark. To say that $L^i = 0$ means that for all $a_1, a_2, \ldots, a_i \in L$ we have that $[a_1, [a_2, [\ldots, a_i]]]] = 0$. With different notation this means that the operator

$$\text{(4.3.2)} \qquad \operatorname{ad}(a_1)\operatorname{ad}(a_2)\ldots\operatorname{ad}(a_{i-1}) = 0, \quad \forall a_1, a_2, \ldots, a_i \in L.$$

Proposition 1. *A subalgebra of a solvable (resp. nilpotent) Lie algebra is solvable (nilpotent). If L is a Lie algebra and I is an ideal, L is solvable if and only if L/I and I are both solvable. The sum of two solvable ideals is a solvable ideal.*

Proof. The proof is straightforward. □

Warning. It is *not true* that L/I nilpotent and I nilpotent implies L nilpotent (for instance B_n/U_n is abelian).

Remark. The following identity can be proved by a simple induction and will be used in the next proposition:

$$[\operatorname{ad}(b), \operatorname{ad}(a_1)\operatorname{ad}(a_2)\ldots\operatorname{ad}(a_i)] = \sum_{h=1}^{i} \operatorname{ad}(a_1)\operatorname{ad}(a_2)\ldots\operatorname{ad}([b, a_h])\ldots\operatorname{ad}(a_i).$$

Proposition 2. *The sum of two nilpotent ideals A and B is a nilpotent ideal.*

Proof. Note first that using the Jacobi identity, for each i, A^i and B^i are ideals. Assume $A^k = B^h = 0$. We claim that $(A + B)^{h+k-1} = 0$. We need to show that the product of $h - 1 + k - 1$ factors $\operatorname{ad}(a_i)$ with $a_i \in A$ or $a_i \in B$ is 0. At least $k - 1$ of these factors are in A, or $h - 1$ of these factors are in B. Suppose we are in the first case. Each time we have a factor $\operatorname{ad}(a)$ with $a \in B$, which comes to the left of some factor in A, we can apply the previous commutation relation and obtain a sum of terms in which the number of factors in A is not changed, but one factor in B is either dropped or it moves to the right of the monomial. Iterating this procedure we get a sum of monomials each starting with a product of $k-1$ factors in A, which is thus equal to 0. □

In a finite-dimensional Lie algebra L the previous propositions allow us to define the *solvable radical* as the maximal solvable ideal of L, and the *nilpotent radical* as the maximal nilpotent ideal of L.

4.4 Killing Form

In the following discussion all Lie algebras will be assumed to be finite dimensional. If L is finite dimensional one can define a symmetric bilinear form on L:

(4.4.1) $$(x, y) := \mathrm{tr}(\mathrm{ad}(x)\,\mathrm{ad}(y)) \qquad \text{the Killing form.}$$

It has the following *invariance* or *associativity* property $([x, y], z) = (x, [y, z])$.

Proof.

$$\begin{aligned}([x, y], z) &= \mathrm{tr}(\mathrm{ad}([x, y])\,\mathrm{ad}(z)) = \mathrm{tr}(\mathrm{ad}(x)\,\mathrm{ad}(y)\,\mathrm{ad}(z) - \mathrm{ad}(y)\,\mathrm{ad}(x)\,\mathrm{ad}(z)) \\ &= \mathrm{tr}(\mathrm{ad}(x)\,\mathrm{ad}(y)\,\mathrm{ad}(z) - \mathrm{ad}(z)\,\mathrm{ad}(y)\,\mathrm{ad}(x)) \\ &= \mathrm{tr}(\mathrm{ad}(x)\,\mathrm{ad}([y, z])) = (x, [y, z]).\end{aligned}$$ □

Remark. The associativity formula means also that $\mathrm{ad}(x)$ is skew adjoint with respect to the Killing form, or

(4.4.2) $$(\mathrm{ad}(x)y, z) = -(y, \mathrm{ad}(x)z).$$

Definition 1. A Lie algebra L is *simple* if it has no nontrivial ideals and it is not abelian.

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

Over $\mathbb{C}$ there are several equivalent definitions of semisimple Lie algebra. For a Lie algebra L the following are equivalent (cf. [Hu1], [Se2],[J1]) and the next sections.

(1) L is a direct sum of simple Lie algebras.
(2) The Killing form $(x, y) := \mathrm{tr}(\mathrm{ad}(x)\,\mathrm{ad}(y))$ is nondegenerate.
(3) L has no abelian ideals.

For the moment let us see at least:

Lemma 1. *If L has a nonzero solvable radical, then it also has an abelian ideal.*

Proof. Let N be its solvable radical, We claim that for each i, $N^{(i)}$ is an ideal. By induction, $[L, N^{(i+1)}] = [L, [N^{(i)}, N^{(i)}]] \subset [[L, N^{(i)}], N^{(i)}] + [N^{(i)}, [L, N^{(i)}]] \subset [N^{(i)}, N^{(i)}] = N^{(i+1)}$. Thus it is enough to take the last i for which $N^{(i)} \neq 0$, $N^{(i+1)} = 0$. □

Lemma 2. *The Killing form is invariant under any automorphism of the Lie algebra.*

Proof. Let ρ be an automorphism. We have

$$d(\rho(a)) = \rho \circ \mathrm{ad}(a) \circ \rho^{-1} \implies (\rho(a), \rho(b)) = \mathrm{tr}(\mathrm{ad}(\rho(a))\,\mathrm{ad}(\rho(b)))$$

$$= \mathrm{tr}(\rho \circ \mathrm{ad}(a)\,\mathrm{ad}(b) \circ \rho^{-1}) = \mathrm{tr}(\mathrm{ad}(a)\,\mathrm{ad}(b)). \tag{4.4.3}$$

□

5 Basic Examples

5.1 Classical Groups

We list here a few of the interesting groups and Lie algebras, which we will study in the book. One should look specifically at Chapter 6 for a more precise discussion of orthogonal and symplectic groups and Chapter 10 for the general theory.

We have already seen the linear groups:

$$GL(n, \mathbb{C}), \quad GL(n, \mathbb{R}), \quad SL(n, \mathbb{C}), \quad SL(n, \mathbb{R})$$

which are readily seen to have real dimension $2n^2, n^2, 2(n^2 - 1), n^2 - 1$. We also have:

The *unitary group* $U(n, \mathbb{C}) := \{X \in GL(n, \mathbb{C}) \mid \overline{X}^t X = 1\}$.

Notice in particular $U(1, \mathbb{C})$ is the set of complex numbers of absolute value 1, the circle group, denoted also by S^1.

The *special unitary group* $SU(n, \mathbb{C}) := \{X \in SL(n, \mathbb{C}) \mid \overline{X}^t X = 1\}$.

The *complex and real orthogonal groups*

$$O(n, \mathbb{C}) := \{X \in GL(n, \mathbb{C}) \mid X^t X = 1\}, \quad O(n, \mathbb{R}) := \{X \in GL(n, \mathbb{R}) \mid X^t X = 1\}.$$

The *special complex and real orthogonal groups*

$$SO(n, \mathbb{C}) := \{X \in SL(n, \mathbb{C}) \mid X^t X = 1\}, \quad SO(n, \mathbb{R}) := \{X \in SL(n, \mathbb{R}) \mid X^t X = 1\}.$$

There is another rather interesting group called by Weyl the *symplectic group* $Sp(2n, \mathbb{C})$. We can define it starting from the $2n \times 2n$ skew symmetric block matrix $J_n := \begin{vmatrix} 0 & 1_n \\ -1_n & 0 \end{vmatrix}$ (where 1_n denotes the identity matrix of size n) as

$$Sp(2n, \mathbb{C}) := \{X \in GL(2n, \mathbb{C}) \mid X^t J_n X = J_n\}.$$

Since $J_n^2 = -1$ the last condition is better expressed by saying that $X^s X = 1$ where $X^s := -J_n X^t J_n = J_n X^t J_n^{-1}$ is the *symplectic transpose.*

Write X as 4 blocks of size n and have

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix}^s = \begin{vmatrix} D^t & -B^t \\ -C^t & A^t \end{vmatrix}. \tag{5.1.1}$$

Finally the compact symplectic group:

$$Sp(n) := \{X \in U(2n, \mathbb{C}) \mid X^t J_n X = J_n\}.$$

Out of these groups we see that

$$GL(n, \mathbb{C}), \quad SL(n, \mathbb{C}), \quad SO(n, \mathbb{C}), \quad O(n, \mathbb{C}), \quad Sp(2n, \mathbb{C})$$

are *complex algebraic*, that is they are defined by polynomial equations in complex space (see Chapter 7 for a detailed discussion), while

$$U(n, \mathbb{C}), \quad SU(n, \mathbb{C}), \quad SO(n, \mathbb{R}), \quad O(n, \mathbb{R}), \quad Sp(n)$$

are *compact* as topological spaces. In fact they are all closed bounded sets of some complex space of matrices. Their local structure can be deduced from the exponential map but also from another interesting device, the *Cayley transform.*

The Cayley transform is a map from matrices to matrices defined by the formula

$$C(X) := \frac{1 - X}{1 + X}. \tag{5.1.2}$$

Of course this map is not everywhere defined but only on the open set U of matrices with $1+X$ invertible, that is without the eigenvalue -1. We see that $1+C(X) = \frac{2}{1+X}$, so rather generally if 2 is invertible (and not only over the complex numbers) we have $C(U) = U$, and we can iterate the Cayley transform and see that $C(C(X)) = X$.

The Cayley transform maps U to U and 0 to 1. In comparison with the exponential, it has the advantage that it is algebraic but the disadvantage that it works in less generality.

Some immediate properties of the Cayley transform are (conjugation applies to complex matrices):

$$C(-X) = C(X)^{-1},\ C(X^t) = C(X)^t,\ C(\overline{X}) = \overline{C(X)},\ C(AXA^{-1}) = AC(X)A^{-1},$$

and finally $C(X^s) = C(X)^s$.

Therefore we obtain:

$$\begin{aligned} C(X) \in U(n, \mathbb{C}) &\quad \text{if and only if} \quad -X = \overline{X}^t. \\ C(X) \in O(n, \mathbb{C}) &\quad \text{if and only if} \quad -X = X^t. \\ C(X) \in Sp(n, \mathbb{C}) &\quad \text{if and only if} \quad -X = X^s. \end{aligned}$$

There are similar conditions for $O(n, \mathbb{R})$ and $Sp(n)$.

It is now easy to see that the matrices which satisfy any one of these conditions form a Lie algebra. In fact the conditions are that $-x = x^*$ where $x \to x^*$ is a ($\mathbb{R}$-linear map) satisfying $(xy)^* = y^*x^*,\ (x^*)^* = x$, cf. Chapter 5 where such a map is called an involution.

Now we have

$$-x = x^*, -y = y^* \implies -[x, y] = -xy + yx = -x^*y^* + y^*x^*$$
$$= -(yx)^* + (xy)^* = [x, y]^*.$$

Proposition. *In an associative algebra A the space of elements which satisfy $-x = x^*$ for an involution $x \to x^*$ is closed under Lie bracket and so it is a Lie subalgebra.*

Remark. A priori it is not obvious that these are the Lie algebras defined via the exponential map. This is clear if one identifies the Lie algebra as the tangent space to the group at 1, which can be computed using any parameterization around 1.

5.2 Quaternions

Denote by $\mathbb{H}$ the algebra of quaternions.

$$\mathbb{H} = \{a + bi + cj + dk \mid a, b, c, d \in \mathbb{R}\}, \quad i^2 = j^2 = k^2 = -1, \ ij = k.$$

Proposition. *$Sp(n)$ can be defined as the set of quaternionic $n \times n$ matrices X with $\overline{X}^t X = 1$.*

In the previous formula, conjugation is the conjugation of quaternions. In this representation $Sp(n)$ will be denoted by $Sp(n, \mathbb{H})$. This is not the same presentation we have given, but rather it is in a different basis.

Proof. We identify the $2n \times 2n$ complex matrices with $n \times n$ matrices with entries 2×2 matrices and replace J_n with a *diagonal block matrix* J_n' made of n diagonal blocks equal to the 2×2 matrix J_1.

We use Cayley's representation of quaternions $\mathbb{H}$ as elements $q = \alpha + j\beta$ with $\alpha, \beta \in \mathbb{C}$ and $j\alpha = \overline{\alpha} j,\ j^2 = -1$. Setting $k := -ji$ we have $a + bi + cj + dk = (a + bi) + j(c - di)$.

Consider $\mathbb{H}$ as a right vector space over $\mathbb{C}$ with basis $1, j$. An element $\alpha + j\beta = q \in \mathbb{H}$ induces a matrix by left multiplication. We have $q1 = q, qj = -\overline{\beta} + j\overline{\alpha}$, thus the matrix

$$q = \begin{vmatrix} \alpha & -\overline{\beta} \\ \beta & \overline{\alpha} \end{vmatrix}, \quad \det(q) = \alpha\overline{\alpha} + \beta\overline{\beta}. \tag{5.2.1}$$

From the formula of symplectic involution we see that the symplectic transpose of a quaternion is a quaternion $q^s = \overline{\alpha} - j\beta, \quad (a + bi + cj + dk)^s = a - bi - cj - dk$.

From 5.1.1 and 5.2.1 it follows that *a 2×2 matrix q is a quaternion if and only if $q^s = \overline{q}^t$.*

We define symplectic transposition using the matrix J'_n. Take an $n \times n$ matrix $X = (a_{i,j})$ of block 2×2 matrices $a_{i,j}$. We see that $X^s = -J'_n X^t J'_n = (a^s_{j,i})$ while $\overline{X}^t = (\overline{a}^t_{j,i})$. Thus $X^s = \overline{X}^t$ if and only if X is a quaternionic matrix. Thus if $X^s = X^{-1} = \overline{X}^t$ we must have X quaternionic and the proof is complete. □

5.3 Classical Lie Algebras

This short section anticipates ideas which will be introduced systematically in the next chapters. We use freely some tensor algebra which will be developed in Chapter 5.

The list of classical Lie algebras, where n is called the rank, is a reformulation of the last sections. It is, apart from some special cases, a list of simple algebras. The verification is left to the reader but we give an example:

(1) $sl(n+1,\mathbb{C})$ is the Lie algebra of the special linear group $SL(n+1,\mathbb{C})$, $\dim sl(n+1,\mathbb{C}) = (n+1)^2 - 1$. $sl(n+1,\mathbb{C})$ is the set of $(n+1) \times (n+1)$ complex matrices with trace 0; if $n > 0$ it is a simple algebra, said to be *of type* A_n.

Hint as to how to prove simplicity: Let I be an ideal. I is stable under the adjoint action of the diagonal matrices so it has a basis of eigenvectors. Choose one such eigenvector; it is either a diagonal matrix or a matrix unit. Commuting a nonzero diagonal matrix with a suitable off-diagonal matrix unit we can always find a matrix unit $e_{i,j}$, $i \neq j$ in the ideal. Then by choosing appropriate other matrix units we can find all matrices in the ideal.

(2) $so(2n,\mathbb{C})$ is the Lie algebra of the special orthogonal group $SO(2n,\mathbb{C})$, we have $\dim so(2n,\mathbb{C}) = 2n^2 - n$. In order to describe it in matrix form it is convenient to choose a hyperbolic basis $e_1,\dots,e_n, f_1,\dots,f_n$ where the matrix of the form is (cf. Chapter 5, §3.5):

$$I_{2n} := \begin{pmatrix} 0 & 1_n \\ 1_n & 0 \end{pmatrix}, \qquad so(2n,\mathbb{C}) := \{A \in M_{2n}(\mathbb{C}) \mid A^t I_{2n} = -I_{2n} A\}.$$

If $n > 3$, $so(2n,\mathbb{C})$ is a simple algebra, said to be *of type* D_n.

Proposition 1. *For $n = 3$ we have the special isomorphism $so(6,\mathbb{C}) = sl(4,\mathbb{C})$.*

Proof. If V is a 4-dimensional vector space, we have an action of $SL(V)$ on $\bigwedge^2 V$ which is 6-dimensional. The action preserves the symmetric pairing $\bigwedge^2 V \times \bigwedge^2 V \to \bigwedge^4 V = \mathbb{C}$. So we have a map $SL(V) \to SO(\bigwedge^2 V)$. We leave to the reader to verify that it is surjective and at the level of Lie algebras induces the required isomorphism. □

Proposition 2. *For $n = 2$ we have the special isomorphism $so(4,\mathbb{C}) = sl(2,\mathbb{C}) \oplus sl(2,\mathbb{C})$.*

Proof. Let V be a 2-dimensional space, V has the symplectic form $V \times V \to \bigwedge^2 V = \mathbb{C}$ and $V \otimes V$ has an induced symmetric form $(u_1 \otimes u_2, v_1 \otimes v_2) = (u_1 \wedge v_1)(u_2 \wedge v_2)$. Then $SL(V) \times SL(V)$ acting as a tensor product preserves this form, and we have a surjective homomorphism $SL(V) \times SL(V) \to SO(V \otimes V)$ which at the level of Lie algebras is an isomorphism.[23] □

(3) $so(2n+1, \mathbb{C})$ is the Lie algebra of the special orthogonal group $SO(2n+1, \mathbb{C})$. We have $\dim so(2n+1, \mathbb{C}) = 2n^2 + n$. In order to describe it in matrix form it is convenient to choose a hyperbolic basis $e_1, \dots, e_n, f_1, \dots, f_n, u$ where the matrix of the form is $I_{2n+1} := \begin{vmatrix} 0 & 1_n & 0 \\ 1_n & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix}$ and $so(2n+1, \mathbb{C}) := \{A \in M_{2n+1}(\mathbb{C}) | A^t I_{2n+1} = -I_{2n+1} A\}$. If $n > 0$, it is a simple algebra, said to be *of type B_n for $n > 1$*.

Proposition 3. *For $n = 1$ we have the special isomorphism $so(3, \mathbb{C}) = sl(2, \mathbb{C})$.*

Proof. This can be realized by acting with $SL(2, \mathbb{C})$ by conjugation on its Lie algebra, the 3-dimensional space of trace 0, 2×2 matrices. This action preserves the form $\mathrm{tr}(AB)$ and so it induces a map $SL(2, \mathbb{C}) \to SO(3, \mathbb{C})$ that is surjective and, at the level of Lie algebras, induces the required isomorphism. □

(4) $sp(2n, \mathbb{C})$ is the Lie algebra of the symplectic group $Sp(2n, \mathbb{C})$, $\dim sp(2n, \mathbb{C}) = 2n^2 + n$. In order to describe it in matrix form it is convenient to choose a symplectic basis $e_1, \dots, e_n, f_1, \dots, f_n$, where the matrix of the form is $J := \begin{pmatrix} 0 & 1_n \\ -1_n & 0 \end{pmatrix}$ and $sp(2n, \mathbb{C}) := \{A \in M_{2n}(\mathbb{C}) | A^t J = -JA\}$. If $n > 1$ it is a simple algebra, said to be *of type C_n for $n > 1$*.

Proposition 4. *For $n = 1$ we have the special isomorphism $sp(2, \mathbb{C}) = sl(2, \mathbb{C})$.*
For $n = 2$ we have the isomorphism $sp(4, \mathbb{C}) = so(5, \mathbb{C})$, hence $B_2 = C_2$.

Proof. This is seen as follows. As previously done, take a 4-dimensional vector space V with a symplectic form, which we identify with $I = e_1 \wedge f_1 + e_2 \wedge f_2 \in \bigwedge^2 V$. We have an action of $Sp(V)$ on $\bigwedge^2 V$ which is 6-dimensional and preserves the symmetric pairing $\bigwedge^2 V \times \bigwedge^2 V \to \bigwedge^4 V = \mathbb{C}$. So we have a map $Sp(V) \to SO(\bigwedge^2 V)$. The element I is fixed and its norm $I \wedge I \neq 0$, hence $Sp(V)$ fixes the 5-dimensional orthogonal complement $I^\perp$ and we have an induced map $Sp(4, \mathbb{C}) \to SO(5, \mathbb{C})$. We leave to the reader to verify that it is surjective and at the level of Lie algebras induces the required isomorphism. □

No further isomorphisms arise between these algebras. This follows from the theory of root systems (cf. Chapter 10, §2,3). The list of all complex simple Lie algebras is completed by adding the five exceptional types, called G_2, F_4, E_6, E_7, E_8.

The reason to choose these special bases is that in these bases it is easy to describe a Cartan subalgebra and the corresponding theory of roots (cf. Chapter 10).

[23] Since there are spin groups one should check that these maps we found are in fact not isomorphisms of groups, but rather covering maps.

Remark. We have described parameterizations around the identity element 1. If we take any element g in the group, we can find a parameterization around g, remarking that the map $x \to gx$ maps a neighborhood of 1 into one of g, and the group into the group.

6 Basic Structure Theorems for Lie Algebras

6.1 Jordan Decomposition

The theory of Lie algebras is a generalization of the theory of a single linear operator. For such an operator on a finite-dimensional vector space the basic fact is the theorem of the Jordan canonical form. Of this theorem we will use the *Jordan decomposition.* Let us for simplicity work over an algebraically closed field.

Definition. A linear map a on a finite-dimensional vector space V is semisimple if it has a basis of eigenvectors. It is nilpotent if $a^k = 0$ for some $k > 0$.

A linear operator is nilpotent and semisimple only if it is 0.

Theorem. *Given a linear operator a on V, there exist unique operators a_s semisimple, a_n nilpotent, such that $a = a_s + a_n$, $[a_s, a_n] = 0$.*

Moreover a_s can be written as a polynomial without constant term in a. If $V \supset A \supset B$ are linear subspaces and $aA \subset B$ we have $a_s A \subset B$, $a_n A \subset B$.

Proof. Let $\alpha_1, \dots, \alpha_k$ be the distinct eigenvalues of a and $n := \dim V$. One can decompose $V = \oplus_i V_i$ where $V_i := \{v \in V \mid (a - \alpha_i)^n v = 0\}$. By the Chinese Remainder Theorem let $f(x)$ be a polynomial with $f(x) \equiv \alpha_i \mod (x - \alpha_i)^n$, $f(x) \equiv 0 \mod x$. Clearly $f(a) = a_s$ is the semisimple operator α_i on V_i. $a_n := a - a_s$ is nilpotent. The rest follows. □

6.2 Engel's Theorem

There are three basic theorems needed to found the theory. The first, Engel's theorem, is true in all characteristics. The other two, the theorem of Lie and Cartan's criterion, hold only in characteristic 0.

In a way these theorems are converses of the basic examples of Section 4.2.

We start with a simple lemma:

Lemma. *Let a be a semisimple matrix with eigenvalues α_i and eigenvectors a basis e_i.*

$\mathrm{ad}(a)$ *is also semisimple with eigenvalues $\alpha_i - \alpha_j$ and eigenvectors the matrix units in the basis e_i.*

If a is nilpotent, $a^k = 0$ we have $\mathrm{ad}(a)^{2k-1} = 0$.

Proof. The statement for semisimple matrices is clear. For nilpotent we use the identity $ad(a) = a_L - a_R$, $a_L(b) = ab$, $a_R(b) = ba$. Since left and right multiplications a_L, a_R are commuting operators,

$$\mathrm{ad}(a)^{2k-1} = (a_L - a_R)^{2k-1} = \sum_{i=0}^{2k-1} \binom{2k-1}{i} a_L^i a_R^{2k-1-i} = 0. \qquad \square$$

Corollary. *If $a \in gl(V)$ and $a = a_s + a_n$ is its Jordan decomposition, then we have that* $\mathrm{ad}(a) = \mathrm{ad}(a_s) + \mathrm{ad}(a_n)$ *is the Jordan decomposition of* $\mathrm{ad}(a)$.

Engel's Theorem. *Let V be a finite-dimensional vector space, and let $L \subset \mathrm{End}(V)$ be a linear Lie algebra all of whose elements are nilpotent. There is a basis of V in which L is formed by strictly upper triangular matrices. In particular L is nilpotent.*

Proof. The proof can be carried out by induction on both $\dim V$, $\dim L$. The essential point is to prove that there is a nonzero vector $v \in V$ with $Lv = 0$, since if we can prove this, then we repeat the argument with L acting on V/Fv. If L is 1-dimensional, then this is the usual fact that a nilpotent linear operator is strictly upper triangular in a suitable basis.

Let $A \subset L$ be a proper maximal Lie subalgebra of L. We are going to apply induction in the following way. By the previous lemma, $\mathrm{ad}(A)$ consists of nilpotent operators on $gl(V)$, hence the elements of $\mathrm{ad}(A)$ are also nilpotent acting on L and L/A.

By induction there is an element $u \in L$, $u \notin A$ with $\mathrm{ad}(A)u = 0$ modulo A or $[A, u] \subset A$. This implies that $A + Fu$ is a larger subalgebra; by maximality of A we must have that $L = A + Fu$ and also that A is an ideal of L. Now let $W := \{v \in V \mid Au = 0\}$. By induction, $W \neq 0$. We have, if $w \in W, a \in A$, that $a(uw) = [a, u]w + uaw = 0$ since $[a, u] \in A$. So W is stable under u, and since u is nilpotent, there is a nonzero vector $v \in W$ with $Aw = uw = 0$. Hence $Lw = 0$. $\square$

6.3 Lie's Theorem

Lie's Theorem. *Let V be a finite-dimensional vector space over the complex numbers, and let $L \subset \mathrm{End}(V)$ be a linear solvable Lie algebra. There is a basis of V in which L is formed by upper triangular matrices.*

Proof. The proof can be carried out by induction on $\dim L$. The essential point is again to prove that there is a nonzero vector $v \in V$ which is an eigenvector for L. If we can prove this, then we repeat the argument with L acting on V/Fv. If L is 1-dimensional, then this is the usual fact that a linear operator has a nonzero eigenvector.

We start as in Engel's Theorem. Since L is solvable any proper maximal subspace $A \supset [L, L]$ is an ideal of codimension 1. We have again $L = A + \mathbb{C}u$, $[u, A] \subset A$. Let, by induction, $v \in V$ be a nonzero eigenvector for A. Denote by $av := \lambda(a)v$ the eigenvalue (a linear function on A).

Consider the space $W := \{v \in V \mid av = \lambda(a)v, \forall a \in A\}$. If we can prove that W is stabilized by u, then we can finish by choosing a nonzero eigenvector for u in W.

Then let $v \in W$; for some m we have the linearly independent vectors $v, uv, u^2v, \dots, u^m v$ and $u^{m+1}v$ dependent on the preceding ones. We need to prove that $uv \in W$. In any case we have the following identity for $a \in A$: $auv = [a,u]v + uav = \lambda(a)uv + \lambda([u,v])v$. We have thus to prove that $\lambda([u,v]) = 0$. We repeat $au^i v = \lambda(a)u^i v + \cdots + u[a,u]u^{i-2}v + [a,u]u^{i-1}v$, inductively $[a,u]u^{i-2}v = \lambda([a,u])u^{i-2}v + \sum_{k<i-2} c_k u^k v$. Thus a acts as an upper triangular matrix on the span $M := \langle v, uv, u^2v, \dots, u^m v\rangle$ of the vectors $u^i v$ with $\lambda(a)$ on the diagonal. On the other hand, since M is stable under u and a we have that $[u,a]$ is a commutator of two operators on M. Thus the trace of the operator $[u,a]$ restricted to M is 0. On the other hand, by the explicit triangular form of $[u,a]$ we obtain for this trace $(m+1)\lambda([u,v])$. Since we are in characteristic 0, we have $m+1 \neq 0$, hence $\lambda([u,v]) = 0$. □

Corollary. *If $L \subset gl(V)$ is a solvable Lie algebra, then $[L,L]$ is made of nilpotent elements and is thus nilpotent.*

For a counterexample to this theorem in positive characteristic, see [Hu], p. 20.

6.4 Cartan's Criterion

Criterion. *Let V be a finite-dimensional vector space over $\mathbb{C}$, and let $L \subset \mathrm{End}(V)$ be a linear Lie algebra. If $\mathrm{tr}(ab) = 0$ for all $a \in L,\ b \in [L,L]$, then L is solvable.*

Proof. There are several proofs in the literature ([Jac], [Hu]). We follow the *slick proof.* The goal is to prove that $[L,L]$ is nilpotent, or using Engel's theorem, that all elements of $[L,L]$ are nilpotent. First we show that we can make the statement more abstract. Let $M := \{x \in gl(V) \mid [x,L] \subset [L,L]\}$. Of course $M \supset L$ and we want to prove that we still have $\mathrm{tr}(ab) = 0$ when $a \in M, b \in [L,L]$. In fact if $b = [x,y], x, y \in L$, then we have by the associativity of the trace form $\mathrm{tr}([x,y]a) = \mathrm{tr}(x[y,a]) = 0$ since $x \in L, [y,a] \in [L,L]$. Thus the theorem will follow from the next general lemma, applied to $A = [L,L], B = L$. □

Lemma. *Let V be a vector space of finite dimension n over the complex numbers, and let $A \subset B \subset \mathrm{End}(V)$ be linear spaces. Let $M := \{x \in gl(V) \mid [x,B] \subset A\}$. If an element $a \in M$ satisfies $\mathrm{tr}(ab) = 0$ for all $b \in M$, then a is nilpotent.*

Proof. The first remark to make is that if $a \in M$, also $a_s, a_n \in M$. This follows from Lemma 6.2 and the properties of the Jordan decomposition. We need to show that $a_s = 0$. Let $\alpha_1, \dots, \alpha_n$ be the eigenvalues of a_s and $e_1, \dots, e_n$ a corresponding basis of eigenvectors. For any semisimple element b, which in the same basis is diagonal with eigenvalues $\beta_1, \dots, \beta_n$, if $b \in M$ we have $\sum_i \alpha_i \beta_i = 0$. A sufficient condition for $b \in M$ is that $\mathrm{ad}(b)$ is a polynomial in $\mathrm{ad}(a_s)$ without constant terms. In turn this is true if one can find a polynomial $f(x)$ without constant term with $f(\alpha_i - \alpha_j) = \beta_i - \beta_j$. By Lagrange interpolation, the only condition we need is that if $\alpha_i - \alpha_j = \alpha_h - \alpha_k$ we must have $\beta_i - \beta_j = \beta_h - \beta_k$. For this it is sufficient to choose the β_i as the value of any $\mathbb{Q}$-linear function on the $\mathbb{Q}$-vector space spanned by the α_i in $\mathbb{C}$. Take such a linear form g. We have $\sum_i \alpha_i g(\alpha_i) = 0$ which implies $\sum_i g(\alpha_i)^2 = 0$. If the numbers $g(\alpha_i)$ are rationals, then this implies that $g(\alpha_i) = 0, \forall i$. Since g can be any $\mathbb{Q}$ linear form on the given $\mathbb{Q}$ vector space, this is possible only if all the $\alpha_i = 0$. □

6.5 Semisimple Lie Algebras

Theorem 1. *Let L be a finite-dimensional Lie algebra over $\mathbb{C}$. Then L is semisimple if and only if the Killing form is nondegenerate.*

Proof. Assume first that the Killing form is nondegenerate. We have seen in §4.4 (Lemma 1) that, if L is not semisimple, then it has an abelian ideal A. Let us show that A must be in the kernel of the Killing form. If $a \in L, b \in A$ we have that $\mathrm{ad}(a)$ is a linear map that preserves A, while $\mathrm{ad}(b)$ is a linear map which maps L into A and is 0 on A.

Hence $\mathrm{ad}(a)\,\mathrm{ad}(b)$ maps L into A and it is 0 on A, so its trace is 0, and A is in the kernel of the Killing form.

Conversely, let A be the kernel of the Killing form. A is an ideal. In fact by associativity, if $a \in A, b, c \in L$, we have $(b, [c, a]) = ([b, c], a) = 0$. Next consider the elements $\mathrm{ad}(A)$. They form a Lie subalgebra with $\mathrm{tr}(ab) = 0$ for all $a, b \in \mathrm{ad}(A)$. By Cartan's criterion $\mathrm{ad}(A)$ is solvable. Finally the kernel of the adjoint representation is the center of L so if $\mathrm{ad}(A)$ is solvable, also A is solvable, and we have found a nonzero solvable ideal. □

Theorem 2. *A finite-dimensional semisimple Lie algebra is a direct sum of simple Lie algebras L_i which are mutually orthogonal under the Killing form.*

Proof. If L is simple there is nothing to prove, otherwise let A be a minimal ideal in L. Let $A^\perp$ be its orthogonal complement with respect to the Killing form. We have always by the associativity that also $A^\perp$ is an ideal. We claim that $A \cap A^\perp = 0$ so that $L = A \oplus A^\perp$. In fact by minimality the only other possibility is $A \subset A^\perp$. But this implies that A is solvable by Cartan's criterion, which is impossible. Since $L = A \oplus A^\perp$, by minimality A is a simple Lie algebra, $A^\perp$ is semisimple, and we can proceed by induction. □

6.6 Real Versus Complex Lie Algebras

Although we have privileged complex Lie algebras, real Lie algebras are also interesting. We want to make some elementary remarks.

First, given a real Lie algebra L, we can complexify it, getting $L_{\mathbb{C}} := L \otimes_{\mathbb{R}} \mathbb{C}$. The algebra $L_{\mathbb{C}}$ continues to carry some of the information of L, although different Lie algebras may give rise to the same complexification. We say that L *is a real form* of $L_{\mathbb{C}}$. For instance, $sl(2, \mathbb{R})$ and $su(2, \mathbb{C})$ are different and both complexify to $sl(2, \mathbb{C})$. Some properties are easily verified. For instance, $M \subset L$ is a subalgebra or ideal if and only if the same is true for $M_{\mathbb{C}} \subset L_{\mathbb{C}}$. The reader can verify also the compatibility of the derived and lower central series:

$$(6.6.1) \qquad (L^{(i)})_{\mathbb{C}} = (L_{\mathbb{C}})^{(i)}, \qquad (L^i)_{\mathbb{C}} = (L_{\mathbb{C}})^i.$$

Therefore the notions of solvability, nilpotency, and semisimplicity are compatible with the complexification. Finally, the Killing form of $L_{\mathbb{C}}$ is just the complexification of the Killing form of L, thus it is nondegenerate if and only if L is semisimple.

In this case there is an interesting invariant, since the Killing form in the real case is a real symmetric form one can consider its signature (cf. Chapter 5, §3.3) which is thus an invariant of the real form. Often it suffices to detect the real form. In particular we have:

Exercise. Let L be the Lie algebra of a semisimple group K. Then the Killing form of L is negative definite if and only if K is compact.

In fact to prove in full this exercise is not at all easy and the reader should see Chapter 10, §5. It is not too hard when K is the adjoint group (see Chapter 10).

When one studies real Lie algebras, it is interesting to study also real representations, then one can use the methods of Chapter 6, 3.2.

7 Comparison between Lie Algebras and Lie Groups

7.1 Basic Comparisons

In this section we need to compare the concepts of nilpotency, solvability, and semisimplicity introduced for Lie algebras with their analogues for Lie groups.

For Lie groups and algebras *abelian* is the same as *commutative*.

We have already seen in §4.2 that a connected Lie group G is abelian if and only if its Lie algebra is abelian. We want to prove now that the derived group of a connected Lie group has, as Lie algebra, the derived algebra.

For groups there is a parallel theory of central and derived series.[24] In a group G the *commutator of two elements* is the element $\{x, y\} := xyx^{-1}y^{-1}$. Given two subgroups H, K of a group G, one defines $\{H, K\}$ to be the subgroup generated by the commutators $\{x, y\},\ x \in H,\ y \in K$.

The *derived group* of a group G is the subgroup $\{G, G\}$ generated by the commutators. $\{G, G\}$ is clearly the minimal normal subgroup K of G such that G/K is abelian.

The *derived series* is defined inductively:

$$G^{(1)} = G, \ldots, G^{(i+1)} := \{G^{(i)}, G^{(i)}\}.$$

The *lower central series* is defined inductively:

$$G^1 = G, \ldots, G^{i+1} := \{G, G^i\}.$$

[24] In fact all these concepts were first developed for finite groups.

In a topological group we define the *derived group* to be the closed subgroup generated by the commutators, and similar definitions for derived and lower central series.[25]

One has thus the notions of solvable and nilpotent group. Let G be a connected Lie group with Lie algebra L.[26]

Proposition 1. *The derived group of G has Lie algebra* $[L, L]$.

Proof. The derived group is the minimal closed normal subgroup H of G such that G/H is abelian. Since a Lie group is abelian if and only if its Lie algebra is abelian, the proposition follows from Proposition 3.4 since subgroups corresponding to Lie ideals are closed. □

Proposition 2. *The Lie algebra of* $G^{(i+1)}$ *is* $L^{(i+1)}$. *The Lie algebra of* G^{i+1} *is* L^{i+1}.

Proof. The first statement follows from the previous proposition. For the second, let H^i be the connected Lie subgroup of Lie algebra L^i. Assume by induction $H^i = G^i$. Observe that G^{i+1} is the minimal normal subgroup K of G contained in G^i with the property that the conjugation action of G on G^i/K is trivial. Since G is connected, if G acts on a manifold, it acts trivially if and only if its Lie algebra acts by trivial vector fields. If K is a normal subgroup of G contained in G^i with Lie algebra M, G acts trivially by conjugation on G^i/K if and only if the Lie algebra of G acts by trivial vector fields. In particular the restriction of the adjoint action of G on L^i/M is trivial and so $K \supset H^{i+1}$. Conversely it is clear that $H^{i+1} \supset G^{i+1}$. □

Thus a connected Lie group G has a maximal closed connected normal solvable subgroup, the *solvable radical*, whose Lie algebra is the solvable radical of the Lie algebra of G. The nilpotent radical is defined similarly.

Proposition 3. *For a Lie group G the following two conditions are equivalent:*

(1) The Lie algebra of G is semisimple.
(2) G has no connected solvable normal subgroups.

Definition. A group G satisfying the previous two conditions is called a *semisimple* group.

Remark. G may have a nontrivial discrete center. Semisimple Lie groups are among the most interesting Lie groups. They have been completely classified. Of this classification, we will see the complex case (the real case is more intricate and beyond the purpose of this book). This classification is strongly related to algebraic and compact groups as we will illustrate in the next chapters.

---

25 In fact for a Lie group the two notions coincide; cf. [OV].

26 The connectedness hypothesis is obviously necessary.