Chapter 1

Lie Groups and Algebraic Groups

Hermann Weyl, in his famous book (Weyl [1946]), gave the name *classical groups* to certain families of matrix groups. In this chapter we introduce these groups and develop the basic ideas of Lie groups, Lie algebras, and linear algebraic groups. We show how to put a Lie group structure on a closed subgroup of the general linear group and determine the Lie algebras of the classical groups. We develop the theory of complex linear algebraic groups far enough to obtain the basic results on their Lie algebras, rational representations, and Jordan–Chevalley decompositions (we defer the deeper results about algebraic groups to Chapter 11). We show that linear algebraic groups are Lie groups, introduce the notion of a *real form* of an algebraic group (considered as a Lie group), and show how the classical groups introduced at the beginning of the chapter appear as real forms of linear algebraic groups.

1.1 The Classical Groups

1.1.1 General and Special Linear Groups

Let $\mathbb{F}$ denote either the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$, and let V be a finite-dimensional vector space over $\mathbb{F}$. The set of all invertible linear transformations from V to V will be denoted as $\mathrm{GL}(V)$. This set has a group structure under composition of transformations, with identity element the identity transformation $\mathrm{Id}(x) = x$ for all $x \in V$. The group $\mathrm{GL}(V)$ is the first of the classical groups. To study it in more detail, we recall some standard terminology related to linear transformations and their matrices.

Let V and W be finite-dimensional vector spaces over $\mathbb{F}$. Let $\{v_1, \ldots, v_n\}$ and $\{w_1, \ldots, w_m\}$ be bases for V and W, respectively. If $T : V \longrightarrow W$ is a linear map

then

$$Tv_j = \sum_{i=1}^{m} a_{ij} w_i \qquad \text{for } j = 1, \ldots, n$$

with $a_{ij} \in \mathbb{F}$. The numbers a_{ij} are called the *matrix coefficients* or *entries* of T with respect to the two bases, and the $m \times n$ array

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

is the *matrix* of T with respect to the two bases. When the elements of V and W are identified with column vectors in $\mathbb{F}^n$ and $\mathbb{F}^m$ using the given bases, then action of T becomes multiplication by the matrix A.

Let $S : W \longrightarrow U$ be another linear transformation, with U an l-dimensional vector space with basis $\{u_1, \ldots, u_l\}$, and let B be the matrix of S with respect to the bases $\{w_1, \ldots, w_m\}$ and $\{u_1, \ldots, u_l\}$. Then the matrix of $S \circ T$ with respect to the bases $\{v_1, \ldots, v_n\}$ and $\{u_1, \ldots, u_l\}$ is given by BA—the product being the usual product of matrices.

We denote the space of all $n \times n$ matrices over $\mathbb{F}$ by $M_n(\mathbb{F})$, and we denote the $n \times n$ identity matrix by I (or I_n if the size of the matrix needs to be indicated); it has entries $\delta_{ij} = 1$ if $i = j$ and 0 otherwise. Let V be an n-dimensional vector space over $\mathbb{F}$ with basis $\{v_1, \ldots, v_n\}$. If $T : V \longrightarrow V$ is a linear map we write $\mu(T)$ for the matrix of T with respect to this basis. If $T, S \in \mathrm{GL}(V)$ then the preceding observations imply that $\mu(S \circ T) = \mu(S)\mu(T)$. Furthermore, if $T \in \mathrm{GL}(V)$ then $\mu(T \circ T^{-1}) = \mu(T^{-1} \circ T) = \mu(\mathrm{Id}) = I$. The matrix $A \in M_n(\mathbb{F})$ is said to be *invertible* if there is a matrix $B \in M_n(\mathbb{F})$ such that $AB = BA = I$. We note that a linear map $T : V \longrightarrow V$ is in $\mathrm{GL}(V)$ if and only if its matrix $\mu(T)$ is invertible. We also recall that a matrix $A \in M_n(\mathbb{F})$ is invertible if and only if its determinant is nonzero.

We will use the notation $\mathrm{GL}(n, \mathbb{F})$ for the set of $n \times n$ invertible matrices with coefficients in $\mathbb{F}$. Under matrix multiplication $\mathrm{GL}(n, \mathbb{F})$ is a group with the identity matrix as identity element. We note that if V is an n-dimensional vector space over $\mathbb{F}$ with basis $\{v_1, \ldots, v_n\}$, then the map $\mu : \mathrm{GL}(V) \longrightarrow \mathrm{GL}(n, \mathbb{F})$ corresponding to this basis is a group isomorphism. The group $\mathrm{GL}(n, \mathbb{F})$ is called the *general linear group of rank* n.

If $\{w_1, \ldots, w_n\}$ is another basis of V, then there is a matrix $g \in \mathrm{GL}(n, \mathbb{F})$ such that

$$w_j = \sum_{i=1}^{n} g_{ij} v_i \quad \text{and} \quad v_j = \sum_{i=1}^{n} h_{ij} w_i \quad \text{for } j = 1, \ldots, n,$$

with $[h_{ij}]$ the inverse matrix to $[g_{ij}]$. Suppose that T is a linear transformation from V to V, that $A = [a_{ij}]$ is the matrix of T with respect to a basis $\{v_1, \ldots, v_n\}$, and

that $B = [b_{ij}]$ is the matrix of T with respect to another basis $\{w_1, \ldots, w_n\}$. Then

$$\begin{aligned} Tw_j &= T\Big(\sum_i g_{ij} v_i\Big) = \sum_i g_{ij} T v_i \\ &= \sum_i g_{ij} \Big(\sum_k a_{ki} v_k\Big) = \sum_l \Big(\sum_k \sum_i h_{lk} a_{ki} g_{ij}\Big) w_l \end{aligned}$$

for $j = 1, \ldots, n$. Thus $B = g^{-1}Ag$ is similar to the matrix A.

Special Linear Group

The special linear group $\mathrm{SL}(n, \mathbb{F})$ is the set of all elements, A, of $M_n(\mathbb{F})$ such that $\det(A) = 1$. Since $\det(AB) = \det(A)\det(B)$ and $\det(I) = 1$, we see that the special linear group is a subgroup of $\mathrm{GL}(n, \mathbb{F})$.

We note that if V is an n-dimensional vector space with basis $\{v_1, \ldots, v_n\}$ and if $\mu : \mathrm{GL}(V) \longrightarrow \mathrm{GL}(n, \mathbb{F})$ is the map previously defined, then the group

$$\mu^{-1}(\mathrm{SL}(n, \mathbb{F})) = \{T \in \mathrm{GL}(V) \,:\, \det(\mu(T)) = 1\}$$

is independent of the choice of basis, by the change of basis formula. We denote this group by $\mathrm{SL}(V)$.

1.1.2 Isometry Groups of Bilinear Forms

Let V be an n-dimensional vector space over $\mathbb{F}$. A bilinear map $B : V \times V \longrightarrow \mathbb{F}$ is called a *bilinear form*. We denote by $\mathrm{O}(V, B)$ (or $\mathrm{O}(B)$ when V is understood) the set of all $g \in \mathrm{GL}(V)$ such that $B(gv, gw) = B(v, w)$ for all $v, w \in V$. We note that $\mathrm{O}(V, B)$ is a subgroup of $\mathrm{GL}(V)$; it is called the *isometry group* of the form B.

Let $\{v_1, \ldots, v_n\}$ be a basis of V and let $\Gamma \in M_n(\mathbb{F})$ be the matrix with $\Gamma_{ij} = B(v_i, v_j)$. If $g \in \mathrm{GL}(V)$ has matrix $A = [a_{ij}]$ relative to this basis, then

$$B(gv_i,\, gv_j) = \sum_{k,l} a_{ki} a_{lj} B(v_k,\, v_l) = \sum_{k,l} a_{ki} \Gamma_{kl} a_{lj}.$$

Thus if A^t denotes the transposed matrix $[c_{ij}]$ with $c_{ij} = a_{ji}$, then the condition that $g \in \mathrm{O}(B)$ is that

$$\Gamma = A^t \Gamma A. \tag{1.1}$$

Recall that a bilinear form B is *nondegenerate* if $B(v, w) = 0$ for all w implies that $v = 0$, and likewise $B(v, w) = 0$ for all v implies that $w = 0$. In this case we have $\det \Gamma \neq 0$. Suppose B is nondegenerate. If $T : V \longrightarrow V$ is linear and satisfies $B(Tv, Tw) = B(v, w)$ for all $v, w \in V$, then $\det(T) \neq 0$ by formula (1.1). Hence $T \in \mathrm{O}(B)$. The next two subsections will discuss the most important special cases of this class of groups.

Orthogonal Groups

We start by introducing the matrix groups; later we will identify these groups with isometry groups of certain classes of bilinear forms. Let $\mathrm{O}(n, \mathbb{F})$ denote the set of all $g \in \mathrm{GL}(n, \mathbb{F})$ such that $gg^t = I$. That is,

$$g^t = g^{-1}.$$

We note that $(AB)^t = B^t A^t$ and if $A, B \in \mathrm{GL}(n, \mathbb{F})$ then $(AB)^{-1} = B^{-1}A^{-1}$. It is therefore obvious that $\mathrm{O}(n, \mathbb{F})$ is a subgroup of $\mathrm{GL}(n, \mathbb{F})$. This group is called the *orthogonal group* of $n \times n$ matrices over $\mathbb{F}$. If $\mathbb{F} = \mathbb{R}$ we introduce the *indefinite orthogonal groups*, $\mathrm{O}(p, q)$, with $p + q = n$ and $p, q \in \mathbb{N}$. Let

$$I_{p,q} = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix}$$

with I_r denoting the $r \times r$ identity matrix. Then we define

$$\mathrm{O}(p, q) = \{g \in M_n(\mathbb{R}) \,:\, g^t I_{p,q} g = I_{p,q}\}.$$

We note that $\mathrm{O}(n, 0) = \mathrm{O}(0, n) = \mathrm{O}(n, \mathbb{R})$. Also, if

$$\sigma = \begin{bmatrix} 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \end{bmatrix}$$

is the matrix with entries 1 on the skew diagonal ($j = n + 1 - i$) and all other entries 0, then $\sigma I_{p,q} \sigma^{-1} = \sigma I_{p,q} \sigma = \sigma I_{p,q} \sigma^t = -I_{q,p}$. Thus the map

$$\varphi : \mathrm{O}(p, q) \longrightarrow \mathrm{GL}(n, \mathbb{R})$$

given by $\varphi(g) = \sigma g \sigma$ defines an isomorphism of $\mathrm{O}(p, q)$ onto $\mathrm{O}(q, p)$.

We will now describe these groups in terms of bilinear forms.

Definition 1.1.1. *Let V be a vector space over $\mathbb{R}$ and let M be a symmetric bilinear form on V. The form M is* positive definite *if $M(v, v) > 0$ for every $v \in V$ with $v \neq 0$.*

Lemma 1.1.2. *Let V be an n-dimensional vector space over $\mathbb{F}$ and let B be a symmetric nondegenerate bilinear form over $\mathbb{F}$.*

(1) *If $\mathbb{F} = \mathbb{C}$ then there exists a basis $\{v_1, \ldots, v_n\}$ of V such that $B(v_i, v_j) = \delta_{ij}$.*

(2) *If $\mathbb{F} = \mathbb{R}$ then there exist integers $p, q \geq 0$ with $p + q = n$ and a basis $\{v_1, \ldots, v_n\}$ of V such that $B(v_i, v_j) = \varepsilon_i \delta_{ij}$ with $\varepsilon_i = 1$ for $i \leq p$ and $\varepsilon_i = -1$ for $i > p$. Furthermore, if we have another such basis then the corresponding integers (p, q) are the same.*

Remark 1.1.3. The basis for V in part (2) is called a *pseudo-orthonormal basis* relative to B, and the number $p - q$ is called the *signature* of the form (we will also call (p, q) the signature of B). A form is positive-definite if and only if its signature is n. In this case a pseudo-orthonormal basis is an orthonormal basis in the usual sense

Proof. We first observe that if M is a symmetric bilinear form on V such that $M(v, v) = 0$ for all $v \in V$, then $M = 0$. Indeed, using the symmetry and bilinearity we have

$$4M(v, w) = M(v + w,\, v + w) - M(v - w,\, v - w) = 0 \tag{1.2}$$

for all $v, w \in V$.

We now construct a basis $\{w_1, \ldots, w_n\}$ of V such that

$$B(w_i,\, w_j) = 0 \quad \text{for } i \neq j \text{ and} \quad B(w_i,\, w_i) \neq 0$$

(such a basis is called an *orthogonal basis* with respect to B). The argument is by induction on n. Since B is nondegenerate, there exists a vector $w_n \in V$ with $B(w_n, w_n) \neq 0$ by (1.2). If $n = 1$ we are done. If $n > 1$, set

$$V' = \{v \in V \,:\, B(w_n,\, v) = 0\}.$$

For $v \in V$ set

$$v' = v - \frac{B(v,\, w_n)}{B(w_n,\, w_n)} w_n.$$

Clearly, $v' \in V'$, so we have $V = V' + \mathbb{F}w_n$. In particular, this shows that $\dim V' = n - 1$. We assert that the form $B' = B|_{V' \times V'}$ is nondegenerate on V'. Indeed, if $v \in V'$ satisfies $B(v',\, w) = 0$ for all $w \in V'$, then $B(v',\, w) = 0$ for all $w \in V$, since $B(v',\, w_n) = 0$. Hence $v' = 0$, proving nondegeneracy of B'. We may assume by induction that there exists a B'-orthogonal basis $\{w_1, \ldots, w_{n-1}\}$ for V'. Then it is clear that $\{w_1, \ldots, w_n\}$ is a B-orthogonal basis for V.

If $\mathbb{F} = \mathbb{C}$ let $\{w_1, \ldots, w_n\}$ be an orthogonal basis of V with respect to B and let $z_i \in \mathbb{C}$ be a choice of square root of $B(w_i,\, w_i)$. Setting $v_i = (z_i)^{-1} w_i$, we then obtain the desired normalization $B(v_i,\, v_j) = \delta_{ij}$.

Now let $\mathbb{F} = \mathbb{R}$. We rearrange the indices (if necessary) so that $B(w_i,\, w_i) \geq B(w_{i+1},\, w_{i+1})$ for $i = 1, \ldots, n - 1$. Let $p = 0$ if $B(w_1,\, w_1) < 0$. Otherwise, let

$$p = \max\{i \,:\, B(w_i,\, w_i) > 0\}.$$

Then $B(w_i,\, w_i) < 0$ for $i > p$. Take z_i to be a square root of $B(w_i,\, w_i)$ for $i \leq p$, and take z_i to be a square root of $-B(w_i,\, w_i)$ for $i > p$. Setting $v_i = (z_i)^{-1} w_i$, we now have $B(v_i,\, v_j) = \varepsilon_i \delta_{ij}$.

We are left with proving that the integer p is intrinsic to B. Take any basis $\{v_1, \ldots, v_n\}$ such that $B(v_i,\, v_j) = \varepsilon_i \delta_{ij}$ with $\varepsilon_i = 1$ for $i \leq p$ and $\varepsilon_i = -1$ for $i > p$. Set

$$V_+ = \mathrm{Span}\{v_1, \ldots, v_p\}, \qquad V_- = \mathrm{Span}\{v_{p+1}, \ldots, v_n\}.$$

Then $V = V_+ \oplus V_-$ (direct sum). Let $\pi : V \longrightarrow V_+$ be the projection onto the first factor. We note that $B|_{V_+ \times V_+}$ is positive definite. Let W be any subspace of V such that $B|_{W \times W}$ is positive definite. Suppose that $w \in W$ and $\pi(w) = 0$. Then $w \in V_-$, so it can be written as $w = \sum_{i>p} a_i v_i$. Hence

$$B(w,w) = \sum_{i,j>p} a_i a_j \, B(v_i,\, v_j) = -\sum_{i>p} a_i^2 \leq 0.$$

Since $B|_{W \times W}$ has been assumed to be positive definite, it follows that $w = 0$. This implies that $\pi : W \longrightarrow V_+$ is injective, and hence $\dim W \leq \dim V_+ = p$. Thus p is uniquely determined as the maximum dimension of a subspace on which B is positive definite. ⧫

The following result follows immediately from Lemma 1.1.2.

Proposition 1.1.4. *Let B be a nondegenerate symmetric bilinear form on an n-dimensional vector space V over $\mathbb{F}$.*

(1) *Let $\mathbb{F} = \mathbb{C}$. If $\{v_1, \ldots, v_n\}$ is an orthonormal basis for V with respect to B, then $\mu : \mathrm{O}(V,B) \longrightarrow \mathrm{O}(n,\mathbb{F})$ defines a group isomorphism.*

(2) *Let $\mathbb{F} = \mathbb{R}$. If B has signature $(p, n-p)$ and $\{v_1, \ldots, v_n\}$ is a pseudo-orthonormal basis of V, then $\mu : \mathrm{O}(V,B) \longrightarrow \mathrm{O}(p, n-p)$ is a group isomorphism.*

Here $\mu(g)$, for $g \in \mathrm{GL}(V)$, is the matrix of g with respect to the given basis.

The *special orthogonal group* over $\mathbb{F}$ is the subgroup

$$\mathrm{SO}(n,\mathbb{F}) = \mathrm{O}(n,\mathbb{F}) \cap \mathrm{SL}(n,\mathbb{F})$$

of $\mathrm{O}(n,\mathbb{F})$. The *indefinite special orthogonal groups* are the groups

$$\mathrm{SO}(p,q) = \mathrm{O}(p,q) \cap \mathrm{SL}(p+q,\mathbb{R}).$$

Symplectic Group

We set

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$$

with I the $n \times n$ identity matrix. The symplectic group of rank n over $\mathbb{F}$ is defined to be

$$\mathrm{Sp}(n,\mathbb{F}) = \{g \in M_{2n}(\mathbb{F}) \,:\, g^t J g = J\}.$$

As in the case of the orthogonal groups one sees without difficulty that $\mathrm{Sp}(n,\mathbb{F})$ is a subgroup of $\mathrm{GL}(2n,\mathbb{F})$.

We will now look at the coordinate-free version of these groups. A bilinear form B is called *skew symmetric* if $B(v,w) = -B(w,v)$. If B is skew-symmetric and nondegenerate, then $m = \dim V$ must be even, since the matrix of B relative to any basis for V is skew-symmetric and has nonzero determinant.

Lemma 1.1.5. *Let V be a $2n$-dimensional vector space over $\mathbb{F}$ and let B be a nondegenerate, skew-symmetric bilinear form on V. Then there exists a basis $\{v_1, \ldots, v_{2n}\}$ for V such that the matrix $[B(v_i, v_j)] = J$ (call such a basis a B-*symplectic basis*).*

Proof. Let v be a nonzero element of V. Since B is nondegenerate, there exists $w \in V$ with $B(v, w) \neq 0$. Replacing w with $B(v, w)^{-1}w$, we may assume that $B(v, w) = 1$. Let

$$W = \{x \in V \,:\, B(v, x) = 0 \text{ and } B(w, x) = 0\}.$$

For $x \in V$ we set $x' = x - B(v, x)w - B(x, w)v$. Then

$$B(v, x') = B(v, x) - B(v, x)B(v, w) - B(w, x)B(v, v) = 0,$$

since $B(v, w) = 1$ and $B(v, v) = 0$ (by skew symmetry of B). Similarly,

$$B(w, x') = B(w, x) - B(v, x)B(w, w) + B(w, x)B(w, v) = 0,$$

since $B(w, v) = -1$ and $B(w, w) = 0$. Thus $V = U \oplus W$, where U is the span of v and w. It is easily verified that $B|_{U\times U}$ is nondegenerate, and so $U \cap W = \{0\}$. This implies that $\dim W = m - 2$. We leave to the reader to check that $B|_{W\times W}$ also is nondegenerate.

Set $v_n = v$ and $v_{2n} = w$ with v, w as above. Since $B|_{W\times W}$ is nondegenerate, by induction there exists a B-symplectic basis $\{w_1, \ldots, w_{2n-2}\}$ of W. Set $v_i = w_i$ and $v_{n+1-i} = w_{n-i}$ for $i \leq n - 1$. Then $\{v_1, \ldots, v_{2n}\}$ is a B-symplectic basis for V. ♦

The following result follows immediately from Lemma 1.1.5.

Proposition 1.1.6. *Let V be a $2n$-dimensional vector space over $\mathbb{F}$ and let B be a nondegenerate skew-symmetric bilinear form on V. Fix a a B-symplectic basis of V and let $\mu(g)$, for $g \in \mathrm{GL}(V)$, be the matrix of g with respect to this basis. Then $\mu : \mathrm{O}(V, B) \longrightarrow \mathrm{Sp}(n, \mathbb{F})$ is a group isomorphism.*

1.1.3 Unitary Groups

Another family of classical subgroups of $\mathrm{GL}(n, \mathbb{C})$ consists of the unitary groups and special unitary groups for definite and indefinite Hermitian forms. If $A \in M_n(\mathbb{C})$ we will use the standard notation $A^* = \overline{A}^t$ for its adjoint matrix, where $\overline{A}$ is the matrix obtained from A by complex conjugating all of the entries. The *unitary group* of rank n is the group

$$\mathrm{U}(n) = \{g \in M_n(\mathbb{C}) \,:\, g^*g = I\}.$$

The *special unitary group* is $\mathrm{SU}(n) = \mathrm{U}(n) \cap \mathrm{SL}(n, \mathbb{C})$. Let the matrix $I_{p,q}$ be as in Section 1.1.2. We define the *indefinite unitary group* of signature (p, q) to be

$$\mathrm{U}(p, q) = \{g \in M_n(\mathbb{C}) \,:\, g^*I_{p,q}g = I_{p,q}\}.$$

The special indefinite unitary group of signature (p, q) is $\mathrm{SU}(p, q) = \mathrm{U}(p, q) \cap \mathrm{SL}(n, \mathbb{C})$.

We will now obtain a coordinate-free description of these groups. Let V be an n-dimensional vector space over $\mathbb{C}$. An $\mathbb{R}$ bilinear map $B : V \times V \longrightarrow \mathbb{C}$ (where we view V as a vector space over $\mathbb{R}$) is said to be a *Hermitian* form if it satisfies

(1) $B(av, w) = aB(v, w)$ for all $a \in \mathbb{C}$ and all $v, w \in V$.

(2) $B(w, v) = \overline{B(v, w)}$ for all $v, w \in V$.

By the second condition, we see that a Hermitian form is nondegenerate provided $B(v, w) = 0$ for all $w \in V$ implies that $v = 0$. The form is said to be *positive definite* if $B(v, v) > 0$ for all $v \in V$ with $v \neq 0$. (Note that if M is a Hermitian form, then $M(v, v) \in \mathbb{R}$ for all $v \in V$.) We define $\mathrm{U}(V, B)$ (also denoted $\mathrm{U}(B)$ when V is understood) to be the group of all elements, g, of $\mathrm{GL}(V)$ such that $B(gv, gw) = B(v, w)$ for all $v, w \in V$. We call $\mathrm{U}(B)$ the *unitary group* of B.

Lemma 1.1.7. *Let V be an n-dimensional vector space over $\mathbb{C}$ and let B be a nondegenerate Hermitian form on V. Then there exists an integer p, with $n \geq p \geq 0$, and a basis $\{v_1, \ldots, v_n\}$ of V, such that*

$$B(v_i, v_j) = \varepsilon_i \delta_{ij},$$

with $\varepsilon_i = 1$ for $i \leq p$ and $\varepsilon_i = -1$ for $i > p$. The number p depends only on B and not on the choice of basis.

The proof of Lemma 1.1.7 is almost identical to that of Lemma 1.1.2 and will be left as an exercise.

If V is an n-dimensional vector space over $\mathbb{C}$ and B is a nondegenerate Hermitian form on V, then a basis as in Lemma 1.1.7 will be called a *pseudo-orthonormal basis* (if $p = n$ then it is an *orthonormal basis* in the usual sense). The pair $(p, n - p)$ will be called the *signature* of B. The following result is proved in exactly the same way as the corresponding result for orthogonal groups.

Proposition 1.1.8. *Let V be a finite dimensional vector space over $\mathbb{C}$ and let B be a nondegenerate Hermitian form on V of signature (p, q). Fix a pseudo-orthonormal basis of V relative to B and let $\mu(g)$, for $g \in \mathrm{GL}(V)$, be the matrix of g with respect to this basis. Then $\mu : U(V, B) \longrightarrow U(p, q)$ is a group isomorphism.*

1.1.4 Quaternionic Groups

We recall some basic properties of the quaternions. Consider the four-dimensional real vector space $\mathbb{H}$ consisting of the 2×2 complex matrices

$$w = \begin{bmatrix} x & -\overline{y} \\ y & \overline{x} \end{bmatrix} \quad \text{with } x, y \in \mathbb{C}. \tag{1.3}$$

One checks directly that $\mathbb{H}$ is closed under multiplication in $M_2(\mathbb{C})$. If $w \in \mathbb{H}$ then $w^* \in \mathbb{H}$ and

$$w^* w = w w^* = (|x|^2 + |y|^2) I$$

(where w^* denotes the conjugate-transpose matrix). Hence every nonzero element of $\mathbb{H}$ is invertible. Thus $\mathbb{H}$ is a *division algebra* (or skew field) over $\mathbb{R}$. This division algebra is a realization of the quaternions.

The more usual way of introducing the quaternions is to consider the vector space, $\mathbb{H}$, over $\mathbb{R}$ with basis $\{\mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$. Define a multiplication so that $\mathbf{1}$ is the identity and

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -\mathbf{1},$$

$$\mathbf{ij} = -\mathbf{ji} = \mathbf{k}, \quad \mathbf{ki} = -\mathbf{ik} = \mathbf{j}, \quad \mathbf{jk} = -\mathbf{kj} = \mathbf{i},$$

and extend the multiplication to $\mathbb{H}$ by linearity. To obtain an isomorphism between this version of $\mathbb{H}$ and the 2×2 complex matrix version, take

$$\mathbf{1} = I, \quad \mathbf{i} = \begin{bmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{bmatrix}, \quad \mathbf{j} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \mathbf{k} = \begin{bmatrix} 0 & \mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix},$$

where i is a fixed choice of $\sqrt{-1}$. The *conjugation* $w \mapsto w^*$ satisfies $(uv)^* = v^*u^*$. In terms of real components, $(a+b\mathbf{i}+c\mathbf{j}+d\mathbf{k})^* = a-b\mathbf{i}-c\mathbf{j}-d\mathbf{k}$ for $a, b, c, d \in \mathbb{R}$. It is useful to write quaternions in complex form as $x + \mathbf{j}y$ with $x, y \in \mathbb{C}$; however, note that the conjugation is then given as

$$(x + \mathbf{j}y)^* = \overline{x} + \overline{y}\mathbf{j} = \overline{x} - \mathbf{j}y.$$

On the $4n$-dimensional real vector space $\mathbb{H}^n$ we define multiplication by $a \in \mathbb{H}$ on the right:

$$(u_1, \ldots, u_n) \cdot a = (u_1 a, \ldots, u_n a).$$

We note that $u \cdot \mathbf{1} = u$ and $u \cdot (ab) = (u \cdot a) \cdot b$. We can therefore think of $\mathbb{H}^n$ as a vector space over $\mathbb{H}$. Viewing elements of $\mathbb{H}^n$ as $n \times 1$ column vectors, we define Au for $u \in \mathbb{H}^n$ and $A \in M_n(\mathbb{H})$ by matrix multiplication. Then $A(u \cdot a) = (Au) \cdot a$ for $a \in \mathbb{H}$; hence A defines a quaternionic linear map. Here matrix multiplication is defined as usual but one must be careful about the order of multiplication of the entries.

We can make $\mathbb{H}^n$ into a $2n$-dimensional vector space over $\mathbb{C}$ in many ways; for example, we can embed $\mathbb{C}$ into $\mathbb{H}$ as any of the subfields

$$\mathbb{R}\mathbf{1} + \mathbb{R}\mathbf{i}, \quad \mathbb{R}\mathbf{1} + \mathbb{R}\mathbf{j}, \quad \mathbb{R}\mathbf{1} + \mathbb{R}\mathbf{k}. \tag{1.4}$$

Using the first of these embeddings, we write $z = x + \mathbf{j}y \in \mathbb{H}^n$ with $x, y \in \mathbb{C}^n$, and likewise $C = A + \mathbf{j}B \in M_n(\mathbb{H})$ with $A, B \in M_n(\mathbb{C})$. The maps

$$z \mapsto \begin{bmatrix} x \\ y \end{bmatrix} \quad \text{and} \quad C \mapsto \begin{bmatrix} A & -\overline{B} \\ B & \overline{A} \end{bmatrix}$$

identify $\mathbb{H}^n$ with $\mathbb{C}^{2n}$ and $M_n(\mathbb{H})$ with the real subalgebra of $M_{2n}(\mathbb{C})$ consisting of matrices T such that

$$JT = \overline{T}J, \quad \text{where} \quad J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}. \tag{1.5}$$

We define $\mathrm{GL}(n,\mathbb{H})$ to be the group of all invertible $n \times n$ matrices over $\mathbb{H}$. Then $\mathrm{GL}(n,\mathbb{H})$ acts on $\mathbb{H}^n$ by complex linear transformations relative to each of the complex structures (1.4). If we use the embedding of $M_n(\mathbb{H})$ into $M_{2n}(\mathbb{C})$ just described, then from (1.5) we see that

$$\mathrm{GL}(n,\mathbb{H}) = \{g \in \mathrm{GL}(2n,\mathbb{C}) \,:\, Jg = \overline{g}J\}.$$

Quaternionic Special Linear Group

We leave it to the reader to prove that the determinant of $A \in \mathrm{GL}(n,\mathbb{H})$ as a complex linear transformation with respect to any of the complex structures (1.4) is the same. We can thus define $\mathrm{SL}(n,\mathbb{H})$ to be the elements of determinant one in $\mathrm{GL}(n,\mathbb{H})$ with respect to any of these complex structures. This group is usually denoted as $\mathrm{SU}^*(2n)$.

The Quaternionic Unitary Groups

For $X = [x_{ij}] \in M_n(\mathbb{H})$ we define $X^* = [x_{ji}^*]$ (here we take the quaternionic matrix entries $x_{ij} \in M_2(\mathbb{C})$ given by (1.3)). Let the diagonal matrix $I_{p,q}$ (with $p+q=n$) be as in Section 1.1.2. The *indefinite quaternionic unitary groups* are the groups

$$\mathrm{Sp}(p,q) = \{g \in \mathrm{GL}(p+q,\mathbb{H}) \,:\, g^* I_{p,q} g = I_{p,q}\}.$$

We leave it to the reader to prove that this set is a subgroup of $\mathrm{GL}(p+q,\mathbb{H})$.

The group $\mathrm{Sp}(p,q)$ is the isometry group of the nondegenerate *quaternionic Hermitian form*

$$B(w,z) = w^* I_{p,q} z, \quad \text{for } w,z \in \mathbb{H}^n. \tag{1.6}$$

(Note that this form satisfies $B(w,z) = B(z,w)^*$ and $B(w\alpha, z\beta) = \alpha^* B(w,z)\beta$ for $\alpha,\beta \in \mathbb{H}$.) If we write $w = u + \mathbf{j}v$ and $z = x + \mathbf{j}y$ with $u,v,x,y \in \mathbb{C}^n$, and set $K_{p,q} = \mathrm{diag}[I_{p,q}\ \ I_{p,q}] \in M_{2n}(\mathbb{R})$, then

$$B(w,z) = \begin{bmatrix} u^* & v^* \end{bmatrix} K_{p,q} \begin{bmatrix} x \\ y \end{bmatrix} + \mathbf{j} \begin{bmatrix} u^t & v^t \end{bmatrix} K_{p,q} \begin{bmatrix} -y \\ x \end{bmatrix}.$$

Thus the elements of $\mathrm{Sp}(p,q)$, viewed as linear transformations of $\mathbb{C}^{2n}$, preserve both a Hermitian form of signature $(2p,2q)$ and a nondegenerate skew-symmetric form.

The Group $\mathrm{SO}^*(2n)$

Let J be the $2n \times 2n$ skew-symmetric matrix from Section 1.1.2. Since $J^2 = -I_{2n}$ (the $2n \times 2n$ identity matrix), the map of $\mathrm{GL}(2n,\mathbb{C})$ to itself given by $\theta(g) = -JgJ$ defines an automorphism whose square is the identity. Our last family of classical groups is

$$\mathrm{SO}^*(2n) = \{g \in \mathrm{SO}(2n,\mathbb{C}) \,:\, \theta(\overline{g}) = g\}.$$

We identify $\mathbb{C}^{2n}$ with $\mathbb{H}^n$ as a vector space over $\mathbb{C}$ by the map $\begin{bmatrix} a \\ b \end{bmatrix} \mapsto a + \mathbf{j}b$, where $a, b \in \mathbb{C}^n$. The group $\mathrm{SO}^*(2n)$ then becomes the isometry group of the nondegenerate *quaternionic skew-Hermitian form*

$$C(x, y) = x^* \mathbf{j} y, \quad \text{for } x, y \in \mathbb{H}^n. \tag{1.7}$$

This form satisfies $C(x, y) = -C(y, x)^*$ and $C(x\alpha, y\beta) = \alpha^* C(x, y)\beta$ for $\alpha, \beta \in \mathbb{H}$.

We have now completed the list of the classical groups associated with $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{H}$. We will return to this list at the end of the chapter when we consider *real forms* of complex algebraic groups. Later we will define *covering groups*; any group covering one of the groups on this list—for example, a *spin group* in Chapter 7—will also be called a classical group.

1.1.5 Exercises

In these exercises $\mathbb{F}$ denotes either $\mathbb{R}$ or $\mathbb{C}$. See Appendix B.2 for notations and properties of tensor and exterior products of vector spaces.

1. Let $\{v_1, \ldots, v_n\}$ and $\{w_1, \ldots, w_n\}$ be bases for an $\mathbb{F}$ vector space V, and let $T : V \longrightarrow V$ be a linear map with matrices A and B, respectively, relative to these bases. Show that $\det A = \det B$.

2. Determine the signature of the form $B(x, y) = \sum_{i=1}^{n} x_i y_{n+1-i}$ on $\mathbb{R}^n$.

3. Let V be a vector space over $\mathbb{F}$ and let B be a skew-symmetric or symmetric nondegenerate bilinear form on V. Assume that W is a subspace of V on which B restricts to a nondegenerate form. Prove that the restriction of B to $W^\perp = \{v \in V \,:\, B(v, w) = 0 \text{ for all } w \in W\}$ is nondegenerate.

4. Let V denote the vector space of symmetric 2×2 matrices over $\mathbb{F}$. If $x, y \in V$ define $B(x, y) = \det(x + y) - \det(x) - \det(y)$.

   (a) Show that B is nondegenerate, and that if $\mathbb{F} = \mathbb{R}$ then the signature of the form B is $(1,\ 2)$.

   (b) If $g \in \mathrm{SL}(2, \mathbb{F})$ define $\varphi(g) \in \mathrm{GL}(V)$ by $\varphi(g)(v) = gvg^t$. Show that $\varphi : \mathrm{SL}(2, \mathbb{F}) \longrightarrow \mathrm{SO}(V, B)$ is a group homomorphism with kernel $\{\pm I\}$.

5. The purpose of this exercise is to prove Lemma 1.1.7 by the method of proof of Lemma 1.1.2.

   (a) Prove that if M is Hermitian form such that $M(v, v) = 0$ for all v then $M = 0$. (HINT: Show that $M(v + sw, v + sw) = sM(w, v) + \overline{s}M(v, w)$ for all $s \in \mathbb{C}$, then substitute values for s to see that $M(v, w) = 0$.)

   (b) Use the result of part (a) to complete the proof of Lemma 1.1.7.

   (HINT: Note that $M(v, v) \in \mathbb{R}$ since M is Hermitian.)

6. Let V be a $2n$-dimensional vector space over $\mathbb{F}$. Consider the space $W = \bigwedge^n V$. Fix a basis ω of the one-dimensional vector space $\bigwedge^{2n} V$. Consider the bilinear form $B(u, v)$ on W defined by $u \wedge v = B(u, v)\omega$.

   (a) Show that B is nondegenerate.

   (b) Show that B is skew symmetric if n is odd and symmetric if n is even.

   (c) Determine the signature of B when n is even and $\mathbb{F} = \mathbb{R}$,

7. *(Notation of the previous exercise)* Let $V = \mathbb{F}^4$ with basis $\{e_1, e_2, e_3, e_4\}$ and let $\omega = e_1 \wedge e_2 \wedge e_3 \wedge e_4$. Define $\varphi(g)(u \wedge v) = gu \wedge gv$ for $g \in \mathrm{SL}(4, \mathbb{F})$ and $u, v \in \mathbb{F}^4$. Show that $\varphi : \mathrm{SL}(4, \mathbb{F}) \longrightarrow \mathrm{SO}(\bigwedge^2 \mathbb{F}^4, B)$ is a group homomorphism with kernel $\{\pm I\}$. (HINT: Use Jordan canonical form to determine the kernel.)

8. *(Notation of the previous exercise)* Let ψ be the restriction of φ to $\mathrm{Sp}(2, \mathbb{F})$. Let $\nu = e_1 \wedge e_3 + e_2 \wedge e_4$.

   (a) Show that $\psi(g)\nu = \nu$ and $B(\nu, \nu) = -2$. (HINT: Show that the map $e_i \wedge e_j \mapsto e_{ij} - e_{ji}$ is a linear isomorphism between $\bigwedge^2 \mathbb{F}^4$ and the subspace of skew-symmetric matrices in $M_4(\mathbb{F})$ that takes ν to J, and that $\varphi(g)$ becomes the transformation $A \mapsto gAg^t$.)

   (b) Let $W = \{w \in \bigwedge^2 \mathbb{F}^4 \,:\, B(\nu, w) = 0\}$. Show that $B|_{W \times W}$ is nondegenerate and has signature $(3, 2)$ when $\mathbb{F} = \mathbb{R}$.

   (c) Set $\rho(g) = \psi(g)|_W$. Show that ρ is a group homomorphism from $\mathrm{Sp}(2, \mathbb{F})$ to $\mathrm{SO}(W, B|_{W \times W})$ with kernel $\{\pm 1\}$. (HINT: Use the previous exercise to determine the kernel.)

9. Let $V = M_2(\mathbb{F})$. For $x, y \in V$ define $B(x, y) = \det(x+y) - \det(x) - \det(y)$.

   (a) Show that B is a symmetric nondegenerate form on V, and calculate the signature of B when $\mathbb{F} = \mathbb{R}$.

   (b) Let $G = \mathrm{SL}(2, \mathbb{F}) \times \mathrm{SL}(2, \mathbb{F})$ and define $\varphi : G \longrightarrow \mathrm{GL}(V)$ by $\varphi(a, b)v = axb^t$ for $a, b \in \mathrm{SL}(2, \mathbb{F})$ and $v \in V$. Show that φ is a group homomorphism and $\varphi(G) \subset \mathrm{SO}(V, B)$. Determine $\mathrm{Ker}(\varphi)$. (HINT: Use Jordan canonical form to determine the kernel.)

10. Identify $\mathbb{H}^n$ with $\mathbb{C}^{2n}$ as a vector space over $\mathbb{C}$ by the map $a + \mathbf{j}b \mapsto \begin{bmatrix} a \\ b \end{bmatrix}$, where $a, b \in \mathbb{C}^n$. Let $T = A + \mathbf{j}B \in M_n(\mathbb{H})$ with $A, B \in M_n(\mathbb{C})$.

    (a) Show that left multiplication by T on $\mathbb{H}^n$ corresponds to multiplication by the matrix $\begin{bmatrix} A & -\bar{B} \\ B & \bar{A} \end{bmatrix} \in M_{2n}(\mathbb{C})$ on $\mathbb{C}^{2n}$.

    (b) Show that multiplication by $\mathbf{i}$ on $M_n(\mathbb{H})$ becomes the transformation

$$\begin{bmatrix} A & -\bar{B} \\ B & -\bar{A} \end{bmatrix} \mapsto \begin{bmatrix} \mathrm{i}A & -\mathrm{i}\bar{B} \\ -\mathrm{i}B & -\mathrm{i}\bar{A} \end{bmatrix}.$$

11. Use the identification of $\mathbb{H}^n$ with $\mathbb{C}^{2n}$ in the previous exercise to view the form $B(x, y)$ in equation (1.6) as an $\mathbb{H}$-valued function on $\mathbb{C}^{2n} \times \mathbb{C}^{2n}$.

    (a) Show that $B(x, y) = \overline{B_0(x, y)} + \mathbf{j}B_1(x, y)$, where B_0 is a $\mathbb{C}$-Hermitian form on $\mathbb{C}^{2n}$ of signature $(2p, 2q)$ and B_1 is a nondegenerate skew-symmetric $\mathbb{C}$-bilinear form on $\mathbb{C}^{2n}$.

    (b) Use part (a) to prove that $\mathrm{Sp}(p, q) = \mathrm{Sp}(\mathbb{C}^{2n}, B_1) \cap \mathrm{U}(\mathbb{C}^{2n}, B_0)$.

12. Use the identification of $\mathbb{H}^n$ with $\mathbb{C}^{2n}$ in the previous exercise to view the form $C(x, y)$ from equation (1.7) as an $\mathbb{H}$-valued function on $\mathbb{C}^{2n} \times \mathbb{C}^{2n}$.

    (a) Show that $C(x, y) = \overline{C_0(x, y)} + \mathbf{j}x^t y$ for $x, y \in \mathbb{C}^{2n}$, where $C_0(x, y)$ is a $\mathbb{C}$-Hermitian form on $\mathbb{C}^{2n}$ of signature (n, n). .

    (b) Use the result of part (a) to prove that $\mathrm{SO}^*(2n) = \mathrm{SO}(2n, \mathbb{C}) \cap \mathrm{U}(\mathbb{C}^{2n}, C_0)$.

13. Why can't we just define $\mathrm{SL}(n, \mathbb{H})$ by taking all $g \in \mathrm{GL}(n, \mathbb{H})$ such that the usual formula for the determinant of g yields 1?

14. Consider the three embeddings of $\mathbb{C}$ in $\mathbb{H}$ given by the subfields (1.4). These give three ways of writing $X \in M_n(\mathbb{H})$ as a $2n \times 2n$ matrix over $\mathbb{C}$. Show that these three matrices have the same determinant.

1.2 The Classical Lie Algebras

Let V be a vector space over $\mathbb{F}$. Let $\mathrm{End}(V)$ denote the algebra (under composition) of $\mathbb{F}$-linear maps of V to V. If $X, Y \in \mathrm{End}(V)$ then we set $[X, Y] = XY - YX$. This defines a new product on $\mathrm{End}(V)$ that satisfies two properties:

(1) $[X, Y] = -[Y, X]$ for all X, Y *(skew symmetry)*.

(2) $[X, [Y, Z]] = [[X, Y], Z] + [Y, [X, Z]]$ for all X, Y, Z *(Jacobi identity)*.

Definition 1.2.1. *A vector space $\mathfrak{g}$ over $\mathbb{F}$ together with a bilinear map $X, Y \mapsto [X, Y]$ of $\mathfrak{g} \times \mathfrak{g}$ to $\mathfrak{g}$ is said to be a* Lie algebra *if conditions* (1) *and* (2) *are satisfied.*

Thus, in particular, we see that $\mathrm{End}(V)$ is a Lie algebra under the binary operation $[X, Y] = XY - YX$. Condition (2) is a substitute for the associative rule; it says that for fixed X, the linear transformation $Y \mapsto [X, Y]$ is a *derivation* of the (non associative) algebra $(\mathfrak{g}, [\cdot\,, \cdot])$.

If $\mathfrak{g}$ is a Lie algebra and if $\mathfrak{h}$ is a subspace such that $X, Y \in \mathfrak{h}$ implies that $[X, Y]$ in $\mathfrak{h}$, then $\mathfrak{h}$ is a Lie algebra under the restriction of $[\cdot\,, \cdot]$. We will call $\mathfrak{h}$ a *Lie subalgebra* of $\mathfrak{g}$ (or *subalgebra*, when the Lie algebra context is clear).

Suppose that $\mathfrak{g}$ and $\mathfrak{h}$ are Lie algebras over $\mathbb{F}$. A *Lie algebra homomorphism* of $\mathfrak{g}$ to $\mathfrak{h}$ is an $\mathbb{F}$-linear map $T : \mathfrak{g} \longrightarrow \mathfrak{h}$ such that $T[X, Y] = [TX, TY]$ for all $X, Y \in \mathfrak{g}$. A Lie algebra homomorphism is an *isomorphism* if it is bijective.

1.2.1 General and Special Linear Lie Algebras

If V is a vector space over $\mathbb{F}$, we write $\mathfrak{gl}(V)$ for $\mathrm{End}(V)$ looked upon as a Lie algebra under $[X, Y] = XY - YX$. We write $\mathfrak{gl}(n, \mathbb{F})$ to denote $M_n(\mathbb{F})$ as a Lie algebra under the matrix commutator bracket. If $\dim V = n$ and we fix a basis for V, then the correspondence between linear transformations and their matrices gives a Lie-algebra isomorphism $\mathfrak{gl}(V) \cong \mathfrak{gl}(n, \mathbb{R})$. These Lie algebras will be called the *general linear Lie algebras*.

If $A = [a_{ij}] \in M_n(\mathbb{F})$ then its *trace* is $\mathrm{tr}(A) = \sum_i a_{ii}$. We note that

$$\mathrm{tr}(AB) = \mathrm{tr}(BA).$$

This implies that if A is the matrix of $T \in \mathrm{End}(V)$ with respect to some basis, then $\mathrm{tr}(A)$ is independent of the choice of basis. We will write $\mathrm{tr}(T) = \mathrm{tr}(A)$. We define

$$\mathfrak{sl}(V) = \{T \in \mathrm{End}(V) \,:\, \mathrm{tr}(T) = 0\}.$$

Since $\mathrm{tr}([S, T]) = 0$ for all $S, T \in \mathrm{End}(V)$, we see that $\mathfrak{sl}(V)$ is a Lie subalgebra of $\mathfrak{gl}(V)$. Choosing a basis for V, we may identify this Lie algebra with

$$\mathfrak{sl}(n, \mathbb{F}) = \{A \in \mathfrak{gl}(n, \mathbb{F}) \,:\, \mathrm{tr}(A) = 0\}.$$

These Lie algebras will be called the *special linear Lie algebras*.

1.2.2 Lie Algebras Associated with Bilinear Forms

Let V be a vector space over $\mathbb{F}$ and let $B : V \times V \longrightarrow \mathbb{F}$ be a bilinear map. We define

$$\mathfrak{so}(V, B) = \{X \in \mathrm{End}(V) \,:\, B(Xv, w) = -B(v, Xw)\}.$$

Thus $\mathfrak{so}(V, B)$ consists of the linear transformations that are *skew symmetric* relative to the form B, and is obviously a linear subspace of $\mathfrak{gl}(V)$. If $X, Y \in \mathfrak{so}(V, B)$, then

$$B(XYv,\, w) = -B(Yv,\, Xw) = B(v,\, YXw).$$

It follows that $B([X, Y]v, w) = -B(v, [X, Y]w)$, and hence $\mathfrak{so}(V, B)$ is a Lie subalgebra of $\mathfrak{gl}(V)$.

Suppose V is finite-dimensional. Fix a basis $\{v_1, \ldots, v_n\}$ for V and let Γ be the $n \times n$ matrix with entries $\Gamma_{ij} = B(v_i, v_j)$. By a calculation analogous to that in Section 1.1.2, we see that $T \in \mathfrak{so}(V, B)$ if and only if its matrix A relative to this basis satisfies

$$A^t\Gamma + \Gamma A = 0. \tag{1.8}$$

When B is nondegenerate then Γ is invertible, and equation (1.8) can be written as $A^t = -\Gamma A \Gamma^{-1}$. In particular, this implies that $\mathrm{tr}(T) = 0$ for all $T \in \mathfrak{so}(V, B)$.

Orthogonal Lie Algebras

Take $V = \mathbb{F}^n$ and the bilinear form B with matrix $\Gamma = I_n$ relative to the standard basis for $\mathbb{F}^n$. Define

$$\mathfrak{so}(n, \mathbb{F}) = \{X \in M_n(\mathbb{F}) \,:\, X^t = -X\}.$$

Since B is nondegenerate, $\mathfrak{so}(n, \mathbb{F})$ is a Lie subalgebra of $\mathfrak{sl}(n, \mathbb{F})$.

When $\mathbb{F} = \mathbb{R}$ we take integers $p, q \geq 0$ such that $p+q = n$ and B be the bilinear form on $\mathbb{R}^n$ whose matrix relative to the standard basis is $I_{p,q}$ (as in Section 1.1.2). Define

$$\mathfrak{so}(p, q) = \{X \in M_n(\mathbb{R}) \,:\, X^t I_{p,q} = -I_{p,q} X\}.$$

Since B is nondegenerate, $\mathfrak{so}(p, q)$ is a Lie subalgebra of $\mathfrak{sl}(n, \mathbb{R})$.

To obtain a basis-free definition of this family of Lie algebras, let B be a nondegenerate symmetric bilinear form on an n-dimensional vector space V over $\mathbb{F}$. Let $\{v_1, \ldots, v_n\}$ be a basis for V that is orthonormal (when $\mathbb{F} = \mathbb{C}$) or pseudo-orthonormal (when $\mathbb{F} = \mathbb{R}$) relative to B (see Lemma 1.1.2). Let $\mu(T)$ be the matrix of $T \in \mathrm{End}(V)$ relative to this basis . When $\mathbb{F} = \mathbb{C}$, then μ defines a Lie algebra isomorphism of $\mathfrak{so}(V, B)$ onto $\mathfrak{so}(n, \mathbb{C})$. When $\mathbb{F} = \mathbb{R}$ and B has signature (p, q), then μ defines a Lie algebra isomorphism of $\mathfrak{so}(V, B)$ onto $\mathfrak{so}(p, q)$.

Symplectic Lie Algebra

Let J be the $2n \times 2n$ skew-symmetric matrix from Section 1.1.2. We define

$$\mathfrak{sp}(n, \mathbb{F}) = \{X \in M_{2n}(\mathbb{F}) \,:\, X^t J = -JX\}.$$

This subspace of $\mathfrak{gl}(n, \mathbb{F})$ is a Lie subalgebra that we call the *symplectic Lie algebra of rank* n.

To obtain a basis-free definition of this family of Lie algebras, let B be a nondegenerate skew-symmetric bilinear form on a $2n$-dimensional vector space V over $\mathbb{F}$. Let $\{v_1, \ldots, v_{2n}\}$ be a B-symplectic basis for V (see Lemma 1.1.5). The map μ that assigns to an endomorphism of V its matrix relative to this basis defines an isomorphism of $\mathfrak{so}(V, B)$ onto $\mathfrak{sp}(n, \mathbb{F})$.

1.2.3 Unitary Lie Algebras

Let $p, q \geq 0$ be integers such that $p + q = n$ and let $I_{p,q}$ be the $n \times n$ matrix from Section 1.1.2. We define

$$\mathfrak{u}(p, q) = \{X \in M_n(\mathbb{C}) \,:\, X^* I_{p,q} = -I_{p,q} X\}$$

(notice that this space is a *real* subspace of $M_n(\mathbb{C})$). One checks directly that $\mathfrak{u}(p, q)$ is a Lie subalgebra of $\mathfrak{gl}_n(\mathbb{C})$ (considered as a Lie algebra over $\mathbb{R}$). We define $\mathfrak{su}(p, q) = \mathfrak{u}(p, q) \cap \mathfrak{sl}(n, \mathbb{C})$.

To obtain a basis-free description of this family of Lie algebras, let V be an n-dimensional vector space over $\mathbb{C}$, and let B be a nondegenerate Hermitian form on V. We define

$$\mathfrak{u}(V,B) = \{T \in \mathrm{End}_{\mathbb{C}}(V) \,:\, B(Tv,w) = -B(v,Tw) \text{ for all } v,w \in V\}.$$

We set $\mathfrak{su}(V,B) = \mathfrak{u}(V,B) \cap \mathfrak{sl}(V)$. If B has signature (p,q) and if $\{v_1,\ldots,v_n\}$ is a pseudo-orthogonal basis of V relative to B (see Lemma 1.1.7), then the assignment $T \mapsto \mu(T)$ of T to its matrix relative to this basis defines a Lie algebra isomorphism of $\mathfrak{u}(V,B)$ with $\mathfrak{u}(p,q)$ and of $\mathfrak{su}(V,B)$ with $\mathfrak{su}(p,q)$.

1.2.4 Quaternionic Lie Algebras

Quaternionic General and Special Linear Lie Algebras

We follow the notation of Section 1.1.4. Consider the $n \times n$ matrices over the quaternions with the usual matrix commutator. We will denote this Lie algebra by $\mathfrak{gl}(n,\mathbb{H})$, considered as a Lie algebra over $\mathbb{R}$ (we have not defined Lie algebras over skew fields). We can identify $\mathbb{H}^n$ with $\mathbb{C}^{2n}$ by using one of the isomorphic copies of $\mathbb{C}$ ($\mathbb{R}\mathbf{1} + \mathbb{R}\mathbf{i}$, $\mathbb{R}\mathbf{1} + \mathbb{R}\mathbf{j}$, or $\mathbb{R}\mathbf{1} + \mathbb{R}\mathbf{k}$) in $\mathbb{H}$. Define

$$\mathfrak{sl}(n,\mathbb{H}) = \{X \in \mathfrak{gl}(n,\mathbb{H}) \,:\, \mathrm{tr}(X) = 0\}.$$

Then $\mathfrak{sl}(n,\mathbb{H})$ is the real Lie algebra that is usually labeled as $\mathfrak{su}^*(2n)$.

Quaternionic Unitary Lie Algebras

For $n = p + q$ with p, q non-negative integers, we define

$$\mathfrak{sp}(p,q) = \{X \in \mathfrak{gl}(n,\mathbb{H}) \,:\, X^* I_{p,q} = -I_{p,q} X\}$$

(the quaternionic adjoint X^* was defined in Section 1.1.4). We leave it as an exercise to check that $\mathfrak{sp}(p,q)$ is a real Lie subalgebra of $\mathfrak{gl}(n,\mathbb{H})$. Let the quaternionic Hermitian form $B(x,y)$ be defined as in (1.6). Then $\mathfrak{sp}(p,q)$ consists of the matrices $X \in M_n(\mathbb{H})$ that satisfy

$$B(Xx,y) = -B(x,X^*y) \quad \text{for all } x,y \in \mathbb{H}^n.$$

The Lie Algebra $\mathfrak{so}^*(2n)$

Let the automorphism θ of $M_{2n}(\mathbb{C})$ be as defined in Section 1.1.4 ($\theta(A) = -JAJ$). Define

$$\mathfrak{so}^*(2n) = \{X \in \mathfrak{so}(2n,\mathbb{C}) \,:\, \theta(\bar{X}) = X\}.$$

This real vector subspace of $\mathfrak{so}(2n,\mathbb{C})$ is a real Lie subalgebra of $\mathfrak{so}(2n,\mathbb{C})$ (considered as a Lie algebra over $\mathbb{R}$). Identify $\mathbb{C}^{2n}$ with $\mathbb{H}^n$ as Section 1.2.4 and let the quaternionic skew-Hermitian form $C(x,y)$ be defined as in (1.7). Then $\mathfrak{so}^*(2n)$ corresponds to the matrices $X \in M_n(\mathbb{H})$ that satisfy

$$C(Xx,y) = -C(x,X^*y) \quad \text{for all } x,y \in \mathbb{H}^n.$$

1.2.5 Lie Algebras Associated with Classical Groups

The Lie algebras $\mathfrak{g}$ described in the preceding sections constitute the list of *classical Lie algebras* over $\mathbb{R}$ and $\mathbb{C}$. These Lie algebras will be a major subject of study throughout the remainder of this book. We will find, however, that the given matrix form of $\mathfrak{g}$ is not always the most convenient; other choices of bases will be needed to determine the structure of $\mathfrak{g}$. This is one of the reasons that we have stressed the intrinsic basis-free characterizations.

Following the standard convention, we have labeled each classical Lie algebra by a fraktur-font version of the name of a corresponding classical group. This passage from a Lie group to a Lie algebra, which is fundamental to Lie theory, arises by *differentiating* the defining equations for the group. In brief, each classical group G is a subgroup of $\mathrm{GL}(V)$ (where V is a real vector space) that is defined by a set $\mathcal{R}$ of algebraic equations. The corresponding Lie subalgebra $\mathfrak{g}$ of $\mathfrak{gl}(V)$ is determined by taking differentiable curves $\sigma : (-\varepsilon, \varepsilon) \to \mathrm{GL}(V)$ such that $\sigma(0) = I$ and $\sigma(t)$ satisfies the equations in $\mathcal{R}$. Then $\sigma'(0) \in \mathfrak{g}$, and all elements of $\mathfrak{g}$ are obtained in this way. This is the reason why $\mathfrak{g}$ is called the *infinitesimal form* of G.

For example, if G is the subgroup $\mathrm{O}(V, B)$ of $\mathrm{GL}(V)$ defined by a bilinear form B, then the curve σ must satisfy $B(\sigma(t)v, \sigma(t)w) = B(v, w)$ for all $v, w \in V$ and $t \in (-\varepsilon, \varepsilon)$. If we differentiate these relations we have

$$0 = \frac{d}{dt} B(\sigma(t)v, \sigma(t)w)\Big|_{t=0} = B(\sigma'(0)v, \sigma(0)w) + B(\sigma(0)v, \sigma'(0)w)$$

for all $v, w \in V$. Since $\sigma(0) = I$ we see that $\sigma'(0) \in \mathfrak{so}(V, B)$, as asserted.

We will return to these ideas in Section 1.3.4 after developing some basic aspects of Lie group theory.

1.2.6 Exercises

1. Prove that the Jacobi identity (2) holds for $\mathrm{End}(V)$.

2. Prove that the inverse of a bijective Lie algebra homomorphism is a Lie algebra homomorphism.

3. Let B be a bilinear form on a finite-dimensional vector space V over $\mathbb{F}$.

   (a) Prove that $\mathfrak{so}(V, B)$ is a Lie subalgebra of $\mathfrak{gl}(V)$.

   (b) Suppose that B is nondegenerate. Prove that $\mathrm{tr}(X) = 0$ for all $X \in \mathfrak{so}(V, B)$.

4. Prove that $\mathfrak{u}(p, q)$, $\mathfrak{sp}(p, q)$, and $\mathfrak{so}^*(2n)$ are real Lie algebras.

5. Let $B_0(x, y)$ be the Hermitian form and $B_1(x, y)$ the skew-symmetric form on $\mathbb{C}^{2n}$ in Exercises 1.1.5 #11.

   (a) Show that $\mathfrak{sp}(p, q) = \mathfrak{su}(\mathbb{C}^{2n}, B_0) \cap \mathfrak{sp}(\mathbb{C}^{2n}, B_1)$ when $M_n(\mathbb{H})$ is identified with a real subspace of $M_{2n}(\mathbb{C})$ as in Exercises 1.1.5 #10.

   (b) Use part (a) to show that $\mathfrak{sp}(p, q) \subset \mathfrak{sl}(p + q, \mathbb{H})$.

6. Let $X \in M_n(\mathbb{H})$. For each of the three choices of a copy of $\mathbb{C}$ in $\mathbb{H}$ given by (1.4) write out the corresponding matrix of X as an element of $M_{2n}(\mathbb{C})$. Use this formula to show that the trace of X is independent of the choice.

1.3 Closed Subgroups of $\mathrm{GL}(n, \mathbb{R})$

In this section we introduce some Lie-theoretic ideas that motivate the later developments in this book. We begin with the definition of a topological group and then emphasize the topological groups that are closed subgroups of $\mathrm{GL}(n, \mathbb{R})$. Our main tool is the *exponential map*, which we treat by explicit matrix calculations.

1.3.1 Topological Groups

Let G be a group with a Hausdorff topology. If the multiplication and inversion maps

$$G \times G \longrightarrow G \quad (g, h \mapsto gh) \quad \text{and} \quad G \longrightarrow G \quad (g \mapsto g^{-1})$$

are continuous, G is called a *topological group* (in this definition, the set $G \times G$ is given the product topology). For example, $\mathrm{GL}(n, \mathbb{F})$ is a topological group when endowed with the topology of the open subset $\{X : \det(X) \neq 0\}$ of $M_n(\mathbb{F})$. The multiplication is continuous and Cramer's rule implies that the inverse is continuous.

If G is a topological group, each element $g \in G$ defines *translation maps*

$$L_g : G \longrightarrow G \quad \text{and} \quad R_g : G \longrightarrow G,$$

given by $L_g(x) = gx$ and $R_g(x) = xg$. The group properties and continuity imply that R_g and L_g are homeomorphisms.

If G is a topological group and H is a subgroup that is closed as a subspace of G, then H is also a topological group (in the relative topology). We call H a *topological subgroup* of G. For example, the defining equations of each classical group show that it is a closed subset of $\mathrm{GL}(V)$ for some V, and hence it is a topological subgroup of $\mathrm{GL}(V)$.

A *topological group homomorphism* will mean a continuous topological group homomorphism. A topological group homomorphism is said to be a *topological group isomorphism* if it is bijective and its inverse is also a topological group homomorphism. An isomorphism of a group with itself is called an *automorphism*. For example, if G is a topological group then each element $g \in G$ defines an automorphism $\tau(g)$ by conjugation: $\tau(g)x = gxg^{-1}$. Such automorphisms are called *inner*.

Before we study our main examples we prove two useful general results about topological groups.

Proposition 1.3.1. *If H is an open subgroup of a topological group G, then H is closed in G.*

Proof. We note that G is a disjoint union of left cosets. If $g \in G$ then the left coset $gH = L_g(H)$ is open since L_g is a homeomorphism. Hence the union of all the left cosets other than H is open, and so H is closed. ⧫

Proposition 1.3.2. *Let G be a topological group. Then the identity component of G (that is, the connected component that contains the identity element e) is a normal subgroup.*

Proof. Let H be the identity component of G. If $h \in H$ then $h \in L_h(H)$ because $e \in H$. Since L_h is a homeomorphism and $H \cap L_h(H)$ is nonempty, it follows that $L_h(H) = H$, showing that H is closed under multiplication. Since $e \in L_h H$ we also have $h^{-1} \in H$, and so H is a subgroup. If $g \in G$ the inner automorphism $\tau(g)$ is a homeomorphism that fixes e and hence maps H into H. ⧫

1.3.2 Exponential Map

On $M_n(\mathbb{R})$ we define the inner product $\langle X, Y\rangle = \mathrm{tr}(XY^t)$. The corresponding norm

$$\|X\| = \langle X, X\rangle^{\frac{1}{2}} = \Big(\sum_{i,j=1}^{n} x_{ij}^2\Big)^{\frac{1}{2}}$$

has the following properties (where $X, Y \in M_n(\mathbb{R})$ and $c \in \mathbb{R}$):

(1) $\|X+Y\| \le \|X\| + \|Y\|, \quad \|cX\| = |c|\,\|X\|,$

(2) $\|XY\| \le \|X\|\,\|Y\|,$

(3) $\|X\| = 0$ if and only if $X = 0$.

Properties (1) and (3) follow by identifying $M_n(\mathbb{R})$ as a real vector space with $\mathbb{R}^{n^2}$ using the matrix entries. To verify property (2), observe that

$$\|XY\|^2 = \sum_{i,j} \Big(\sum_k x_{ik} y_{kj}\Big)^2.$$

Now $\big|\sum_k x_{ik}y_{kj}\big|^2 \le \big(\sum_k x_{ik}^2\big)\big(\sum_k y_{kj}^2\big)$ by the usual Cauchy-Schwarz inequality. Hence

$$\|XY\|^2 \le \sum_{i,j} \Big(\sum_k x_{ik}^2\Big)\Big(\sum_k y_{kj}^2\Big) = \Big(\sum_{i,k} x_{ik}^2\Big)\Big(\sum_{k,j} y_{kj}^2\Big) = \|X\|^2\,\|Y\|^2.$$

Taking the square root of both sides completes the proof.

We define matrix-valued analytic functions by substitution in convergent power series. Let $\{a_m\}$ be a sequence of real numbers such that

$$\sum_{m=0}^{\infty} |a_m| r^m < \infty \quad \text{for some } r > 0.$$

For $A \in M_n(\mathbb{R})$ and $r > 0$ let

$$B_r(A) = \{X \in M_n(\mathbb{R}) \,:\, \|X - A\| < r\}$$

(the open ball of radius r around A). If $X \in B_r(0)$ and $k \geq l$ then by properties (1) and (2) of the norm we have

$$\begin{aligned}\Big\| \sum_{0\leq m\leq k} a_m X^m - \sum_{0\leq m\leq l} a_m X^m \Big\| = \Big\| \sum_{l< m\leq k} a_m X^m \Big\| &\leq \sum_{l< m\leq k} |a_m| \, \|X^m\| \\ &\leq \sum_{l< m\leq k} |a_m| \, \|X\|^m \leq \sum_{m>l}^{\infty} |a_m| r^m.\end{aligned}$$

The last series goes to 0 as $l \to \infty$ by the convergence assumption. Thus we can define the function

$$f(X) = \sum_{m=0}^{\infty} a_m X^m$$

on $B_r(0)$. The function $X \mapsto f(X)$ is *real analytic* (each entry in the matrix $f(X)$ is a convergent power series in the entries of X when $\|X\| < r$).

Substitution Principle: Any equation involving power series in a variable x that holds as an identity of absolutely convergent scalar series when $|x| < r$, also holds as an identity of matrix series that converge absolutely in the matrix norm when $||X|| < r$.

This follows by rearranging the power series, which is permissible by absolute convergence.

In Lie theory two functions play a special role:

$$\exp(X) = \sum_{m=0}^{\infty} \frac{1}{m!} X^m \quad \text{and} \quad \log(1+X) = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{1}{m} X^m.$$

The exponential series converges absolutely for all X, and the logarithm series converges absolutely for $\|X\| < 1$. We therefore have two analytic matrix-valued functions

$$\exp : M_n(\mathbb{R}) \longrightarrow M_n(\mathbb{R}) \quad \text{and} \quad \log : B_1(I) \longrightarrow M_n(\mathbb{R}).$$

If $X, Y \in M_n(\mathbb{R})$ and $XY = YX$, then each term $(X+Y)^m$ can be expanded by the binomial formula. Rearranging the series for $\exp(X+Y)$ (which is justified by absolute convergence), we obtain the identity

$$\exp(X+Y) = \exp(X)\exp(Y). \tag{1.9}$$

In particular, this implies that $\exp(X)\exp(-X) = \exp(0) = I$. Thus

$$\exp : M_n(\mathbb{R}) \longrightarrow \mathrm{GL}(n, \mathbb{R}).$$

The power series for the exponential and logarithm satisfy the identities

$$\log(\exp(x)) = x \qquad \text{for } |x| < \log 2, \tag{1.10}$$

$$\exp(\log(1+x)) = 1+x \quad \text{for } |x| < 1. \tag{1.11}$$

To verify (1.10), use the chain rule to show that the derivative of $\log(\exp(x))$ is 1; since this function vanishes at $x = 0$, it is x. To verify (1.11), use the chain rule twice to show that the second derivative of $\exp(\log(1+x))$ is zero; thus the function is a polynomial of degree one. This polynomial and its first derivative have the value 1 at $x = 0$, hence it is $x + 1$.

We use these identities to show that the matrix logarithm function gives a local inverse to the exponential function.

Lemma 1.3.3. *Suppose $g \in \mathrm{GL}(n,\mathbb{R})$ satisfies $\|g - I\| < \log 2/(1+\log 2)$. Then $\|\log(g)\| < \log 2$ and $\exp(\log(g)) = g$. Furthermore, if $X \in B_{\log 2}(0)$ and $\exp X = g$, then $X = \log(g)$.*

Proof. Since $\log 2/(1+\log 2) < 1$, the power series for $\log(g)$ is absolutely convergent and

$$\|\log(g)\| \le \sum_{m=1}^{\infty} \|g-I\|^m = \frac{\|g-I\|}{1-\|g-I\|} < \log 2.$$

Since $\|g - I\| < 1$, we can replace z by $g - I$ in identity (1.11) by the substitution principle. Hence $\exp(\log(g)) = g$.

If $X \in B_{\log 2}(0)$ then

$$\|\exp(X) - I\| \le \mathrm{e}^{\|X\|} - 1 < 1.$$

Hence we can replace x by X in identity (1.10) by the substitution principle. If $\exp X = g$, this identity yields $X = \log(g)$. ♦

Remark 1.3.4. Lemma 1.3.3 asserts that the exponential map is a bijection from a neighborhood of 0 in $M_n(\mathbb{R})$ onto a neighborhood of I in $\mathrm{GL}(n,\mathbb{R})$. However, if $n > 1$ then the map

$$\exp : M_n(\mathbb{R}) \longrightarrow \{g \,:\, \det(g) > 0\} \subset \mathrm{GL}(n,\mathbb{R})$$

is neither injective nor surjective, unlike the scalar case.

If $X \in M_n(\mathbb{R})$, then the continuous function $\varphi(t) = \exp(tX)$ from $\mathbb{R}$ to $\mathrm{GL}(n,\mathbb{R})$ satisfies $\varphi(0) = I$ and $\varphi(s+t) = \varphi(s)\varphi(t)$ for all $s, t \in \mathbb{R}$, by equation (1.9). Thus given X, we obtain a homomorphism φ from the additive group of real numbers to the group $\mathrm{GL}(n,\mathbb{R})$. We call this homomorphism the *one-parameter group generated by X*. It is a fundamental result in Lie theory that all homomorphisms from $\mathbb{R}$ to $\mathrm{GL}(n,\mathbb{R})$ are obtained this way.

Theorem 1.3.5. *Let $\varphi : \mathbb{R} \longrightarrow \mathrm{GL}(n,\mathbb{R})$ be a continuous homomorphism from the additive group $\mathbb{R}$ to $\mathrm{GL}(n,\mathbb{R})$. Then there exists a unique $X \in M_n(\mathbb{R})$ such that $\varphi(t) = \exp(tX)$ for all $t \in \mathbb{R}$.*

Proof. The uniqueness of X is immediate, since

$$\frac{d}{dt}\exp(tX)\Big|_{t=0} = X.$$

To prove the existence of X, let $\varepsilon > 0$ and set $\varphi_\varepsilon(t) = \varphi(\varepsilon t)$. Then φ_ε is also a continuous homomorphism of $\mathbb{R}$ into $\mathrm{GL}(n,\mathbb{R})$. Since φ is continuous and $\varphi(0) = I$, from Lemma 1.3.3 we can choose ε so that $\varphi_\varepsilon(t) \in \exp B_r(0)$ for $|t| < 2$, where $r = \frac{1}{2}\log 2$. If we can show that $\varphi_\varepsilon(t) = \exp(tX)$ for some $X \in M_n(\mathbb{R})$ and all $t \in \mathbb{R}$, then $\varphi(t) = \exp\big((t/\varepsilon)X\big)$. Thus it suffices to treat the case $\varepsilon = 1$.

Assume now that $\varphi(t) \in \exp B_r(0)$ for $|t| < 2$, with $r = \frac{1}{2}\log 2$. Then there exists $X \in B_r(0)$ such that $\varphi(1) = \exp X$. Likewise, there exists $Z \in B_r(0)$ such that $\varphi\big(\frac{1}{2}\big) = \exp Z$. But

$$\varphi(1) = \varphi\big(\tfrac{1}{2}\big)\cdot\varphi\big(\tfrac{1}{2}\big) = \exp(Z)\cdot\exp(Z) = \exp(2Z).$$

Since $\|2Z\| < \log 2$ and $\|X\| < \log 2$, Lemma 1.3.3 implies that $Z = \frac{1}{2}X$. Since $\varphi(\frac{1}{4}) = \exp(W)$ with $W \in B_r(0)$, we likewise have $W = \frac{1}{2}Z = \frac{1}{4}X$. Continuing this argument, we conclude that

$$\varphi\big(\tfrac{1}{2^k}\big) = \exp\big(\tfrac{1}{2^k}X\big) \quad \text{for all integers } k \geq 0.$$

Let $a = \frac{1}{2}a_1 + \frac{1}{4}a_2 + \cdots + \frac{1}{2^k}a_k + \cdots$, with $a_j \in \{0,1\}$, be the dyadic expansion of the real number $0 \leq a < 1$. Then by continuity and the assumption that φ is a group homomorphism we have

$$\begin{aligned}\varphi(a) &= \lim_{k\to\infty} \varphi\big(\tfrac{1}{2}a_1 + \tfrac{1}{4}a_2 + \cdots + \tfrac{1}{2^k}a_k\big) = \lim_{k\to\infty} \varphi\big(\tfrac{1}{2}\big)^{a_1}\varphi\big(\tfrac{1}{4}\big)^{a_2}\cdots\varphi\big(\tfrac{1}{2^k}\big)^{a_k}\\ &= \lim_{k\to\infty} \big(\exp\tfrac{1}{2}X\big)^{a_1}\cdots\big(\exp\tfrac{1}{2^k}X\big)^{a_k}\\ &= \lim_{k\to\infty} \exp\big\{\big(\tfrac{1}{2}a_1 + \tfrac{1}{4}a_2 + \cdots + \tfrac{1}{2^k}a_k\big)X\big\} = \exp(aX).\end{aligned}$$

Now if $0 \leq a < 1$ then $\varphi(-a) = \varphi(a)^{-1} = \exp(aX)^{-1} = \exp(-aX)$. Finally, given $a \in \mathbb{R}$ choose an integer $k > |a|$. Then

$$\varphi(a) = \varphi\big(\tfrac{a}{k}\big)^k = \big(\exp\tfrac{a}{k}X\big)^k = \exp(aX).$$

This shows that φ is the one-parameter subgroup generated by X. ♦

1.3.3 Lie Algebra of a Closed Subgroup of $\mathrm{GL}(n,\mathbb{R})$

Let G be a closed subgroup of $\mathrm{GL}(n,\mathbb{R})$. We define

$$\mathrm{Lie}(G) = \{X \in M_n(\mathbb{R}) \,:\, \exp(tX) \in G \text{ for all } t \in \mathbb{R}\} \tag{1.12}$$

Thus by Theorem 1.3.5 each matrix in $\mathrm{Lie}(G)$ corresponds to a unique continuous one-parameter subgroup of G. To show that $\mathrm{Lie}(G)$ is a Lie subalgebra of $\mathfrak{gl}(n,\mathbb{R})$, we need more information about the product $\exp X \exp Y$.

Fix $X, Y \in M_n(\mathbb{R})$. By Lemma 1.3.3 there is an analytic matrix-valued function $Z(s,t)$, defined for (s,t) in a neighborhood of zero in $\mathbb{R}^2$, such that $Z(0,0) = 0$ and

$$\exp(sX)\exp(tY) = \exp(Z(s,t)).$$

It is easy to calculate the linear and quadratic terms in the power series of $Z(s,t)$. Since $Z(s,t) = \log(\exp(sX)\exp(tY))$, the power series for the logarithm and exponential functions give

$$\begin{aligned} Z(s,t) &= \big(\exp(sX)\exp(tY) - I\big) - \tfrac{1}{2}\big(\exp(sX)\exp(tY) - I\big)^2 + \cdots \\ &= \big((I + sX + \tfrac{1}{2}s^2X^2)(I + tY + \tfrac{1}{2}t^2Y^2) - I\big) - \tfrac{1}{2}\big(sX + tY\big)^2 + \cdots \\ &= \big(sX + tY + \tfrac{1}{2}s^2X^2 + stXY + \tfrac{1}{2}t^2Y^2\big) \\ &\qquad - \tfrac{1}{2}\big(s^2X^2 + st(XY + YX) + t^2Y^2\big) + \cdots, \end{aligned}$$

where $\cdots$ indicates terms that are of total degree three and higher in s, t. The first-degree term is $sX+tY$, as expected (the series terminates after this term when X and Y commute). The quadratic terms involving only X or Y cancel; the only remaining term involving both X and Y is the commutator:

$$Z(s,t) = sX + tY + \frac{st}{2}[X,Y] + \cdots. \tag{1.13}$$

Rescaling X and Y, we can state formula (1.13) as follows:

Lemma 1.3.6. *There exists $\varepsilon > 0$ and an analytic matrix-valued function $R(X,Y)$ on $B_\varepsilon(0) \times B_\varepsilon(0)$ such that*

$$\exp X \exp Y = \exp\big(X + Y + \tfrac{1}{2}[X,Y] + R(X,Y)\big)$$

when $X, Y \in B_\varepsilon(0)$. Furthermore, $\|R(X,Y)\| \le C(\|X\| + \|Y\|)^3$ for some constant C and all $X, Y \in B_\varepsilon(0)$.

From Lemma 1.3.6 we now obtain the fundamental identities relating the Lie algebra structure of $\mathfrak{gl}(n,\mathbb{R})$ to the group structure of $\mathrm{GL}(n,\mathbb{R})$.

Proposition 1.3.7. *For $X, Y \in M_n(\mathbb{R})$ one has*

$$\exp(X+Y) = \lim_{k\to\infty}\Big(\exp\big(\tfrac{1}{k}X\big)\exp\big(\tfrac{1}{k}Y\big)\Big)^k \tag{1.14}$$

$$\exp([X,Y]) = \lim_{k\to\infty}\Big(\exp\big(\tfrac{1}{k}X\big)\exp\big(\tfrac{1}{k}Y\big)\exp\big(-\tfrac{1}{k}X\big)\exp\big(-\tfrac{1}{k}Y\big)\Big)^{k^2} \tag{1.15}$$

Proof. For k a sufficiently large integer Lemma 1.3.6 implies that

$$\exp\big(\tfrac{1}{k}X\big)\exp\big(\tfrac{1}{k}Y\big) = \exp\big(\tfrac{1}{k}(X+Y) + \mathrm{O}(1/k^2)\big),$$

where $\mathrm{O}(r)$ denotes a matrix function of r whose norm is bounded by Cr for some constant C (depending only on $||X|| + ||Y||$) and all small r. Hence

$$\begin{aligned}\Big(\exp\big(\tfrac{1}{k}X\big)\exp\big(\tfrac{1}{k}Y\big)\Big)^k &= \exp k\Big(\tfrac{1}{k}\big(X+Y\big)+\mathrm{O}(1/k^2)\Big)\\ &= \exp\Big(X+Y+\mathrm{O}\big(1/k\big)\Big).\end{aligned}$$

Letting $k \to \infty$, we obtain formula (1.14).

Likewise, we have

$$\begin{aligned}&\exp\big(\tfrac{1}{k}X\big)\exp\big(\tfrac{1}{k}Y\big)\exp\big(-\tfrac{1}{k}X\big)\exp\big(-\tfrac{1}{k}Y\big)\\ &\qquad= \exp\big(\tfrac{1}{k}(X+Y)+\tfrac{1}{2k^2}[X,Y]+\mathrm{O}(1/k^3)\big)\\ &\qquad\qquad\cdot\exp\big(-\tfrac{1}{k}(X+Y)+\tfrac{1}{2k^2}[X,Y]+\mathrm{O}(1/k^3)\big)\\ &\qquad= \exp\big(\tfrac{1}{k^2}[X,Y]+\mathrm{O}(1/k^3)\big).\end{aligned}$$

(Of course, each occurrence of $\mathrm{O}\big(1/k^3\big)$ in these formulas stands for a different function.) Thus

$$\begin{aligned}&\Big(\exp\big(\tfrac{1}{k}X\big)\exp\big(\tfrac{1}{k}Y\big)\exp\big(-\tfrac{1}{k}X\big)\exp\big(-\tfrac{1}{k}Y\big)\Big)^{k^2}\\ &\qquad= \exp k^2\Big(\tfrac{1}{k^2}[X,Y]+\mathrm{O}\big(1/k^3\big)\Big) = \exp\big([X,Y]+\mathrm{O}(1/k)\big).\end{aligned}$$

This implies formula (1.15). ⧫

Theorem 1.3.8. *If G is a closed subgroup of $\mathrm{GL}(n,\mathbb{R})$ then $\mathrm{Lie}(G)$ is a Lie subalgebra of $M_n(\mathbb{R})$.*

Proof. If $X \in \mathrm{Lie}(G)$ then $tX \in \mathrm{Lie}(G)$ for all $t \in \mathbb{R}$. If $X, Y \in \mathrm{Lie}(G)$ and $t \in \mathbb{R}$, then

$$\exp\big(t(X+Y)\big) = \lim_{k\to\infty}\Big(\exp\big(\tfrac{t}{k}X\big)\exp\big(\tfrac{t}{k}Y\big)\Big)^k$$

is in G since G is a closed subgroup. Similarly,

$$\exp(t[X,Y]) = \lim_{k\to\infty}\Big(\exp(\tfrac{t}{k}X)\exp\big(\tfrac{t}{k}Y\big)\exp\big(-\tfrac{t}{k}X\big)\exp\big(-\tfrac{t}{k}Y\big)\Big)^{k^2}$$

is in G. ⧫

If G is a closed subgroup of $\mathrm{GL}(n,\mathbb{R})$, then the elements of G act on the one-parameter subgroups in G by conjugation. Since $gX^kg^{-1} = (gXg^{-1})^k$ for $g \in G$, $X \in \mathrm{Lie}(G)$, and all positive integers k, we have

$$g(\exp tX)g^{-1} = \exp(tgXg^{-1}) \quad \text{for all } t \in \mathbb{R}$$

The left side of this equation is a one-parameter subgroup of G. Hence $gXg^{-1} \in \mathrm{Lie}(G)$. We define $\mathrm{Ad}(g) \in \mathrm{GL}(\mathrm{Lie}(G))$ by

$$\mathrm{Ad}(g)X = gXg^{-1} \quad \text{for } X \in \mathrm{Lie}(G). \tag{1.16}$$

Clearly $\mathrm{Ad}(g_1g_2) = \mathrm{Ad}(g_1)\,\mathrm{Ad}(g_2)$ for $g_1, g_2 \in G$, so $g \mapsto \mathrm{Ad}(g)$ is a continuous group homomorphism from G to $\mathrm{GL}(\mathrm{Lie}(G))$. Furthermore, if $X, Y \in \mathrm{Lie}(G)$ and $g \in G$, then the relation $gXYg^{-1} = (gXg^{-1})(gYg^{-1})$ implies that

$$\mathrm{Ad}(g)([X,Y]) = [\mathrm{Ad}(g)X, \mathrm{Ad}(g)Y] \tag{1.17}$$

Hence $\mathrm{Ad}(g)$ is an *automorphism* of the Lie algebra structure.

Remark 1.3.9. There are several ways of associating a Lie algebra with a closed subgroup of $\mathrm{GL}(n,\mathbb{R})$; in the course of chapter we shall prove that the different Lie algebras are all isomorphic.

1.3.4 Lie Algebras of the Classical Groups

To determine the Lie algebras of the classical groups, we fix the following embeddings of $\mathrm{GL}(n,\mathbb{F})$ as a closed subgroup of $\mathrm{GL}(dn,\mathbb{R})$. Here $\mathbb{F}$ is $\mathbb{C}$ or $\mathbb{H}$ and $d = \dim_{\mathbb{R}} \mathbb{F}$.

We take $\mathbb{C}^n$ to be $\mathbb{R}^{2n}$ and let multiplication by $\sqrt{-1}$ be given by the matrix

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix},$$

with I the $n \times n$ identity matrix. Then $M_n(\mathbb{C})$ is identified with the matrices in $M_{2n}(\mathbb{R})$ that commute with J, and $\mathrm{GL}(n,\mathbb{C})$ is identified with the invertible matrices in $M_n(\mathbb{C})$. Thus $\mathrm{GL}(n,\mathbb{C})$ is a closed subgroup of $\mathrm{GL}(2n,\mathbb{R})$.

The case of the quaternionic groups is handled similarly. We take $\mathbb{H}^n$ to be $\mathbb{R}^{4n}$ and use the $4n \times 4n$ matrices

$$J_1 = \begin{bmatrix} J & 0 \\ 0 & -J \end{bmatrix}, \quad J_2 = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}, \quad \text{and} \quad J_3 = \begin{bmatrix} 0 & J \\ J & 0 \end{bmatrix}$$

(with J as above but now I is the $2n \times 2n$ identity matrix) to give multiplication by **i**, **j**, and **k**, respectively. This gives a model for $\mathbb{H}^n$, since these matrices satisfy the quaternion relations

$$\begin{aligned} &J_p^2 = -I, \quad J_1J_2 = J_3, \quad J_2J_3 = J_1, \quad J_3J_1 = J_2 \\ &\text{and} \quad J_pJ_l = -J_lJ_p \quad \text{for } p \neq l. \end{aligned}$$

In this model $M_n(\mathbb{H})$ is identified with the matrices in $M_{4n}(\mathbb{R})$ that commute with J_p ($p = 1, 2, 3$), and $\mathrm{GL}(n,\mathbb{H})$ consists of the invertible matrices in $M_n(\mathbb{H})$. Thus $\mathrm{GL}(n,\mathbb{H})$ is a closed subgroup of $\mathrm{GL}(4n,\mathbb{R})$.

Since each classical group G is a closed subgroup of $\mathrm{GL}(n,\mathbb{F})$ with $\mathbb{F}$ either $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$, the embeddings just defined make G a closed subgroup of $\mathrm{GL}(dn,\mathbb{R})$. With these identifications the names of the Lie algebras in Section 1.2 are consistent with the names attached to the groups in Section 1.1; to obtain the Lie algebra corresponding to a classical group, one replaces the initial capital letters in the group name with

fraktur letters. We work out the details for a few examples and leave the rest as an exercise.

It is clear from the definition that $\mathrm{Lie}(\mathrm{GL}(n,\mathbb{R})) = M_n(\mathbb{R}) = \mathfrak{gl}(n,\mathbb{R})$. The Lie algebra of $\mathrm{GL}(n,\mathbb{C})$ consists of all $X \in M_{2n}(\mathbb{R})$ such that $J^{-1}\exp(tX)J = \exp(tX)$ for all $t \in \mathbb{R}$. Since $A^{-1}\exp(X)A = \exp(A^{-1}XA)$ for any $A \in \mathrm{GL}(n,\mathbb{R})$, we see that $X \in \mathrm{Lie}(\mathrm{GL}(n,\mathbb{C}))$ if and only if $\exp(tJ^{-1}XJ) = \exp(tX)$ for all $t \in \mathbb{R}$. This relation holds if and only if $J^{-1}XJ = X$, so we conclude that $\mathrm{Lie}(\mathrm{GL}(n,\mathbb{C})) = \mathfrak{gl}(n,\mathbb{C})$. The same argument (using the matrices $\{J_i\}$) shows that $\mathrm{Lie}(\mathrm{GL}(n,\mathbb{H})) = \mathfrak{gl}(n,\mathbb{H})$.

We now look at $\mathrm{SL}(n,\mathbb{R})$. For any $X \in M_n(\mathbb{C})$ there exists $U \in \mathrm{U}(n)$ and an upper-triangular matrix $T = [t_{ij}]$ such that $X = UTU^{-1}$ (this is the *Schur triangular form*). Thus $\exp(X) = U\exp(T)U^{-1}$ and so $\det(\exp(X)) = \det(\exp(T))$. Since $\exp(T)$ is upper triangular with ith diagonal entry $\mathrm{e}^{t_{ii}}$, we have $\det(\exp(T)) = \mathrm{e}^{\mathrm{tr}(T)}$. But $\mathrm{tr}(T) = \mathrm{tr}(X)$, so we conclude that

$$\det(\exp(X)) = \mathrm{e}^{\mathrm{tr}(X)}. \tag{1.18}$$

If $X \in M_n(\mathbb{R})$, then from equation (1.18) we see that the one-parameter subgroup $t \mapsto \exp(tX)$ is in $\mathrm{SL}(n,\mathbb{R})$ if and only if $\mathrm{tr}(X) = 0$. Hence

$$\mathrm{Lie}(\mathrm{SL}(n,\mathbb{R})) = \{X \in M_n(\mathbb{R}) \,:\, \mathrm{tr}(X) = 0\} = \mathfrak{sl}(n,\mathbb{R}).$$

For the other classical groups it is convenient to use the following simple result.

Lemma 1.3.10. *Suppose $H \subset G \subset \mathrm{GL}(n,\mathbb{R})$ with H a closed subgroup of G and G a closed subgroup of $\mathrm{GL}(n,\mathbb{R})$. Then H is a closed subgroup of $\mathrm{GL}(n,\mathbb{R})$, and*

$$\mathrm{Lie}(H) = \{X \in \mathrm{Lie}(G) \,:\, \exp(tX) \in H \text{ for all } t \in \mathbb{R}\}.$$

Proof. It is obvious that H is a closed subgroup of $\mathrm{GL}(n,\mathbb{R})$. If $X \in \mathrm{Lie}(H)$ then $\exp(tX) \in H \subset G$ for all $t \in \mathbb{R}$. Thus $X \in \mathrm{Lie}(G)$. ♦

We consider $\mathrm{Lie}(\mathrm{Sp}(n,\mathbb{C}))$. Since $\mathrm{Sp}(n,\mathbb{C}) \subset \mathrm{GL}(2n,\mathbb{C}) \subset \mathrm{GL}(2n,\mathbb{R})$, we can look upon $\mathrm{Lie}(\mathrm{Sp}(n,\mathbb{C}))$ as the set of $X \in M_n(\mathbb{C})$ such that $\exp tX \in \mathrm{Sp}(n,\mathbb{C})$ for all $t \in \mathbb{R}$. This condition can be expressed as

$$\exp(tX^t)J\exp(tX) = J \quad \text{for all } t \in \mathbb{R}. \tag{1.19}$$

Differentiating this equation at $t = 0$, we find that $X^tJ + JX = 0$ for all $X \in \mathrm{Lie}(\mathrm{Sp}(n,\mathbb{C}))$. Conversely, if X satisfies this last equation, then $JXJ^{-1} = -X^t$, and so

$$J\exp(tX)J^{-1} = \exp(tJXJ^{-1}) = \exp(-tX^t) \quad \text{for all } t \in \mathbb{R}.$$

Hence X satisfies condition (1.19). This proves that $\mathrm{Lie}(\mathrm{Sp}(n,\mathbb{C})) = \mathfrak{sp}(n,\mathbb{C})$.

We do one more family of examples. Let $G = \mathrm{U}(p,q) \subset \mathrm{GL}(p+q,\mathbb{C})$. Then

$$\mathrm{Lie}(G) = \{X \in M_n(\mathbb{C}) \,:\, \exp(tX)^* I_{p,q} \exp(tX) = I_{p,q} \text{ for all } t \in \mathbb{R}\}.$$

We note that for $t \in \mathbb{R}$

$$\big(\exp tX\big)^* = \big(I + tX + \tfrac{1}{2}t^2X^2 + \cdots\big)^* = I + tX^* + \tfrac{1}{2}t^2(X^*)^2 + \cdots .$$

Thus if $X \in \mathrm{Lie}(G)$, then

$$\big(\exp tX\big)^* I_{p,q} \exp tX = I_{p,q}.$$

Differentiating this equation with respect to t at $t = 0$, we obtain $X^*I_{p,q} + I_{p,q}X = 0$. This shows that $\mathrm{Lie}(\mathrm{U}(p,q)) \subset \mathfrak{u}(p,q)$. Conversely, if $X \in \mathfrak{u}(p,q)$, then

$$(X^*)^k I_{p,q} = (-1)^k I_{p,q} X^k \quad \text{for all integers } k.$$

Using this relation in the power series for the exponential function, we have

$$\exp(tX^*)I_{p,q} = I_{p,q} \exp(-tX).$$

This equation can be written as $\exp(tX)^* I_{p.q} \exp(tX) = I_{p.q}$; hence $\exp(tX) \in \mathrm{U}(p,q)$ for all $t \in \mathbb{R}$. This proves that $\mathrm{Lie}(\mathrm{U}(p,q)) = \mathfrak{u}(p,q)$.

1.3.5 Exponential Coordinates on Closed Subgroups

We will now study in more detail the relationship between the Lie algebra of a closed subgroup H of $\mathrm{GL}(n,\mathbb{R})$ and the group structure of H. We first note that for $X \in \mathrm{Lie}(H)$ the map $t \mapsto \exp tX$ from $\mathbb{R}$ to H has range in the identity component of H. Hence the Lie algebra of H is the same as the Lie algebra of the identity component of H. It is therefore reasonable to confine our attention to connected groups in this discussion. The following key result is due to J. von Neumann:

Theorem 1.3.11. *Let H be a closed subgroup of $\mathrm{GL}(n,\mathbb{R})$. There exists an open neighborhood V of 0 in $\mathrm{Lie}(H)$ and an open neighborhood Ω of I in $\mathrm{GL}(n,\mathbb{R})$ so that*

(1) $\exp(V) = H \cap \Omega$

(2) $\exp : V \longrightarrow \exp(V)$ *is a homeomorphism onto the open neighborhood $H \cap \Omega$ of I in H*

Proof. Let $K = \{X \in M_n(\mathbb{R}) \,:\, \mathrm{tr}(X^tY) = 0 \text{ for all } Y \in \mathrm{Lie}(H)\}$ be the orthogonal complement of $\mathrm{Lie}(H)$ in $M_n(\mathbb{R})$ relative to the trace form inner product. Then there is an orthogonal direct sum decomposition

$$M_n(\mathbb{R}) = \mathrm{Lie}(H) \oplus K. \tag{1.20}$$

Using decomposition (1.20), we define an analytic map $\varphi : M_n(\mathbb{R}) \longrightarrow \mathrm{GL}(n,\mathbb{R})$ by $\varphi(X) = \exp(X_1)\exp(X_2)$ when $X = X_1 + X_2$ with $X_1 \in \mathrm{Lie}(H)$ and $X_2 \in K$. We note that $\varphi(0) = I$ and

$$\varphi(tX) = \big(I + tX_1 + \mathrm{O}(t^2)\big)\big(I + tX_2 + \mathrm{O}(t^2)\big) = I + tX + \mathrm{O}(t^2).$$

Hence the differential of φ at 0 is the identity map. The inverse function theorem implies that there exists $s_1 > 0$ such that $\varphi : B_{s_1}(0) \longrightarrow \mathrm{GL}(n, \mathbb{R})$ has an open image U_1 and the map $\varphi : B_{s_1}(0) \longrightarrow U_1$ is a homeomorphism.

With these preliminaries established we can begin the argument. Suppose, for the sake of obtaining a contradiction, that for every $\varepsilon > 0$ with $\varepsilon \leq s_1$ the set $\varphi(B_\varepsilon(0)) \cap H$ contains an element that is not in $\exp(\mathrm{Lie}(H))$. In this case for each integer $k \geq 1/s_1$ there exists an element in $Z_k \in B_{1/k}(0)$ such that $\exp(Z_k) \in H$ and $Z_k \notin \mathrm{Lie}(H)$. We write $Z_k = X_k + Y_k$ with $X_k \in \mathrm{Lie}(H)$ and $0 \neq Y_k \in K$. Then

$$\varphi(Z_k) = \exp(X_k)\exp(Y_k).$$

Since $\exp(X_k) \in H$, we see that $\exp(Y_k) \in H$. We also observe that $\|Y_k\| \leq 1/k$. Let $\varepsilon_k = \|Y_k\|$. Then $0 < \varepsilon_k \leq 1/k \leq s_1$. For each k there exists a positive integer m_k such that $s_1 \leq m_k\varepsilon_k < 2s_1$. Hence

$$s_1 \leq \|m_k Y_k\| < 2s_1\,. \tag{1.21}$$

Since the sequence $m_k Y_k$ is bounded we can replace it with a subsequence that converges. We may therefore assume that there exists $Y \in W$ with $\lim_{k\to\infty} m_k Y_k = Y$. Then $\|Y\| \geq s_1 > 0$ by inequalities (1.21), so $Y \neq 0$.

We claim that $\exp(tY) \in H$ for all $t \in \mathbb{R}$. Indeed, given t, we write $tm_k = a_k + b_k$ with $a_k \in \mathbb{Z}$ the integer part of tm_k and $0 \leq b_k < 1$. Then

$$tm_k Y_k = a_k Y_k + b_k Y_k.$$

Hence

$$\exp(tm_k Y_k) = \big(\exp Y_k\big)^{a_k} \exp\big(b_k Y_k\big).$$

We have $\big(\exp Y_k\big)^{m_k} \in H$ for all n. Since $\lim_{k\to\infty} Y_k = 0$ and $0 \leq b_k < 1$, it follows that $\lim_{k\to\infty} \exp(b_k Y_k) = I$. Hence

$$\exp(tY) = \lim_{k\to\infty} \exp\big(tm_k Y_k\big) = \lim_{k\to\infty} \big(\exp Y_k\big)^{a_k} \in H$$

since H is closed. But this implies that $Y \in \mathrm{Lie}(H) \cap K = \{0\}$, which is a contradiction since $Y \neq 0$. This proves that there must exist an $\varepsilon > 0$ such that $\varphi(B_\varepsilon(0)) \cap H \subset \exp(\mathrm{Lie}(H))$. Set $V = B_\epsilon(0) \cap \mathrm{Lie}\, H$. Then

$$\exp V = \varphi(B_\epsilon(0)) \cap H$$

is an open neighborhood of I in H, by definition of the relative topology on H, and the restriction of $\exp$ to V is a homeomorphism onto $\exp V$. ♦

A topological group G is a *Lie group* if there is a differentiable manifold structure on G (see Appendix D.1.1) such that the following conditions are satisfied:

(i) The manifold topology on G is the same as the group topology.

(ii) The multiplication map $G \times G \longrightarrow G$ and the inversion map $G \longrightarrow G$ are C^∞.

The group $\mathrm{GL}(n,\mathbb{R})$ is a Lie group, with its manifold structure as an open subset of the vector space $M_n(\mathbb{R})$. The multiplication and inversion maps are real analytic.

Theorem 1.3.12. *Let H be a closed subgroup of $\mathrm{GL}(n,\mathbb{R})$. View H as a topological group with the relative topology from $\mathrm{GL}(n,\mathbb{R})$. Then H has a Lie group structure that is compatible with its topology.*

Proof. Let $K \subset M_n(\mathbb{R})$ be the orthogonal complement to $\mathrm{Lie}(H)$, as in equation (1.20). The map $X \oplus Y \mapsto \exp(X)\exp(Y)$ from $\mathrm{Lie}(H) \oplus K$ to $\mathrm{GL}(n,\mathbb{R})$ has differential $X \oplus Y \mapsto X + Y$ at 0 by Lemma 1.3.6. Hence by the inverse function theorem, Lemma 1.3.3, and Theorem 1.3.11 there are open neighborhoods of 0:

$$U \subset \mathrm{Lie}(H), \quad V \subset K, \quad W \subset M_n(\mathbb{R}),$$

with the properties:

(1) If $\Omega = \{g_1 g_2 g_3 \,:\, g_i \in \exp W\}$, then the map $\log : \Omega \longrightarrow M_n(\mathbb{R})$ is a diffeomorphism onto its image. Furthermore, $W = -W$.

(2) There are real-analytic maps $\varphi : \Omega \longrightarrow U$ and $\psi : \Omega \longrightarrow V$ such that $g \in \Omega$ can be factored as $g = \exp(\varphi(g))\exp(\psi(g))$.

(3) $H \cap \Omega = \{g \in \Omega \,:\, \psi(g) = 0\}$.

Give H the relative topology as a closed subset of $\mathrm{GL}(n,\mathbb{R})$. We will define a C^∞ d-atlas for H as follows ($d = \dim \mathrm{Lie}(H)$): For any $h \in H$, $h\exp U = (h\exp W) \cap H$ by (3). Hence the set $U_h = h\exp U$ is an open neighborhood of h in H. Define

$$\Phi_h(h\exp X) = X \quad \text{for } X \in U.$$

Then by (2) the map $\Phi_h : U_h \longrightarrow U$ is a homeomorphism. Suppose $h_1 \exp X_1 = h_2 \exp X_2$ with $h_i \in H$ and $X_i \in U$. Then $\Phi_{h_2}(\Phi_{h_1}^{-1}(X_1)) = X_2$. Now $h_2^{-1}h_1 = \exp X_2 \exp(-X_1) \in (\exp W)^2$, so

$$\exp X_2 = h_2^{-1} h_1 \exp X_1 \in \Omega.$$

It follows from (1) and (2) that $X_2 = \log(h_2^{-1}h_1 \exp X_1)$ is a C^∞ function of X_1 with values in $\mathrm{Lie}(H)$ (in fact, it is a real-analytic function). Thus $\{(U_h, \Phi_h)\}_{h\in H}$ is a C^∞ d-atlas for H.

It remains to show that the map $h_1, h_2 \mapsto h_1 h_2^{-1}$ is C^∞ from $H \times H$ to H. Let $X_i \in \mathfrak{h}$. Then

$$h_1 \exp X_1 \exp(-X_2) h_2^{-1} = h_1 h_2^{-1} \exp(\mathrm{Ad}(h_2)X_1)\exp(-\mathrm{Ad}(h_2)X_2).$$

Fix h_2 and set

$$U^{(2)} = \{(X_1, X_2) \in U \times U \,:\, \exp(\mathrm{Ad}(h_2)X_1)\exp(-\mathrm{Ad}(h_2)X_2) \in \exp W\}.$$

Then $U^{(2)}$ is a open neighborhood of $(0,0)$ in $\mathrm{Lie}(H) \times \mathrm{Lie}(H)$. By (3) we have

$$\beta(X_1, X_2) = \log\big(\exp(\mathrm{Ad}(h_2)X_1)\exp(-\mathrm{Ad}(h_2)X_2)\big) \in U$$

for $(X_1, X_2) \in U^{(2)}$. The map $\beta : U^{(2)} \longrightarrow U$ is clearly C^∞ and we can write

$$h_1 \exp X_1 \exp(-X_2)h_2^{-1} = h_1 h_2^{-1} \exp \beta(X_1, X_2)$$

for $(X_1, X_2) \in U^{(2)}$. This shows that multiplication and inversion are C^∞ maps on H. ♦

Remark 1.3.13. An atlas $\mathcal{A} = \{U_\alpha, \Phi_\alpha\}_{\alpha \in I}$ on a C^∞ manifold X is *real analytic* if each transition map $\Phi_\alpha \circ \Phi_\beta^{-1}$ is given by a convergent power series in the local coordinates at each point in its domain. Such an atlas defines a *real-analytic* (class C^ω) manifold structure on X, just as in the C^∞ case, since the composition of real-analytic functions is real analytic. A map between manifolds of class C^ω is *real-analytic* if it is given by convergent power series in local real-analytic coordinates. The exponential coordinate atlas on the subgroup H defined in the proof of Theorem 1.3.12 is real analytic, and the group operations on H are real-analytic maps relative to the C^ω manifold structure defined by this atlas.

1.3.6 Differentials of Homomorphisms

Let $G \subset \mathrm{GL}(n, \mathbb{R})$ and $H \subset \mathrm{GL}(m, \mathbb{R})$ be closed subgroups.

Proposition 1.3.14. *Let $\varphi : H \longrightarrow G$ be a continuous homomorphism. There exists a unique Lie algebra homomorphism $d\varphi : \mathrm{Lie}(H) \longrightarrow \mathrm{Lie}(G)$ such that*

$$\varphi(\exp(X)) = \exp(d\varphi(X)) \quad \textit{for all } X \in \mathrm{Lie}(H).$$

This homomorphism is called the differential *of φ.*

Proof. If $X \in \mathrm{Lie}(H)$ then $t \mapsto \varphi(\exp tX)$ defines a continuous homomorphism of $\mathbb{R}$ into $\mathrm{GL}(n, \mathbb{R})$. Hence Theorem 1.3.5 implies that there exists $\mu(X) \in M_n(\mathbb{R})$ such that

$$\varphi(\exp(tX)) = \exp(t\mu(X)) \quad \text{for all } t \in R.$$

It is clear from the definition that $\mu(tX) = t\mu(X)$ for all $t \in \mathbb{R}$. We will use Proposition 1.3.7 to prove that $\mu : \mathrm{Lie}(H) \longrightarrow \mathrm{Lie}(G)$ is a Lie algebra homomorphism.

If $X, Y \in \mathrm{Lie}(H)$ and $t \in \mathbb{R}$, then by continuity of φ and formula (1.14) we have

$$\begin{aligned}
\varphi\Big(\exp\big(tX + tY\big)\Big) &= \lim_{k\to\infty} \varphi\Big(\exp\big(\tfrac{t}{k}X\big)\exp\big(\tfrac{t}{k}Y\big)\Big)^k \\
&= \lim_{k\to\infty} \Big(\exp\big(\tfrac{t}{k}\mu(X)\big)\exp\big(\tfrac{t}{k}\mu(Y)\big)\Big)^k \\
&= \exp\big(t\mu(X) + t\mu(Y)\big).
\end{aligned}$$

Hence the uniqueness assertion in Theorem 1.3.5 implies that

$$\mu(X+Y) = \mu(X) + \mu(Y).$$

Likewise, now using formula (1.15), we prove that $\mu([X,Y]) = [\mu(X), \mu(Y)]$. We define $d\varphi(X) = \mu(X)$. ⧫

By Remark 1.3.13 G and H are real-analytic manifolds relative to charts given by exponential coordinates.

Corollary 1.3.15. *The homomorphism φ is real analytic.*

Proof. This follows immediately from the definition of the Lie group structures on G and H using exponential coordinates (as in the proof of Theorem 1.3.12), together with Proposition 1.3.14. ⧫

1.3.7 Lie Algebras and Vector Fields

We call the entries x_{ij} in the matrix $X = [x_{ij}] \in M_n(\mathbb{R})$ the *standard coordinates* on $M_n(\mathbb{R})$. That is, the functions x_{ij} are the components of X with respect to the standard basis $\{e_{ij}\}$ for $M_n(\mathbb{R})$ (where the elementary matrix e_{ij} has exactly one nonzero entry, which is a 1 in the i,j position). If U is an open neighborhood of I in $M_n(\mathbb{R})$ and $f \in C^\infty(U)$, then

$$\frac{\partial}{\partial x_{ij}} f(u) = \frac{d}{dt} f(u + te_i)\Big|_{t=0}$$

is the usual partial derivative relative to the standard coordinates.

If we use the multiplicative structure on $M_n(\mathbb{R})$ and the exponential map instead of the additive structure, then we can define

$$\frac{d}{dt} f\big(u\exp(te_{ij})\big)\Big|_{t=0} = \frac{d}{dt} f\big(u + tue_{ij})\big)\Big|_{t=0},$$

since $\exp(tA) = I + tA + \mathrm{O}(t^2)$ for $A \in M_n(\mathbb{R})$. Now $ue_{ij} = \sum_{k=1}^n x_{ki}(u)e_{kj}$. Thus by the chain rule we find that

$$\frac{d}{dt} f\big(u\exp(te_{ij})\big)\Big|_{t=0} = E_{ij}f(u) \quad \text{for } u \in U,$$

where E_{ij} is the vector field

$$E_{ij} = \sum_{k=1}^n x_{ki}\,\frac{\partial}{\partial x_{kj}}. \tag{1.22}$$

on U. In general, if $A = \sum\limits_{i,j=1}^n a_{ij}e_{ij} \in M_n(\mathbb{R})$, then we can define a vector field on $M_n(\mathbb{R})$ by $X_A = \sum\limits_{i,j=1}^n a_{ij}E_{ij}$. By the chain rule we have

$$\sum_{k=1}^n a_{ij}E_{ij}f(u) = \frac{d}{dt} f\big(u\exp\big(\sum_{i,j=1}^n a_{ij}e_{ij}\big)\big)\Big|_{t=0}.$$

Hence

$$X_A f(u) = \frac{d}{dt} f\big(u \exp(tA)\big)\Big|_{t=0}. \tag{1.23}$$

Define the *left translation operator* $L(y)$ by $L(y)f(g) = f(y^{-1}g)$ for f a C^∞ function on $\mathrm{GL}(n,\mathbb{R})$ and $y \in \mathrm{GL}(n,\mathbb{R})$. It is clear from (1.23) that X_A commutes with $L(y)$ for all $y \in \mathrm{GL}(n,\mathbb{R})$. Furthermore, at the identity element we have

$$(X_A)_I = \sum_{i,j} a_{ij} \Big(\frac{\partial}{\partial x_{ij}}\Big)_I \in T(M_n(\mathbb{R}))_I, \tag{1.24}$$

since $(E_{ij})_I = \big(\frac{\partial}{\partial x_{kj}}\big)_I$. It is important to observe that equation (1.24) only holds at I; the vector field X_A is a linear combination (with real coefficients) of the *variable coefficient* vector fields $\{E_{ij}\}$, whereas the *constant coefficient* vector field $\sum a_{ij} \frac{\partial}{\partial x_{ij}}$ does not commute with $L(y)$.

The map $A \mapsto X_A$ is obviously linear; we claim that it also satisfies

$$[X_A, X_B] = X_{[A,B]} \tag{1.25}$$

and hence is a Lie algebra homomorphism. Indeed, using linearity in A and B, we see that it suffices to verify formula (1.25) when $A = e_{ij}$ and $B = e_{kl}$. In this case $[e_{ij},\, e_{kl}] = \delta_{jk} e_{il} - \delta_{il} e_{kj}$ by matrix multiplication, whereas the commutator of the vector fields is

$$\begin{aligned}[E_{ij}, E_{kl}] &= \sum_{p,q} x_{pi} \Big\{ \frac{\partial}{\partial x_{pj}} (x_{qk}) \Big\} \frac{\partial}{\partial x_{ql}} - \sum_{p,q} x_{qk} \Big\{ \frac{\partial}{\partial x_{ql}} (x_{pi}) \Big\} \frac{\partial}{\partial x_{pj}} \\ &= \delta_{jk} E_{il} - \delta_{il} E_{kj}.\end{aligned}$$

Hence $[e_{ij},\, e_{kl}] \mapsto [E_{ij}, E_{kl}]$ as claimed.

Now assume that G is a closed subgroup of $\mathrm{GL}(n,\mathbb{R})$ and let $\mathrm{Lie}(G)$ be defined as in (1.12) using one-parameter subgroups. We know from Corollary 1.3.15 that the injection map $i_G : G \longrightarrow \mathrm{GL}(n,\mathbb{R})$ is C^∞ (in fact, analytic).

Lemma 1.3.16. *One has* $(di_G)_I(T(G)_I) = \{(X_A)_I \,:\, A \in \mathrm{Lie}(G)\}$.

Proof. For $A \in \mathrm{Lie}(G)$ the one-parameter group $t \mapsto \exp(tA)$ is a C^∞ map from $\mathbb{R}$ to G, by definition of the manifold structure of G (see Theorem 1.3.12). We define the tangent vector $v_A \in T(G)_I$ by

$$v_A f = \frac{d}{dt} f\big(\exp(tA)\big)\Big|_{t=0} \quad \text{for } f \in C^\infty(G).$$

By definition of the differential of a smooth map, we then have $(di_G)_I(v_A) f = (X_A)_I$. This shows that

$$(di_G)_I(T(G)_I) \supset \{(X_A)_I \,:\, A \in \mathrm{Lie}(G)\}. \tag{1.26}$$

Since $\dim \mathrm{Lie}(A) = \dim T(G)_I$, the two spaces in (1.26) are the same. ♦

Define the *left translation operator* $L(y)$ on $C^\infty(G)$ by $L(y)f(g) = f(y^{-1}g)$ for $y \in G$ and $f \in C^\infty(G)$. We say that a smooth vector field X on G is *left invariant* if it commutes with the operators $L(y)$ for all $y \in G$:

$$X(L(y)f)(g) = (L(y)Xf)(g) \quad \text{for all } y, g \in G \text{ and } f \in C^\infty(G).$$

If $A \in \mathrm{Lie}(G)$ we set

$$X_A^G f(g) = \frac{d}{dt} f(g \exp(tA))\Big|_{t=0} \quad \text{for } f \in C^\infty(G).$$

Since the map $\mathbb{R} \times G \longrightarrow G$ given by $t, g \mapsto g \exp tA$ is smooth we see that X_A^G is a left-invariant vector field on G. When $G = \mathrm{GL}(n,\mathbb{R})$ then $X_A^G = X_A$ as defined in (1.24).

Proposition 1.3.17. *Every left-invariant regular vector field on G is of the form X_A for a unique $A \in \mathrm{Lie}(G)$. Furthermore, if $A, B \in \mathrm{Lie}(G)$ then $[X_A^G, X_B^G] = X_{[A,B]}^G$.*

Proof. Since a left-invariant vector field X is uniquely determined by the tangent vector X_I at I, the first statement follows from Lemma 1.26. Likewise, to prove the commutator formula it suffices to show that

$$[X_A^G, X_B^G]_I = \left(X_{[A,B]}^G\right)_I \quad \text{for all } A, B \in \mathrm{Lie}(G). \tag{1.27}$$

From Theorem 1.3.12 there is a coordinate chart for $\mathrm{GL}(n,\mathbb{R})$ at I whose first $d = \dim G$ coordinates are the linear coordinates given by a basis for $\mathrm{Lie}(G)$. Thus there is a neighborhood Ω of I in $\mathrm{GL}(n,\mathbb{R})$ so every C^∞ function f on the corresponding neighborhood $U = \Omega \cap G$ of I in G is of the form $\varphi|_U$, with $\varphi \in C^\infty(\Omega)$. If $g \in U$ and $A \in \mathrm{Lie}(G)$, then for $t \in \mathbb{R}$ near zero we have $g \exp tA \in U$. Hence

$$X_A\varphi(g) = \frac{d}{dt}\varphi(g \exp tA)\Big|_{t=0} = \frac{d}{dt} f(g \exp tA)\Big|_{t=0} = X_A^G f(g).$$

Thus $(X_A\varphi)|_U = X_A^G f$. Now take $B \in \mathrm{Lie}(G)$. Then

$$\begin{aligned}[X_A^G, X_B^G]f = X_A^G X_B^G f - X_B^G X_A^G f &= \left(X_A X_B \varphi - X_B X_A \varphi\right)\big|_U \\ &= \left([X_A, X_B]\varphi\right)\big|_U .\end{aligned}$$

But by (1.25) we have $[X_A, X_B]\varphi = X_{[A,B]}\varphi$. Hence

$$[X_A^G, X_B^G]f = \left(X_{[A,B]}\varphi\right)\big|_U = X_{[A,B]}^G f.$$

Since this last equation holds for all $f \in C^\infty(U)$, it proves (1.27). ♦

Let $G \subset \mathrm{GL}(n,\mathbb{R})$ and $H \subset \mathrm{GL}(m,\mathbb{R})$ be closed subgroups. If $\varphi : H \longrightarrow G$ is a continuous homomorphism, we know from Corollary 1.3.15 that φ must be real analytic. We now calculate $d\varphi_I : T(H)_I \longrightarrow T(G)_I$. Using the notation in

the proof of Lemma 1.26, we have the linear map $\mathrm{Lie}(H) \longrightarrow T(H)_I$ given by $A \mapsto v_A$ for $A \in \mathrm{Lie}(H)$. If $f \in C^\infty(G)$ then

$$d\varphi_I(v_A)f = \frac{d}{dt} f(\varphi(\exp tA))\Big|_{t=0}.$$

By Proposition 1.3.14 there is a Lie algebra homomorphism $d\phi : \mathrm{Lie}(H) \longrightarrow \mathrm{Lie}(G)$ with $\varphi(\exp(tA)) = \exp(td\varphi(A))$. Thus $d\varphi_I(v_A) = v_{d\varphi(A)}$. Since the vector field X_A^H on H is left invariant, we conclude that

$$d\varphi_h(X_A^H)_h = (X^G_{d\varphi(A)})_{\varphi(h)} \quad \text{for all } h \in H.$$

Thus for a closed subgroup G of $\mathrm{GL}(n,\mathbb{R})$ the matrix algebra version of its Lie algebra and the geometric version of its Lie algebra as the left invariant vector fields on G are isomorphic under the correspondence $A \mapsto X_A^G$, by Proposition 1.3.17. Furthermore, under this correspondence the differential of a homomorphism given in Proposition 1.3.14 is the same as the differential defined in the general Lie group context (see Appendix D.2.2).

1.3.8 Exercises

1. Show that $\exp : M_n(\mathbb{C}) \longrightarrow \mathrm{GL}(n,\mathbb{C})$ is surjective. (HINT: Use Jordan canonical form.)

2. This exercise illustrates that $\exp : M_n(\mathbb{R}) \longrightarrow \mathrm{GL}(n,\mathbb{R})$ is neither injective nor surjective when $n \geq 2$.

   (a) Let $X = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Calculate the matrix form of the one-parameter subgroup $\varphi(t) = \exp(tX)$ and show that the kernel of the homomorphism $t \mapsto \varphi(t)$ is $2\pi\mathbb{Z}$.

   (b) Let $g = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}$. Show that g is not the exponential of any real 2×2 matrix. (HINT: Assume $g = \exp(X)$. Compare the eigenvectors of X and g to conclude that X can have only one eigenvalue. Then use $\mathrm{tr}(X) = 0$ to show that this eigenvalue must be zero.)

3. Complete the proof that the Lie algebras of the classical groups in Section 1.1 are the Lie algebras with the corresponding fraktur names in Section 1.2, following the same technique used for $\mathfrak{sl}(n,\mathbb{R})$, $\mathfrak{sp}(n,\mathbb{F})$, and $\mathfrak{su}(p,q)$ in Section 1.3.4.

4. *(Notation of Exercises 1.1.5, # 4)* Show that φ is continuous and prove that $d\varphi$ is a Lie algebra isomorphism. Use this result to prove that the image of φ is open (and hence also closed) in $\mathrm{SO}(V,B)$.

5. *(Notation of Exercises 1.1.5, #6 and #7)* Show that φ is continuous and prove that $d\varphi$ is a Lie algebra isomorphism. Use this result to prove that the image of φ is open and closed in the corresponding orthogonal group.

6. *(Notation of Exercises 1.1.5, #8 and #9)* Prove that the differentials of ψ and φ are Lie algebra isomorphisms.

7. Let $X, Y \in M_n(\mathbb{R})$. Use Lemma 1.3.6 to prove that there exists an $\epsilon > 0$ and a constant $C > 0$ so that the following holds for $||X|| + ||Y|| < \epsilon$:

   (a) $\exp X \exp Y \exp(-X) = \exp\big(Y + [X,Y] + Q(X,Y)\big)$, with $Q(X,Y) \in M_n(\mathbb{R})$ and $||Q(X,Y)|| \leq C(||X|| + ||Y||)^3$.

   (b) $\exp X \exp Y \exp(-X) \exp(-Y) = \exp\big([X,Y] + P(X,Y)\big)$, with $P(X,Y) \in M_n(\mathbb{R})$ and $||P(X,Y)|| \leq C(||X|| + ||Y||)^3$.

1.4 Linear Algebraic Groups

1.4.1 Definitions and Examples

Since each classical group $G \subset \mathrm{GL}_n(\mathbb{F})$ is defined by algebraic equations, we can also study G using algebraic techniques instead of analysis. We will take the field $\mathbb{F} = \mathbb{C}$ in this setting (it could be any algebraically closed field of characteristic zero). We also require that the equations defining G are polynomials in the *complex* matrix entries (that is, they do not involve complex conjugation), in the sense of the following definition:

Definition 1.4.1. A subgroup G of $\mathrm{GL}(n,\mathbb{C})$ is a *linear algebraic group* if there is a set $\mathcal{A}$ of polynomial functions on $M_n(\mathbb{C})$ so that

$$G = \{g \in \mathrm{GL}(n,\mathbb{C}) \,:\, f(g) = 0 \text{ for all } f \in \mathcal{A}\}.$$

Here a function f on $M_n(\mathbb{C})$ is a *polynomial function* if

$$f(y) = p(x_{11}(y), x_{12}(y), \ldots, x_{nn}(y)) \quad \text{for all } y \in M_n(\mathbb{C}),$$

where $p \in \mathbb{C}[x_{11}, x_{12}, \ldots, x_{nn}]$ is a polynomial and x_{ij} are the matrix entry functions on $M_n(\mathbb{C})$.

Given a complex vector space V with $\dim V = n$, we fix a basis for V and let $\mu : \mathrm{GL}(V) \longrightarrow \mathrm{GL}(n,\mathbb{C})$ be the corresponding isomorphism obtained in Section 1.1.1. We call a subgroup $G \subset \mathrm{GL}(V)$ a *linear algebraic group* if $\mu(G)$ is an algebraic group in the sense of Definition 1.4.1 (this definition is clearly independent of the choice of basis).

Examples

1. The basic example of a linear algebraic group is $\mathrm{GL}(n,\mathbb{C})$ (take the defining set $\mathcal{A}$ of relations to consist of the zero polynomial). In the case $n = 1$ we have $\mathrm{GL}(1,\mathbb{C}) = \mathbb{C} \setminus \{0\} = \mathbb{C}^\times$, the multiplicative group of the field $\mathbb{C}$.

2. The special linear group $\mathrm{SL}(n,\mathbb{C})$ is algebraic and defined by one polynomial equation $\det(g) - 1 = 0$.

3. Let $D_n \subset \mathrm{GL}(n,\mathbb{C})$ be the subgroup of diagonal matrices. The defining equations for D_n are $x_{ij}(g) = 0$ for all $i \neq j$, so D_n is an algebraic group.

4. Let $N_n \subset \mathrm{GL}(n,\mathbb{C})$ be the subgroup of upper-triangular matrices with diagonal entries 1. The defining equations in this case are $x_{ii}(g) = 1$ for all i and $x_{ij}(g) = 0$ for all $i > j$. When $n = 2$, the group N_2 is isomorphic (as an abstract group) to the additive group of the field $\mathbb{C}$, via the map

$$z \mapsto \begin{bmatrix} 1 & z \\ 0 & 1 \end{bmatrix}$$

from $\mathbb{C}$ to N_2. We will look upon $\mathbb{C}$ as the linear algebraic group N_2.

5. Let $B_n \subset \mathrm{GL}(n,\mathbb{C})$ be the subgroup of upper-triangular matrices. The defining equations for B_n are $x_{ij}(g) = 0$ for all $i > j$, so B_n is an algebraic group.

6. Let $\Gamma \in \mathrm{GL}(n,\mathbb{C})$ and let $B_\Gamma(x,y) = x^t \Gamma y$ for $x, y \in \mathbb{C}^n$. Then B_Γ is a nondegenerate bilinear form on $\mathbb{C}^n$. Let

$$G_\Gamma = \{g \in \mathrm{GL}(n,\mathbb{C}) \,:\, g^t \Gamma g = \Gamma\}$$

be the subgroup that preserves this form. Since G_Γ is defined by quadratic equations in the matrix entries, it is an algebraic group. This shows that the orthogonal groups and the symplectic groups are algebraic subgroups of $\mathrm{GL}(n,\mathbb{C})$.

For the orthogonal or symplectic groups (when $\Gamma^t = \pm\Gamma$), there is another description of G_Γ that will be important in connection with *real forms* in this chapter and *symmetric spaces* in Chapters 11 and 12. Define

$$\sigma_\Gamma(g) = \Gamma^{-1}(g^t)^{-1}\Gamma \quad \text{for } g \in \mathrm{GL}(n,\mathbb{C}).$$

Then $\sigma_\Gamma(gh) = \sigma_\Gamma(g)\sigma_\Gamma(h)$ for $g, h \in \mathrm{GL}(n,\mathbb{C})$, $\sigma_\Gamma(I) = I$, and

$$\sigma_\Gamma(\sigma_\Gamma(g)) = \Gamma^{-1}(\Gamma^t g (\Gamma^t)^{-1})\Gamma = g$$

since $\Gamma^{-1}\Gamma^t = \pm I$. Such a map σ_S is called an *involutory automorphism* of $\mathrm{GL}(n,\mathbb{C})$. We have $g \in G_\Gamma$ if and only if $\sigma_\Gamma(g) = g$, and hence the group G_Γ is the set of *fixed points* of σ_Γ.

1.4.2 Regular Functions

We now establish some basic properties of linear algebraic groups. We begin with the notion of *regular function*. For the group $\mathrm{GL}(n,\mathbb{C})$, the ring of *regular functions* is defined as

$$\mathcal{O}[\mathrm{GL}(n,\mathbb{C})] = \mathbb{C}[x_{11}, x_{12}, \ldots, x_{nn}, \det(x)^{-1}].$$

This is the commutative algebra over $\mathbb{C}$ generated by the matrix entry functions $\{x_{ij}\}$ and the function $\det(x)^{-1}$, with the relation $\det(x) \cdot \det(x)^{-1} = 1$ (where $\det(x)$ is expressed as a polynomial in $\{x_{ij}\}$ as usual).

For any complex vector space V of dimension n, let $\varphi : \mathrm{GL}(V) \longrightarrow \mathrm{GL}(n,\mathbb{C})$ be the group isomorphism defined in terms of a basis for V. The algebra $\mathcal{O}[\mathrm{GL}(V)]$

of *regular functions* on $\mathrm{GL}(V)$ is defined as all functions $f \circ \varphi$, where f is a regular function on $\mathrm{GL}(n, \mathbb{C})$. This definition is clearly independent of the choice of basis for V.

The regular functions on $\mathrm{GL}(V)$ that are linear combination of the matrix entry functions x_{ij}, relative to some basis for V, can be descibed in the following basis-free way: Given $B \in \mathrm{End}(V)$, we define a function f_B on $\mathrm{End}(V)$ by

$$f_B(Y) = \mathrm{tr}_V(YB), \quad \text{for } Y \in \mathrm{End}(V). \tag{1.28}$$

For example, when $V = \mathbb{C}^n$ and $B = e_{ij}$, then $f_{e_{ij}}(Y) = x_{ji}(Y)$. Since the map $B \mapsto f_B$ is linear, it follows that each function f_B on $\mathrm{GL}(n, \mathbb{C})$ is a linear combination of the matrix-entry functions and hence is regular. Furthermore, the algebra $\mathcal{O}[\mathrm{GL}(n, \mathbb{C})]$ is generated by $\{f_B : B \in M_n(\mathbb{C})\}$ and $(\det)^{-1}$. Thus for any finite-dimensional vector space V the algebra $\mathcal{O}[\mathrm{GL}(V)]$ is generated by $(\det)^{-1}$ and the functions f_B, for $B \in \mathrm{End}(V)$.

An element $g \in \mathrm{GL}(V)$ acts on $\mathrm{End}(V)$ by left and right multiplication, and we have

$$f_B(gY) = f_{Bg}(Y), \quad f_B(Yg) = f_{gB}(Y) \quad \text{for } B, Y \in \mathrm{End}(V).$$

Thus the functions f_B allow us to transfer properties of the linear action of g on $\mathrm{End}(V)$ to properties of the action of g on functions on $\mathrm{GL}(V)$, as we will see in later sections.

Definition 1.4.2. *Let $G \subset \mathrm{GL}(V)$ be an algebraic subgroup. A complex-valued function f on G is* regular *if it is the restriction to G of a regular function on $\mathrm{GL}(V)$.*

The set $\mathcal{O}[G]$ of regular functions on G is a commutative algebra over $\mathbb{C}$ under pointwise multiplication. It has a finite set of generators, namely the restrictions to G of $(\det)^{-1}$ and the functions f_B, with B varying over any linear basis for $\mathrm{End}(V)$. Set

$$\mathcal{I}_G = \{f \in \mathcal{O}[\mathrm{GL}(V)] : f(G) = 0\}.$$

This is an ideal in $\mathcal{O}[\mathrm{GL}(V)]$ that we can describe in terms of the algebra $\mathcal{P}(\mathrm{End}(V))$ of polynomials on $\mathrm{End}(V)$ by

$$\mathcal{I}_G = \bigcup_{p \geq 0} \{(\det)^{-p} f : f \in \mathcal{P}(\mathrm{End}(V)), \quad f(G) = 0\}. \tag{1.29}$$

The map $f \mapsto f|_G$ gives an algebra isomorphism

$$\mathcal{O}[G] \cong \mathcal{O}[\mathrm{GL}(V)]/\mathcal{I}_G. \tag{1.30}$$

Let G and H be linear algebraic groups and let $\varphi : G \longrightarrow H$ be a map. For $f \in \mathcal{O}[H]$ define the function $\varphi^*(f)$ on G by $\varphi^*(f)(g) = f(\varphi(g))$. We say that φ is a *regular map* if $\varphi^*(\mathcal{O}[H]) \subset \mathcal{O}[G]$.

Definition 1.4.3. An *algebraic group homomorphism* $\varphi : G \longrightarrow H$ is a group homomorphism that is a regular map. We say that G and H are *isomorphic as algebraic groups* if there exists an algebraic group homomorphism $\varphi : G \longrightarrow H$ that has a regular inverse.

Given linear algebraic groups $G \subset \mathrm{GL}(m, \mathbb{C})$ and $H \subset \mathrm{GL}(n, \mathbb{C})$, we make the group-theoretic direct product $K = G \times H$ into an algebraic group by the natural block diagonal embedding $K \longrightarrow \mathrm{GL}(m+n, \mathbb{C})$ as the block-diagonal matrices

$$k = \begin{bmatrix} g & 0 \\ 0 & h \end{bmatrix} \quad \text{with } g \in G \text{ and } h \in H.$$

Since polynomials in the matrix entries of g and h are polynomials in the matrix entries of k, we see that K is an algebraic subgroup of $\mathrm{GL}(m+n, \mathbb{C})$. The algebra homomorphism carrying $f' \otimes f'' \in \mathcal{O}[G] \otimes \mathcal{O}[H]$ to the function $(g, h) \mapsto f'(g) f''(h)$ on $G \times H$ gives an isomorphism

$$\mathcal{O}[G] \otimes \mathcal{O}[H] \xrightarrow{\cong} \mathcal{O}[K]$$

(see Lemma A.1.9). In particular, $G \times G$ is an algebraic group with the algebra of regular functions $\mathcal{O}[G \times G] \cong \mathcal{O}[G] \otimes \mathcal{O}[G]$.

Proposition 1.4.4. *The maps $\mu : G \times G \longrightarrow G$ and $\eta : G \longrightarrow G$ given by multiplication and inversion are regular. If $f \in \mathcal{O}[G]$ then there exists an integer p and $f_i', f_i'' \in \mathcal{O}[G]$ for $i = 1, \ldots, p$, such that*

$$f(gh) = \sum_{i=1}^{p} f_i'(g)\, f_i''(h) \quad \text{for } g, h \in G. \tag{1.31}$$

Furthermore, for fixed $g \in G$ the maps $x \mapsto L_g(x) = gx$ and $x \mapsto R_g(x) = xg$ on G are regular.

Proof. Cramer's rule says that $\eta(g) = \det(g)^{-1}\mathrm{adj}(g)$, where $\mathrm{adj}(g)$ is the transposed cofactor matrix of g. Since the matrix entries of $\mathrm{adj}(g)$ are polynomials in the matrix entries $x_{ij}(g)$, it is clear from (1.30) that $\eta^* f \in \mathcal{O}[G]$ whenever $f \in \mathcal{O}[G]$.

Let $g, h \in G$. Then

$$x_{ij}(gh) = \sum_r x_{ir}(g)\, x_{rj}(h).$$

Hence (1.31) is valid when $f = x_{ij}|_G$. It also holds when $f = (\det)^{-1}|_G$ by the multiplicative property of the determinant. Let $\mathcal{F}$ be the set of $f \in \mathcal{O}[G]$ for which (1.31) is valid. Then $\mathcal{F}$ is a subalgebra of $\mathcal{O}[G]$, and we have just verified that the matrix entry functions and $\det^{-1}$ are in $\mathcal{F}$. Since these functions generate $\mathcal{O}[G]$ as an algebra, it follows that $\mathcal{F} = \mathcal{O}[G]$.

Using the identification $\mathcal{O}[G \times G] = \mathcal{O}[G] \otimes \mathcal{O}[G]$, we can write (1.31) as

$$\mu^*(f) = \sum_i f_i' \otimes f_i''.$$

This shows that μ is a regular map. Furthermore, $L_g^*(f) = \sum_i f_i'(g) f_i''$ and $R_g^*(f) = \sum_i f_i''(g) f_i'$, which proves that L_g and R_g are regular maps. ♦

Examples

1. Let D_n be the subgroup of diagonal matrices in $\mathrm{GL}(n, \mathbb{C})$. The map

$$(x_1, \ldots, x_n) \mapsto \mathrm{diag}[x_1, \ldots, x_n]$$

from $(\mathbb{C}^\times)^n$ to D_n is obviously an isomorphism of algebraic groups. Since $\mathcal{O}[\mathbb{C}^\times] = \mathbb{C}[x, x^{-1}]$ consists of the Laurent polynomials in one variable, it follows that

$$\mathcal{O}[D_n] \cong \mathbb{C}[x_1, x_1^{-1}, \ldots, x_n, x_n^{-1}]$$

is the algebra of the Laurent polynomials in n variables. We call an algebraic group H that is isomorphic to D_n an *algebraic torus* of *rank* n.

2. Let $N_n \subset \mathrm{GL}(n, \mathbb{C})$ be the subgroup of upper-triangular matrices with unit diagonal. It is easy to show that the functions x_{ij} for $i > j$ and $x_{ii} - 1$ generate $\mathcal{I}_{N_n}$, and that

$$\mathcal{O}[N_n] \cong \mathbb{C}[x_{12}, x_{13}, \ldots, x_{n-1,n}]$$

is the algebra of polynomials in the $n(n-1)/2$ variables $\{x_{ij} : i < j\}$.

In the examples of algebraic groups G just given, the determination of generators for the ideal $\mathcal{I}_G$ and the structure of $\mathcal{O}[G]$ is straightforward because $\mathcal{I}_G$ is generated by linear functions of the matrix entries. In general, it is a difficult problem to find generators for $\mathcal{I}_G$ and to determine the structure of the algebra $\mathcal{O}[G]$.

1.4.3 Lie Algebra of an Algebraic Group

We next associate a Lie algebra of matrices to a linear algebraic group $G \subset \mathrm{GL}(n, \mathbb{C})$. Since the exponential function is not a polynomial, we must proceed somewhat differently than in Section 1.3.3. Our strategy is to adapt the vector field point of view in Section 1.3.7 to the setting of linear algebraic groups; the main change is to replace the algebra of smooth functions on G by the algebra $\mathcal{O}[G]$ of regular (rational) functions. The Lie algebra of G will then be defined as the derivations of $\mathcal{O}[G]$ that commute with left translations. The following notion of a derivation (*infinitesimal transformation*) plays an important role in Lie theory.

Definition 1.4.5. Let $\mathcal{A}$ be an algebra (not assumed to be associative) over a field $\mathbb{F}$. Then $\mathrm{Der}(\mathcal{A}) \subset \mathrm{End}(\mathcal{A})$ is the set of all linear transformations $D : \mathcal{A} \longrightarrow \mathcal{A}$ that satisfy $D(ab) = (Da)b + a(Db)$ for all $a, b \in \mathcal{A}$ (call D a *derivation* of $\mathcal{A}$).

We leave it as an exercise to verify that $\mathrm{Der}(\mathcal{A})$ is a Lie subalgebra of $\mathrm{End}(\mathcal{A})$, called the algebra of *derivations* of $\mathcal{A}$.

We begin with the case $G = \mathrm{GL}(n, \mathbb{C})$, which we view as a linear algebraic group with the algebra of regular functions $\mathcal{O}[G] = \mathbb{C}[x_{11}, x_{12}, \ldots, x_{nn}, \det^{-1}]$. A

regular vector field on G is a complex-linear transformation $X : \mathcal{O}[G] \longrightarrow \mathcal{O}[G]$ of the form

$$Xf(g) = \sum_{i,j=1}^{n} c_{ij}(g) \frac{\partial f}{\partial x_{ij}}(g) \tag{1.32}$$

for $f \in \mathcal{O}[G]$ and $g \in G$, where we assume that the coefficients $c_{ij} \in \mathcal{O}[G]$. In addition to being a linear transformation of $\mathcal{O}[G]$, the operator X satisfies

$$X(f_1 f_2)(g) = (Xf_1)(g) f_2(g) + f_1(g)(Xf_2)(g) \tag{1.33}$$

for $f_1, f_2 \in \mathcal{O}[G]$ and $g \in G$, by the product rule for differentiation. Any linear transformation X of $\mathcal{O}[G]$ that satisfies (1.33) is called a *derivation* of the algebra $\mathcal{O}[G]$. If X_1 and X_2 are derivations, then so is the linear transformation $[X_1, X_2]$, and we write $\mathrm{Der}(\mathcal{O}[G])$ for the Lie algebra of all derivations of $\mathcal{O}[G]$.

We will show that every derivation of $\mathcal{O}[G]$ is given by a regular vector field on G. For this purpose it is useful to consider Equation (1.33) with g fixed. We say that a complex linear map $v : \mathcal{O}[G] \longrightarrow \mathbb{C}$ is a *tangent vector to G at g* if

$$v(f_1 f_2) = v(f_1) f_2(g) + f_1(g) v(f_2). \tag{1.34}$$

The set of all tangent vectors at g is a vector subspace of the complex dual vector space $\mathcal{O}[G]^*$, since Equation (1.34) is linear in v. We call this vector space the *tangent space* to G at g (in the sense of algebraic groups), and denote it by $T(G)_g$. For any $A = [a_{ij}] \in M_n(\mathbb{C})$ we can define a tangent vector v_A at g by

$$v_A(f) = \sum_{i,j=1}^{n} a_{ij} \frac{\partial f}{\partial x_{ij}}(g) \quad \text{for } f \in \mathcal{O}[G]. \tag{1.35}$$

Lemma 1.4.6. *Let $G = \mathrm{GL}(n, \mathbb{C})$ and let $v \in T(G)_g$. Set $a_{ij} = v(x_{ij})$ and $A = [a_{ij}] \in M_n(\mathbb{C})$. Then $v = v_A$. Hence the map $A \mapsto v_A$ is a linear isomorphism from $M_n(\mathbb{C})$ to $T(G)_g$.*

Proof. By (1.34) we have $v(1) = v(1 \cdot 1) = 2v(1)$. Hence $v(1) = 0$. In particular, if $f = \det^k$ for some positive integer k, then

$$0 = v(f \cdot f^{-1}) = v(f) f(g)^{-1} + f(g) v(f^{-1}),$$

and so $v(1/f) = -v(f)/f(g)^2$. Hence v is uniquely determined by its restriction to the polynomial functions on G. Furthermore, $v(f_1 f_2) = 0$ whenever f_1 and f_2 are polynomials on $M_n(\mathbb{C})$ with $f_1(g) = 0$ and $f_2(g) = 0$. Let f be a polynomial function on $M_n(\mathbb{C})$. When v is evaluated on the Taylor polynomial of f centered at g, one obtains zero for the constant term and for all terms of degree greater than one. Also $v(x_{ij} - x_{ij}(g)) = v(x_{ij})$. This implies that $v = v_A$, where $a_{ij} = v(x_{ij})$. ♦

Corollary 1.4.7. *($G = \mathrm{GL}(n, \mathbb{C})$) If $X \in \mathrm{Der}(\mathcal{O}[G])$ then X is given by* (1.32)*, where $c_{ij} = X(x_{ij})$.*

Proof. For fixed $g \in G$, the linear functional $f \mapsto Xf(g)$ is a tangent vector at g. Hence $Xf(g) = v_A(f)$, where $a_{ij} = X(x_{ij})(g)$. Now define $c_{ij}(g) = X(x_{ij})(g)$ for all $g \in G$. Then $c_{ij} \in \mathcal{O}[G]$ by assumption, and Equation (1.32) holds. ♦

We continue to study the group $G = \mathrm{GL}(n, \mathbb{C})$ as an algebraic group. Just as in the Lie group case, we say that a regular vector field X on G is *left invariant* if it commutes with the left translation operators $L(y)$ for all $y \in G$ (where now these operators are understood to act on $\mathcal{O}[G]$).

Let $A \in M_n(\mathbb{C})$. Define a derivation X_A of $\mathcal{O}[G]$ by

$$X_A f(u) = \frac{d}{dt} f\big(u(I + tA)\big)\Big|_{t=0}$$

for $u \in G$ and $f \in \mathcal{O}[G]$, where the derivative is defined algebraically as usual for rational functions of the complex variable t. When $A = e_{ij}$ is an elementary matrix, we write $X_{e_{ij}} = E_{ij}$, as in Section 1.3.7 (but now understood as acting on $\mathcal{O}[G]$). Then the map $A \mapsto X_A$ is complex linear, and when $A = [a_{ij}]$ we have

$$X_A = \sum_{i,j} a_{ij} E_{ij}, \quad \text{with} \quad E_{ij} = \sum_{k=1}^{n} x_{ki} \frac{\partial}{\partial x_{kj}},$$

by the same proof as for (1.22). The commutator correspondence (1.25) holds as an equality of regular vector fields on $\mathrm{GL}(n, \mathbb{C})$ (with the same proof). Thus the map $A \mapsto X_A$ is a complex Lie algebra isomorphism from $M_n(n, \mathbb{C})$ onto the Lie algebra of left-invariant regular vector fields on $\mathrm{GL}(n, \mathbb{C})$. Furthermore,

$$X_A f_B(u) = \frac{d}{dt} \operatorname{tr}\big(u(I + tA)B\big)\Big|_{t=0} = \operatorname{tr}(uAB) = f_{AB}(u) \tag{1.36}$$

for all $A, B \in M_n(\mathbb{C})$, where the trace function f_B is defined by (1.28).

Now let $G \subset \mathrm{GL}(n, \mathbb{C})$ be a linear algebraic group. We define its Lie algebra $\mathfrak{g}$ as a complex Lie subalgebra of $M_n(\mathbb{C})$ as follows: Recall that $\mathcal{I}_G \subset \mathcal{O}[\mathrm{GL}(n, \mathbb{C})]$ is the ideal of regular functions that vanish on G. Define

$$\mathfrak{g} = \{A \in M_n(\mathbb{C}) \,:\, X_A f \in \mathcal{I}_G \quad \text{for all } f \in \mathcal{I}_G\}. \tag{1.37}$$

When $G = \mathrm{GL}(n, \mathbb{C})$, we have $\mathcal{I}_G = 0$, so $\mathfrak{g} = M_n(\mathbb{C})$ in this case, in agreement with the previous definition of $\mathrm{Lie}(G)$. An arbitrary algebraic subgroup G of $\mathrm{GL}(n, \mathbb{C})$ is closed, and hence a Lie group by Theorem 1.3.11. After developing some algebraic tools, we shall show (in Section 1.4.4) that $\mathfrak{g} = \mathrm{Lie}(G)$ is the same set of matrices, whether we consider G as an algebraic group or as a Lie group.

Let $A \in \mathfrak{g}$. Then the left-invariant vector field X_A on $\mathrm{GL}(n, \mathbb{C})$ induces a linear transformation of the quotient algebra $\mathcal{O}[G] = \mathcal{O}[\mathrm{GL}(n, \mathbb{C})]/\mathcal{I}_G$:

$$X_A(f + \mathcal{I}_G) = X_A(f) + \mathcal{I}_G$$

(since $X_A(\mathcal{I}_G) \subset \mathcal{I}_G$). For simplicity of notation we will also denote this transformation as X_A when the domain is clear. Clearly X_A is a derivation of $\mathcal{O}[G]$ that commutes with left translations by elements of G.

Proposition 1.4.8. *Let G be an algebraic subgroup of $\mathrm{GL}(n,\mathbb{C})$. Then $\mathfrak{g}$ is a Lie subalgebra of $M_n(\mathbb{C})$ (viewed as a Lie algebra over $\mathbb{C}$). Furthermore, the map $A \mapsto X_A$ is an injective complex-linear Lie algebra homomorphism from $\mathfrak{g}$ to $\mathrm{Der}(\mathcal{O}[G])$.*

Proof. Since the map $A \mapsto X_A$ is complex linear, it follows that $A + \lambda B \in \mathfrak{g}$ if $A, B \in \mathfrak{g}$ and $\lambda \in \mathbb{C}$. The differential operators $X_A X_B$ and $X_B X_A$ on $\mathcal{O}[\mathrm{GL}(V)]$ leave the subspace $\mathcal{I}_G$ invariant. Hence $[X_A, X_B]$ also leaves this space invariant. But $[X_A, X_B] = X_{[A,B]}$ by (1.25), so we have $[A, B] \in \mathfrak{g}$.

Suppose $A \in \mathrm{Lie}(G)$ and X_A acts by zero on $\mathcal{O}[G]$. Then $X_A f|_G = 0$ for all $f \in \mathcal{O}[\mathrm{GL}(n,\mathbb{C})]$. Since $I \in G$ and X_A commutes with left translations by $\mathrm{GL}(n,\mathbb{C})$, it follows that $X_A f = 0$ for all regular functions f on $\mathrm{GL}(n,\mathbb{C})$. Hence $A = 0$ by Corollary 1.4.7. ⧫

To calculate $\mathfrak{g}$ it is convenient to use the following property: If $G \subset \mathrm{GL}(n,\mathbb{C})$ and $A \in M_n(\mathbb{C})$, then A is in $\mathfrak{g}$ if and only if

$$X_A f|_G = 0 \quad \text{for all } f \in \mathcal{P}(M_n(\mathbb{C})) \cap \mathcal{I}_G. \tag{1.38}$$

This is an easy consequence of the definition of $\mathfrak{g}$ and (1.29), and we leave the proof as an exercise. Another basic relation between algebraic groups and their Lie algebras is the following:

$$\begin{aligned}&\text{If } G \subset H \text{ are linear algebraic groups with Lie algebras } \mathfrak{g} \text{ and } \mathfrak{h},\\ &\text{respectively, then } \mathfrak{g} \subset \mathfrak{h}.\end{aligned} \tag{1.39}$$

This is clear from the definition of the Lie algebras, since $\mathcal{I}_H \subset \mathcal{I}_G$.

Examples

1. Let D_n be the group of invertible diagonal $n \times n$ matrices. Then the Lie algebra $\mathfrak{d}_n$ of D_n (in the sense of algebraic groups) consists of the diagonal matrices in $M_n(\mathbb{C})$. To prove this, take any polynomial f on $M_n(\mathbb{C})$ that vanishes on D_n. Then we can write

$$f = \sum_{i \neq j} x_{ij} f_{ij},$$

where $f_{ij} \in \mathcal{P}(M_n(\mathbb{C}))$ and $1 \leq i, j \leq n$. Hence $A = [a_{ij}] \in \mathfrak{d}_n$ if and only if $X_A x_{ij}|_{D_n} = 0$ for all $i \neq j$. Since

$$X_A x_{ij} = f_{Ae_{ji}} = \sum_{p=1}^{n} a_{pj} x_{ip}$$

by (1.36), we see that $X_A x_{ij}$ vanishes on D_n for all $i \neq j$ if and only if $a_{ij} = 0$ for all $i \neq j$.

2. Let N_n be the group of upper-triangular matrices with diagonal entries 1. Then its Lie algebra $\mathfrak{n}_n$ consists of the strictly upper-triangular matrices in $M_n(\mathbb{C})$. To prove this, let f be any polynomial on $M_n(\mathbb{C})$ that vanishes on N_n. Then we can write

$$f = \sum_{i=1}^{n} (x_{ii} - 1) f_i + \sum_{1 \leq j < i \leq n} x_{ij} f_{ij},$$

where f_i and f_{ij} are polynomials on $M_n(\mathbb{C})$. Hence $A \in \mathfrak{n}_n$ if and only if $X_A x_{ij}|_{N_n} = 0$ for all $1 \leq j \leq i \leq n$. By the same calculation as in Example 1, we have

$$X_A x_{ij}|_{N_n} = a_{ij} + \sum_{p=i+1}^{n} a_{pj} x_{ip}.$$

Hence $A \in \mathfrak{n}_n$ if and only if $a_{ij} = 0$ for all $1 \leq j \leq i \leq n$.

3. Let $1 \leq p \leq n$ and let P be the subgroup of $\mathrm{GL}(n, \mathbb{C})$ consisting of all matrices in block upper triangular form

$$g = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}, \quad \text{where } a \in \mathrm{GL}(p, \mathbb{C}),\ d \in \mathrm{GL}(n-p, \mathbb{C}), \text{ and } b \in M_{p,n-p}(\mathbb{C}).$$

The same arguments as in Example 2 show that the ideal $\mathfrak{I}_P$ is generated by the matrix entry functions x_{ij} with $p < i \leq n$ and $1 \leq j \leq p$ and that the Lie algebra of P (as an algebraic group) consists of all matrices X in block upper triangular form

$$X = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}, \quad \text{where } A \in M_p(\mathbb{C}),\ D \in M_{n-p}(\mathbb{C}), \text{ and } B \in M_{p,n-p}(\mathbb{C}).$$

1.4.4 Algebraic Groups as Lie Groups

We now show that a linear algebraic group over $\mathbb{C}$ is a Lie group and that the Lie algebra defined using continuous one-parameter subgroups coincides with the Lie algebra defined using left-invariant derivations of the algebra of regular functions. For $Z = [z_{pq}] \in M_n(\mathbb{C})$ we write $Z = X + \mathrm{i}Y$, where i is a fixed choice of $\sqrt{-1}$ and $X, Y \in M_n(\mathbb{R})$. Then the map

$$Z \mapsto \begin{bmatrix} X & Y \\ -Y & X \end{bmatrix}$$

is an isomorphism between $M_n(\mathbb{C})$ considered as an associative algebra over $\mathbb{R}$ and the subalgebra of $M_{2n}(\mathbb{R})$ consisting of matrices A such that $AJ = JA$, where

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \quad \text{with } I = I_n$$

as in Section 1.3.4.

Recall that we associate to a closed subgroup G of $\mathrm{GL}(2n, \mathbb{R})$ the matrix Lie algebra

$$\mathrm{Lie}(G) = \{A \in M_{2n}(\mathbb{R}) \,:\, \exp(tA) \in G \text{ for all } t \in \mathbb{R}\} \tag{1.40}$$

and give G the Lie group structure using exponential coordinates (Theorem 1.3.12). For example, when $G = \mathrm{GL}(n, \mathbb{C})$, then the Lie algebra of $\mathrm{GL}(n, \mathbb{C})$ (as a real Lie group) is just $M_n(\mathbb{C})$ looked upon as a subspace of $M_{2n}(\mathbb{R})$ as above. This is the same matrix Lie algebra as in the sense of linear algebraic groups, but with scalar multiplication restricted to $\mathbb{R}$. We now prove that the same relation between the Lie algebras holds for every linear algebraic group.

Theorem 1.4.9. *Let G be an algebraic subgroup of $\mathrm{GL}(n,\mathbb{C})$ with Lie algebra $\mathfrak{g}$ as an algebraic group. Then G has the structure of a Lie group whose Lie algebra as a Lie group is $\mathfrak{g}$ looked upon as a Lie algebra over $\mathbb{R}$. If G and H are linear algebraic groups then a homomorphism in the sense of linear algebraic groups is a Lie group homomorphism.*

Proof. By definition, G is a closed subgroup of $\mathrm{GL}(n,\mathbb{C})$ and hence of $\mathrm{GL}(2n,\mathbb{R})$. Thus G has a Lie group structure and a Lie algebra $\mathrm{Lie}(G)$ defined by (1.40), which is a Lie subalgebra of $M_n(\mathbb{C})$ looked upon as a vector space over $\mathbb{R}$.

If $A \in \mathrm{Lie}(G)$ and $f \in \mathcal{I}_G$ (see Section 1.4.3), then $f(g\exp(tA)) = 0$ for $g \in G$ and all $t \in \mathbb{R}$. Hence

$$0 = \frac{d}{dt} f(g\exp(tA))\big|_{t=0} = \frac{d}{dt} f(g(I+tA))\big|_{t=0} = X_A f(g)$$

(see Section 1.4.3), so we have $A \in \mathfrak{g}$. Thus $\mathrm{Lie}(G)$ is a subalgebra of $\mathfrak{g}$ (looked upon as a real vector space).

Conversely, given $A \in \mathfrak{g}$, we must show that $\exp tA \in G$ for all $t \in \mathbb{R}$. Since G is algebraic, this is the same as showing that

$$f(\exp tA) = 0 \quad \text{for all } f \in \mathcal{I}_G \text{ and all } t \in \mathbb{R}. \tag{1.41}$$

Given $f \in \mathcal{I}_G$, we set $\varphi(t) = f(\exp tA)$ for $t \in \mathbb{C}$. Then $\varphi(t)$ is an analytic function of t, since it is a polynomial in the complex matrix entries $z_{pq}(\exp tA)$ and $\det(\exp -tA)$. Hence by Taylor's theorem

$$\varphi(t) = \sum_{k=0}^{\infty} \varphi^{(k)}(0)\frac{t^k}{k!}$$

with the series converging absolutely for all $t \in \mathbb{C}$. Since $\exp(tA) = I + tA + O(t^2)$, it follows from the definition of the vector field X_A that

$$\varphi^{(k)}(0) = (X_A^k f)(I) \quad \text{for all nonnegative integers } k.$$

But $X_A^k f \in \mathcal{I}_G$ since $A \in \mathfrak{g}$, so we have $(X_A^k f)(I) = 0$. Hence $\varphi(t) = 0$ for all t, which proves (1.41). Thus $\mathfrak{g} \subset \mathrm{Lie}(G)$.

The last assertion of the theorem is clear because polynomials in the matrix entries $\{z_{ij}\}$ and $\det^{-1}$ are C^∞ functions relative to the real Lie group structure. ♦

When G is a linear algebraic group, we shall denote the Lie algebra of G either by $\mathfrak{g}$ or $\mathrm{Lie}(G)$, as a Lie algebra over $\mathbb{C}$. When G is viewed as a real Lie group, then $\mathrm{Lie}(G)$ is viewed as a vector space over $\mathbb{R}$.

1.4.5 Exercises

1. Let $D_n = (\mathbb{C}^\times)^n$ (an algebraic torus of rank n). Suppose D_k is isomorphic to D_n as an algebraic group. Prove that $k = n$. (HINT: The given

group isomorphism induces a surjective algebra homomorphism from $\mathcal{O}[D_k]$ onto $\mathcal{O}[D_n]$; clear denominators to obtain a polynomial relation of the form $x_n f(x_1, \ldots, x_k) = g(x_1, \ldots, x_k)$, which implies $n \leq k$.)

2. Let $\mathcal{A}$ be a finite-dimensional associative algebra over $\mathbb{C}$ with unit 1. Let G be the set of all $g \in \mathcal{A}$ such that g is invertible in $\mathcal{A}$. For $z \in \mathcal{A}$ let $L_a \in \mathrm{End}(\mathcal{A})$ be the operator of left multiplication by a. Define $\Phi : G \longrightarrow \mathrm{GL}(\mathcal{A})$ by $\Phi(g) = L_g$.

   (a) Show that $\Phi(G)$ is a linear algebraic subgroup in $\mathrm{GL}(\mathcal{A})$. (HINT: To find a set of algebraic equations for $\Phi(G)$, prove that $T \in \mathrm{End}(\mathcal{A})$ commutes with all the operators of right multiplication by elements of $\mathcal{A}$ if and only if $T = L_a$ for some $a \in \mathcal{A}$.)

   (b) For $a \in \mathcal{A}$, show that there is a left-invariant vector field X_a on G such that

   $$X_a f(g) = \frac{d}{dt} f(g(1 + ta))|_{t=0}$$

   for $f \in \mathcal{O}[G]$.

   (c) Let $\mathcal{A}_{\mathrm{Lie}}$ be the vector space $\mathcal{A}$ with Lie bracket $[a, b] = ab - ba$. Prove that the map $a \mapsto X_a$ is an isomorphism from $\mathcal{A}_{\mathrm{Lie}}$ onto the left-invariant vector fields on G. (HINT: Adapt the arguments used for $\mathrm{GL}(n, \mathbb{C})$ in Section 1.4.3.)

   (d) Let $\{u_\alpha\}$ be a basis for $\mathcal{A}$ (as a vector space), and let $\{u_\alpha^*\}$ be the dual basis. Define the *structure constants* $c_{\alpha\beta\gamma}$ by $u_\alpha u_\beta = \sum_\gamma c_{\alpha\beta\gamma}\, u_\gamma$. Let $\partial/\partial u_\alpha$ denote the directional derivative in the direction u_α. Prove that

   $$X_{u_\beta} = \sum_\gamma \varphi_{\beta\gamma} \frac{\partial}{\partial u_\gamma},$$

   where $\varphi_{\beta\gamma} = \sum_\alpha c_{\alpha\beta\gamma}\, u_\alpha^*$ is a linear function on $\mathcal{A}$. (HINT: Adapt the argument used for Corollary 1.4.7)

3. Let $\mathcal{A}$ be a finite-dimensional algebra over $\mathbb{C}$. This means that there is a *multiplication map* $\mu : \mathcal{A} \times \mathcal{A} \longrightarrow \mathcal{A}$ that is bilinear (it is not assumed to be associative). Define the *automorphism group* of $\mathcal{A}$ to be

   $$\mathrm{Aut}(\mathcal{A}) = \{g \in \mathrm{GL}(\mathcal{A}) : g\mu(X, Y) = \mu(gX, gY), \text{ for } X, Y \in \mathcal{A}\}.$$

   Show that $\mathrm{Aut}(\mathcal{A})$ is an algebraic subgroup of $\mathrm{GL}(\mathcal{A})$.

4. Suppose $G \subset \mathrm{GL}(n, \mathbb{C})$ is a linear algebraic group. Let $z \mapsto \varphi(z)$ be an analytic map from $\{z \in \mathbb{C} : |z| < r\}$ to $M_n(\mathbb{C})$ for some $r > 0$. Assume that $\varphi(0) = I$ and $\varphi(z) \in G$ for all $|z| < r$. Prove that the matrix $A = (d/dz)\varphi(z)|_{z=0}$ is in $\mathrm{Lie}(G)$. (HINT: Write $\varphi(z) = I + zA + z^2 F(z)$, where $F(z)$ is an analytic matrix-valued function. Then show that $X_A f(g) = (d/dz) f(g\varphi(z))|_{z=0}$ for all $f \in \mathcal{O}[\mathrm{GL}(n, \mathbb{C})]$.)

5. Let $B_\Gamma(x,y) = x^t\Gamma y$ be a nondegenerate bilinear form on $\mathbb{C}^n$, where $\Gamma \in \mathrm{GL}_n(\mathbb{C})$. Let G_Γ be the isometry group of this form. Define the *Cayley transform* $c(A) = (I+A)(I-A)^{-1}$ for $A \in M_n(\mathbb{C})$ with $\det(I-A) \neq 0$.

   (a) Suppose $A \in M_n(\mathbb{C})$ and $\det(I-A) \neq 0$. Prove that $c(A) \in G_\Gamma$ if and only if $A^t\Gamma + \Gamma A = 0$. (HINT: Use the equation $g^t\Gamma g = \Gamma$ characterizing elements $g \in G_\Gamma$.)

   (b) Give an algebraic proof (without using the exponential map) that $\mathrm{Lie}(G_\Gamma)$ consists of all $A \in M_n(\mathbb{C})$ such that

   $$A^t\Gamma + \Gamma A = 0. \qquad (\dagger)$$

   Conclude that the Lie algebra of G_Γ is the same, whether G_Γ be viewed as a Lie group or as a linear algebraic group.

   (HINT: Define $\psi_B(g) = \mathrm{tr}((g^t\Gamma g - \Gamma)B)$ for $g \in \mathrm{GL}(n,\mathbb{C})$ and $B \in M_n(\mathbb{C})$. Show that $X_A\psi_B(I) = \mathrm{tr}((A^t\Gamma + \Gamma A)B)$ for any $A \in M_n(\mathbb{C})$. Since ψ_B vanishes on G_Γ, conclude that every $A \in \mathrm{Lie}(G_\Gamma)$ satisfies (†). For the converse, take A satisfying (†), define $\varphi(z) = c(zA)$, and then apply the previous exercise and part (a).)

6. Let V be a finite-dimensional complex vector space with a nondegenerate skew-symmetric bilinear form Ω. Define $\mathrm{GSp}(V,\Omega)$ to be all $g \in \mathrm{GL}(V)$ for which there is a $\lambda \in \mathbb{C}^\times$ (depending on g) so that $\Omega(gx,gy) = \lambda\Omega(x,y)$ for all $x, y \in V$.

   (a) Show that the homomorphism $\mathbb{C}^\times \times \mathrm{Sp}(V,\Omega) \longrightarrow \mathrm{GSp}(V,\Omega)$ given by $(\lambda, g) \mapsto \lambda g$ is surjective. What is its kernel?

   (b) Show that $\mathrm{GSp}(V,\Omega)$ is an algebraic subgroup of $\mathrm{GL}(V)$.

   (c) Find $\mathrm{Lie}(G)$. (HINT: Show (a) implies $\dim_{\mathbb{C}} \mathrm{Lie}(G) = \dim_{\mathbb{C}} \mathfrak{sp}(\mathbb{C}^{2l},\Omega)+1$.)

7. Let $G = \mathrm{GL}(1,\mathbb{C})$ and let $\varphi : G \longrightarrow G$ by $\varphi(z) = \bar{z}$. Show that φ is a group homomorphism that is *not* regular.

8. Let $P \subset \mathrm{GL}(n,\mathbb{C})$ be the subgroup defined in Example 3 of Section 1.4.3.

   (a) Prove that the ideal $\mathfrak{I}_P$ is generated by the matrix entry functions x_{ij} with $p < i \leq n$ and $1 \leq j \leq p$.

   (b) Use (a) to prove that $\mathrm{Lie}(P)$ consists of all matrices in 2×2 block upper triangular form (with diagonal blocks of sizes $p \times p$ and $(n-p) \times (n-p)$).

9. Let $G \subset \mathrm{GL}(n,\mathbb{C})$. Prove that condition (1.38) characterizes $\mathrm{Lie}(G)$. (HINT: The functions in $\mathfrak{I}_G$ are of the form $\det^{-p} f$, where $f \in \mathcal{P}(M_n(\mathbb{C}))$ vanishes on G. Use this to show that if f and $X_A f$ vanish on G then so does $X_A(\det^{-p} f)$.)

10. Let $\mathcal{A}$ be an algebra over a field $\mathbb{F}$, and let D_1, D_2 be derivations of $\mathcal{A}$. Verify that $[D_1, D_2] = D_1 \circ D_2 - D_2 \circ D_1$ is a derivation of $\mathcal{A}$.

1.5 Rational Representations

Now that we have introduced the classical groups, we turn to the second main theme of the book: linear actions (*representations*) of a classical group G on finite-dimensional vector spaces. Determining all such actions might seem much harder than studying the groups directly, but it turns out, thanks to the work of É. Cartan and H. Weyl, that these representations have a very explicit structure that also yields information about G. Linear representations are also the natural setting for studying *invariants* of G, the third theme of the book.

1.5.1 Definitions and Examples

Let G be a linear algebraic group. A *representation* of G is a pair (ρ, V), where V is a complex vector space (not necessarily finite-dimensional), and $\rho : G \longrightarrow \mathrm{GL}(V)$ is a group homomorphism. We say that the representation is *regular* if $\dim V < \infty$ and the functions on G

$$g \mapsto \langle v^*, \rho(g)v \rangle, \tag{1.42}$$

which we call *matrix coefficients* of ρ, are regular, for all $v \in V$ and $v^* \in V^*$ (recall that $\langle v^*, v \rangle$ denotes the natural pairing between a vector space and its dual).

If we fix a basis for V and write out the matrix for $\rho(g)$ in this basis ($d = \dim V$),

$$\rho(g) = \begin{bmatrix} \rho_{11}(g) & \cdots & \rho_{1d}(g) \\ \vdots & \ddots & \vdots \\ \rho_{d1}(g) & \cdots & \rho_{dd}(g) \end{bmatrix},$$

then all the functions $\rho_{ij}(g)$ on G are regular. Furthermore, ρ is a regular homomorphism from G to $\mathrm{GL}(V)$, since the regular functions on $\mathrm{GL}(V)$ are generated by the matrix entry functions and $\det^{-1}$, and we have $(\det \rho(g))^{-1} = \det \rho(g^{-1})$, which is a regular function on G. Regular representations are often called *rational* representations since each entry $\rho_{ij}(g)$ is a rational function of the matrix entries of g (however, the only denominators that occur are powers of $\det g$, so these functions are defined everywhere on G).

It will be convenient to phrase the definition of regularity as follows: On $\mathrm{End}(V)$ we have the symmetric bilinear form $A, B \mapsto \mathrm{tr}_V(AB)$. This form is nondegenerate, so if $\lambda \in \mathrm{End}(V)^*$ then there exists $A_\lambda \in \mathrm{End}(V)$ such that $\lambda(X) = \mathrm{tr}_V(A_\lambda X)$. For $B \in \mathrm{End}(V)$ define the function f_B^ρ on G by

$$f_B^\rho(g) = \mathrm{tr}_V(\rho(g)B)$$

(when B has rank one, then this function is of the form (1.42)). Then (ρ, V) is regular if and only if f_B^ρ is a regular function on G, for all $B \in \mathrm{End}(V)$. We set

$$E^\rho = \{ f_B^\rho \,:\, B \in \mathrm{End}(V) \}.$$

This is the linear subspace of $\mathcal{O}[G]$ spanned by the functions in the matrix for ρ. It is finite dimensional and invariant under left and right translations by G. We call it the space of *representative functions* associated with ρ.

Suppose ρ is a representation of G on an infinite-dimensional vector space V. We say that (ρ, V) is *locally regular* if every finite-dimensional subspace $E \subset V$ is contained in a finite-dimensional G-invariant subspace F such that the restriction of ρ to F is a regular representation.

If (ρ, V) is a regular representation and $W \subset V$ is a linear subspace, then we say that W is *G-invariant* if $\rho(g)w \in W$ for all $g \in G$ and $w \in W$. In this case we obtain a representation σ of G on W by restriction of $\rho(g)$. We also obtain a representation τ of G on the quotient space V/W by setting $\tau(g)(v + W) = \rho(g)v + W$. If we take a basis for W and complete it to a basis for V, then the matrix of $\rho(g)$ relative to this basis has the block form

$$\rho(g) = \begin{bmatrix} \sigma(g) & * \\ 0 & \tau(g) \end{bmatrix} \tag{1.43}$$

(with the basis for W listed first). This matrix form shows that the representations (σ, W) and $(\tau, V/W)$ are regular.

If (ρ, V) and (τ, W) are representations of G and $T \in \mathrm{Hom}(V, W)$, we say that T is a *G intertwining map* if

$$\tau(g)T = T\rho(g) \quad \text{for all } g \in G.$$

We denote by $\mathrm{Hom}_G(V, W)$ the vector space of all G intertwining maps. The representations ρ and τ are *equivalent* if there exists an invertible G intertwining map. In this case we write $\rho \cong \tau$.

We say that a representation (ρ, V) with $V \neq \{0\}$ is *reducible* if there is a G-invariant subspace $W \subset V$ such that $W \neq \{0\}$ and $W \neq V$. This means that there exists a basis for V so that $\rho(g)$ has the block form (1.43) with all blocks of size at least 1×1. A representation that is not reducible is called *irreducible*.

Consider now the representations L and R of G on the infinite-dimensional vector space $\mathcal{O}[G]$ given by *left* and *right translations*:

$$L(x)f(y) = f(x^{-1}y), \quad R(x)f(y) = f(yx) \text{ for } f \in \mathcal{O}[G].$$

Proposition 1.5.1. *The representation $(L, \mathcal{O}[G])$ and $(R, \mathcal{O}[G])$ are locally regular.*

Proof. For any $f \in \mathcal{O}[G]$, equation (1.31) furnishes functions f_i', $f_i'' \in \mathcal{O}[G]$ so that

$$L(x)f = \sum_{i=1}^{n} f_i'(x^{-1})\, f_i'', \quad R(x)f = \sum_{i=1}^{n} f_i''(x) f_i' \,. \tag{1.44}$$

Thus the subspaces

$$V_L(f) = \mathrm{Span}\{L(x)f \,:\, x \in G\} \quad \text{and} \quad V_R(f) = \mathrm{Span}\{R(x)f \,:\, x \in G\}$$

are finite dimensional. By definition $V_L(f)$ is invariant under left translations while $V_R(f)$ is invariant under right translations. If $E \subset \mathcal{O}[G]$ is any finite-dimensional

subspace, let $f_1, \ldots, f_k$ be a basis for E. Then $F_L = \sum_{i=1}^k V_L(f_i)$ is a finite-dimensional subspace invariant under left translations, while $F_L = \sum_{i=1}^k V_L(f_i)$ is a finite-dimensional subspace invariant under right translations. From (1.44) we see that the restrictions of the representations L to F_L and R to F_R are regular. ♦

We note that $L(x)R(y)f = R(y)L(x)f$ for $f \in \mathcal{O}[G]$. We can thus define a locally regular representation τ of the product group $G \times G$ on $\mathcal{O}[G]$ by $\tau(x, y) = L(x)R(y)$. We recover the left and right translation representations of G by restricting τ to the subgroups $G \times \{1\}$ and $\{1\} \times G$, each of which is isomorphic to G.

We may also embed G into $G \times G$ as the *diagonal subgroup* $\Delta(G) = \{(x, x) : x \in G\}$. The restriction of τ to $\Delta(G)$ gives the *conjugation representation* of G on $\mathcal{O}[G]$, which we denote by Int. It acts by

$$\mathrm{Int}(x)f(y) = f(x^{-1}yx), \text{ for } f \in \mathcal{O}[G],\ x \in G.$$

Clearly $(\mathrm{Int}, \mathcal{O}[G])$ is a locally regular representation.

1.5.2 Differential of a Rational Representation

Let $G \subset \mathrm{GL}(n, \mathbb{C})$ be a linear algebraic group with Lie algebra $\mathfrak{g} \subset \mathfrak{gl}(n, \mathbb{C})$. Let (π, V) be a rational representation of G. Viewing G and $\mathrm{GL}(V)$ as Lie groups, we can apply Proposition 1.3.14 to obtain a homomorphism (of real Lie algebras)

$$d\pi : \mathfrak{g} \longrightarrow \mathfrak{gl}(V).$$

We call $d\pi$ the *differential* of the representation π. Since $\mathfrak{g}$ is a Lie algebra over $\mathbb{C}$ in this case, we have $\pi(\exp(tA)) = \exp(d\pi(tA))$ for $A \in \mathfrak{g}$ and $t \in \mathbb{C}$. The entries in the matrix $\pi(g)$ (relative to any basis for V) are regular functions on G, so it follows that $t \mapsto \pi(\exp(tA))$ is an analytic (matrix-valued) function of the complex variable t. Thus

$$d\pi(A) = \frac{d}{dt}\pi(\exp(tA))\bigg|_{t=0}$$

and the map $A \mapsto d\pi(A)$ is complex linear. Thus $d\pi$ is a homomorphism of complex Lie algebras when G is a linear algebraic group.

This definition of $d\pi$ has made use of the exponential map and the Lie group structure on G. We can also define $d\pi$ in a purely algebraic fashion, as follows: View the elements of $\mathfrak{g}$ as left-invariant vector fields on G by the map $A \mapsto X_A$ and differentiate the entries in the matrix for π using X_A. To express this in a basis-free way, recall that every linear transformation $B \in \mathrm{End}(V)$ defines a linear function f_C on $\mathrm{End}(V)$ by

$$f_C(B) = \mathrm{tr}_V(BC) \quad \text{for } B \in \mathrm{End}(V).$$

The representative function $f_C^\pi = f_C \circ \pi$ on G is then a regular function.

Theorem 1.5.2. *The differential of a rational representation (π, V) is the unique linear map $d\pi : \mathfrak{g} \longrightarrow \mathrm{End}(V)$ such that*

$$X_A(f_C \circ \pi)(I) = f_{d\pi(A)C}(I) \quad \textit{for all } A \in \mathfrak{g} \textit{ and } C \in \mathrm{End}(V). \tag{1.45}$$

Furthermore, for $A \in \mathrm{Lie}(G)$, one has

$$X_A(f \circ \pi) = (X_{d\pi(A)} f) \circ \pi \quad \textit{for all } f \in \mathcal{O}[\mathrm{GL}(V)]. \tag{1.46}$$

Proof. For fixed $A \in \mathfrak{g}$, the map $C \mapsto X_A(f_C \circ \pi)(I)$ is a linear functional on $\mathrm{End}(V)$. Hence there exists a unique $D \in \mathrm{End}(V)$ such that

$$X_A(f_C \circ \pi)(I) = \mathrm{tr}_V(DC) = f_{DC}(I).$$

But $f_{DC} = X_D f_C$ by equation (1.36). Hence to show that $d\pi(A) = D$, it suffices to prove that equation (1.46) holds. Let $f \in \mathcal{O}[\mathrm{GL}(V)]$ and $g \in G$. Then

$$\begin{aligned} X_A(f \circ \pi)(g) &= \frac{d}{dt} f\big(\pi(g \exp(tA))\big)\Big|_{t=0} = \frac{d}{dt} f\big(\pi(g) \exp(t d\pi(A))\big)\Big|_{t=0} \\ &= (X_{d\pi(A)} f)(\pi(g)) \end{aligned}$$

by definition of the vector fields X_A on G and $X_{d\pi(A)}$ on $\mathrm{GL}(V)$. ♦

Remark 1.5.3. An algebraic-group proof of Theorem 1.5.2 and the property that $d\pi$ is a Lie-algebra homomorphism (taking (1.45) as the definition of $d\pi(A)$) is outlined in Exercises 1.5.4.

Let G and H be linear algebraic groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, respectively, and let $\pi : G \longrightarrow H$ be a regular homomorphism. If $H \subset \mathrm{GL}(V)$, then we may view (π, V) as a regular representation of G with differential $d\pi : \mathfrak{g} \longrightarrow \mathrm{End}(V)$.

Proposition 1.5.4. *One has $d\pi(\mathfrak{g}) \subset \mathfrak{h}$ and the map $d\pi : \mathfrak{g} \longrightarrow \mathfrak{h}$ is a homomorphism of complex Lie algebras. Furthermore, if $K \subset \mathrm{GL}(W)$ is a linear algebraic group and $\rho : H \longrightarrow K$ is a regular homomorphism, then $d(\rho \circ \pi) = d\rho \circ d\pi$. In particular, if $G = K$ and $\rho \circ \pi$ is the identity map, then $d\rho \circ d\pi =$ identity, so that isomorphic linear algebraic groups have isomorphic Lie algebras.*

Proof. We first verify that $d\pi(A) \in \mathfrak{h}$ for all $A \in \mathfrak{g}$. Let $f \in \mathcal{I}_H$ and $h \in H$. Then

$$\begin{aligned} (X_{d\pi(A)} f)(h) &= L(h^{-1})(X_{d\pi(A)} f)(I) = X_{d\pi(A)}(L(h^{-1})f)(I) \\ &= X_A((L(h^{-1})f) \circ \pi)(I) \end{aligned}$$

by (1.46). But $L(h^{-1})f \in \mathcal{I}_H$, so $(L(h^{-1})f) \circ \pi = 0$ since $\pi(G) \subset H$. Hence $X_{d\pi(A)} f(h) = 0$ for all $h \in H$. This shows that $d\pi(A) \in \mathfrak{h}$.

Given regular homomorphisms

$$G \xrightarrow{\pi} H \xrightarrow{\rho} K,$$

we set $\sigma = \rho \circ \pi$ and take $A \in \mathfrak{g}$ and $f \in \mathcal{O}[K]$. Then by (1.46) we have

$$(X_{d\sigma(A)} f) \circ \sigma = X_A((f \circ \rho) \circ \pi) = (X_{d\pi(A)}(f \circ \rho)) \circ \pi = (X_{d\rho(d\pi(A))} f) \circ \sigma.$$

Taking $f = f_C$ for $C \in \mathrm{End}(W)$ and evaluating the functions in this equation at I, we conclude from (1.45) that $d\sigma(A) = d\rho(d\pi(A))$. ♦

Corollary 1.5.5. *Suppose $G \subset H$ are algebraic subgroups of $\mathrm{GL}(n, \mathbb{C})$. If (π, V) is a regular representation of H, then the differential of $\pi|_G$ is $d\pi|_{\mathfrak{g}}$.*

Examples

1. Let $G \subset \mathrm{GL}(V)$ be a linear algebraic group. By definition of $\mathcal{O}[G]$, the representation $\pi(g) = g$ on V is regular. We call (π, V) the *defining* representation of G. It follows directly from the definition that $d\pi(A) = A$ for $A \in \mathfrak{g}$.

2. Let (π, V) be a regular representation. Define the *contragredient* (or *dual*) representation (π^*, V^*) by $\pi^*(g)v^* = v^* \circ \pi(g^{-1})$. Then π^* is obviously regular since

$$\langle v^*, \pi(g)v \rangle = \langle \pi^*(g^{-1})v^*, v \rangle \quad \text{for } v \in V \text{ and } v^* \in V^*.$$

If $\dim V = d$ (the *degree* of π) and V is identified with $d \times 1$ column vectors by a choice of basis, then V^* is identified with $1 \times d$ row vectors. Viewing $\pi(g)$ as a $d \times d$ matrix using the basis, we have

$$\langle v^*, \pi(g)v \rangle = v^* \pi(g) v \quad \text{(matrix multiplication)}.$$

Thus $\pi^*(g)$ acts by right multiplication on row vectors by the matrix $\pi(g^{-1})$.

The space of representative functions for π^* consists of the functions $g \mapsto f(g^{-1})$, where f is a representative function for π. If $W \subset V$ is a G-invariant subspace, then

$$W^\perp = \{v^* \in V^* \,:\, \langle v^*, w \rangle = 0 \quad \text{for all } w \in W\}$$

is a G-invariant subspace of V^*. In particular, if (π, V) is irreducible then so is (π^*, V^*). The natural vector-space isomorphism $(V^*)^* \cong V$ gives an equivalence $(\pi^*)^* \cong \pi$.

To calculate the differential of π^*, let $A \in \mathfrak{g}$, $v \in V$, and $v^* \in V^*$. Then

$$\begin{aligned} \langle d\pi^*(A)v^*, v \rangle &= \frac{d}{dt} \langle \pi^*(\exp tA)v^*, v \rangle \Big|_{t=0} = \frac{d}{dt} \langle v^*, \pi(\exp(-tA))v \rangle \Big|_{t=0} \\ &= -\langle v^*, d\pi(A)v \rangle. \end{aligned}$$

Since this holds for all v and v^*, we conclude that

$$d\pi^*(A) = -d\pi(A)^t \quad \text{for } A \in \mathfrak{g} \tag{1.47}$$

Caution: The notation $\pi^*(g)$ for the contragredient representation should not be confused with the notation B^* for the conjugate transpose of a matrix B. The pairing $\langle v^*, v \rangle$ between V^* and V is *complex linear* in each argument.

3. Let (π_1, V_1) and (π_2, V_2) be regular representations of G. Define the *direct sum* representation $\pi_1 \oplus \pi_2$ on $V_1 \oplus V_2$ by

$$(\pi_1 \oplus \pi_2)(g)(v_1 \oplus v_2) = \pi_1(g)v_1 \oplus \pi_2(g)v_2 \quad \text{for } g \in G \text{ and } v_i \in V_i.$$

Then $\pi_1 \oplus \pi_2$ is obviously a representation of G. It is regular since

$$\langle v_1^* \oplus v_2^*, (\pi_1 \oplus \pi_2)(g)(v_1 \oplus v_2)\rangle \;=\; \langle v_1^*, \pi_1(g)v_1\rangle + \langle v_2^*, \pi_2(g)v_2\rangle$$

for $v_i \in V_i$ and $v_i^* \in V_i^*$. This shows that the space of representative functions for $\pi_1 \oplus \pi_2$ is

$$E^{\pi_1 \oplus \pi_2} = E^{\pi_1} + E^{\pi_2}.$$

If $\pi = \pi_1 \oplus \pi_2$, then in matrix form we have

$$\pi(g) = \begin{bmatrix} \pi_1(g) & 0 \\ 0 & \pi_2(g) \end{bmatrix}.$$

Differentiating the matrix entries, we find that

$$d\pi(A) = \begin{bmatrix} d\pi_1(A) & 0 \\ 0 & d\pi_2(A) \end{bmatrix} \quad \text{for } A \in \mathfrak{g}.$$

Thus $d\pi(A) = d\pi_1(A) \oplus d\pi_2(A)$.

4. Let (π_1, V_1) and (π_2, V_2) be regular representations of G. Define the *tensor product* representation $\pi_1 \otimes \pi_2$ on $V_1 \otimes V_2$ by

$$(\pi_1 \otimes \pi_2)(g)(v_1 \otimes v_2) = \pi_1(g)v_1 \otimes \pi_2(g)v_2$$

for $g \in G$ and $v_i \in V$. It is clear that $\pi_1 \otimes \pi_2$ is a representation. It is regular since

$$\langle v_1^* \otimes v_2^*,\, (\pi_1 \otimes \pi_2)(g)(v_1 \otimes v_2)\rangle \;=\; \langle v_1^*,\, \pi_1(g)v\rangle\langle v_2^*, \pi_2(g)v_2\rangle$$

for $v_i \in V$ and $v_i^* \in V_i^*$. In terms of representative functions, we have

$$E^{\pi_1 \otimes \pi_2} = \mathrm{Span}(E^{\pi_1} \cdot E^{\pi_2})$$

(the sums of products of representative functions of π_1 and π_2). Set $\pi = \pi_1 \otimes \pi_2$. Then

$$\begin{aligned} d\pi(A) &= \frac{d}{dt}\Big\{ \exp\big(td\pi_1(A)\big) \otimes \exp\big(td\pi_2(A)\big)\Big\}\Big|_{t=0} \\ &= d\pi_1(A) \otimes I + I \otimes d\pi_2(A) \end{aligned} \tag{1.48}$$

5. Let (π, V) be a regular representation of G and set $\rho = \pi \otimes \pi^*$ on $V \otimes V^*$. Then by Examples 2 and 4 we see that

$$d\rho(A) = d\pi(A) \otimes I - I \otimes d\pi(A)^t. \tag{1.49}$$

However, there is the canonical isomorphism $T : V \otimes V^* \cong \mathrm{End}(V)$, with

$$T(v \otimes v^*)(u) = \langle v^*, u \rangle v.$$

Set $\sigma(g) = T\rho(g)T^{-1}$. If $Y \in \mathrm{End}(V)$ then $T(Y \otimes I) = YT$ and $T(I \otimes Y^t) = TY$. Hence $\sigma(g)(Y) = \pi(g)Y\pi(g)^{-1}$ and

$$d\sigma(A)(Y) = d\pi(A)Y - Yd\pi(A) \quad \text{for } A \in \mathfrak{g}. \tag{1.50}$$

6. Let (π, V) be a regular representation of G and set $\rho = \pi^* \otimes \pi^*$ on $V^* \otimes V^*$. Then by Examples 2 and 4 we see that

$$d\rho(A) = -d\pi(A)^t \otimes I - I \otimes d\pi(A)^t.$$

However, there is a canonical isomorphism between $V^* \otimes V^*$ and the space of bilinear forms on V, where $g \in G$ acts on a bilinear form B by

$$g \cdot B(x, y) = B(\pi(g^{-1})x, \pi(g^{-1})y).$$

If V is identified with column vectors by a choice of a basis and $B(x, y) = y^t \Gamma x$, then $g \cdot \Gamma = \pi(g^{-1})^t \Gamma \pi(g^{-1})$ (matrix multiplication). The action of $A \in \mathfrak{g}$ on B is

$$A \cdot B(x, y) = -B(d\pi(A)x, y) - B(x, d\pi(A)y).$$

We say that a bilinear form B is *invariant* under G if $g \cdot B = B$ for all $g \in G$. Likewise, we say that B is *invariant* under $\mathfrak{g}$ if $A \cdot B = 0$ for all $A \in \mathfrak{g}$. This invariance property can be expressed as

$$B(d\pi(A)x, y) + B(x, d\pi(A)y) = 0 \quad \text{for all } x, y \in V \text{ and } A \in \mathfrak{g}.$$

1.5.3 The Adjoint Representation

Let $G \subset \mathrm{GL}(n, \mathbb{C})$ be an algebraic group with Lie algebra $\mathfrak{g}$. The representation of $\mathrm{GL}(n, \mathbb{C})$ on $M_n(\mathbb{C})$ by similarity ($A \mapsto gAg^{-1}$) is regular (see Example 5 of Section 1.5.2). We now show that the restriction of this representation to G furnishes a regular representation of G. The following lemma is the key point.

Lemma 1.5.6. *Let $A \in \mathfrak{g}$ and $g \in G$. Then $gAg^{-1} \in \mathfrak{g}$.*

Proof. For $A \in M_n(\mathbb{C})$, $g \in \mathrm{GL}(n, \mathbb{C})$, and $t \in \mathbb{C}$ we have

$$g \exp(tA) g^{-1} = \sum_{k=0}^{\infty} \frac{t^k}{k!} (gAg^{-1})^k = \exp(tgAg^{-1}).$$

Now assume $A \in \mathfrak{g}$ and $g \in G$. Since $\mathfrak{g} = \mathrm{Lie}(G)$ by Theorem 1.4.9, we have $\exp(tgAg^{-1}) = g \exp(tA) g^{-1} \in G$ for all $t \in \mathbb{C}$. Hence $gAg^{-1} \in \mathfrak{g}$. ♦

We define $\mathrm{Ad}(g)A = gAg^{-1}$ for $g \in G$ and $A \in \mathfrak{g}$. Then by Lemma 1.5.6, $\mathrm{Ad}(g) : \mathfrak{g} \longrightarrow \mathfrak{g}$. The representation $(\mathrm{Ad}, \mathfrak{g})$ is called the *adjoint representation* of G. For $A, B \in \mathfrak{g}$ we have

$$\begin{aligned} \mathrm{Ad}(g)[A,B] &= gABg^{-1} - gBAg^{-1} = gAg^{-1}gBg^{-1} - gBg^{-1}gAg^{-1} \\ &= [\mathrm{Ad}(g)A, \mathrm{Ad}(g)B], \end{aligned}$$

so that $\mathrm{Ad}(g)$ is a Lie algebra automorphism. Thus $\mathrm{Ad} : G \longrightarrow \mathrm{Aut}(\mathfrak{g})$ (the group of automorphisms of $\mathfrak{g}$).

If $H \subset \mathrm{GL}(n, \mathbb{C})$ is another algebraic group with Lie algebra $\mathfrak{h}$, we denote the adjoint representations of G and H by Ad_G and Ad_H, respectively. Suppose that $G \subset H$. Since $\mathfrak{g} \subset \mathfrak{h}$ by property (1.39), we have

$$\mathrm{Ad}_H(g)A = \mathrm{Ad}_G(g)A \quad \text{for } g \in G \text{ and } A \in \mathfrak{g}. \tag{1.51}$$

Theorem 1.5.7. *The differential of the adjoint representation of G is the representation* $\mathrm{ad} : \mathfrak{g} \longrightarrow \mathrm{End}(\mathfrak{g})$ *given by*

$$\mathrm{ad}(A)(B) = [A, B] \quad \textit{for } A, B \in \mathfrak{g}. \tag{1.52}$$

Furthermore, $\mathrm{ad}(A)$ *is a derivation of* $\mathfrak{g}$*, and hence* $\mathrm{ad}(\mathfrak{g}) \subset \mathrm{Der}(\mathfrak{g})$.

Proof. Equation (1.52) is the special case of Equation (1.50) with π the defining representation of G on $\mathbb{C}^n$ and $d\pi(A) = A$. The derivation property follows from the Jacobi identity. ♦

Remark 1.5.8. If $G \subset \mathrm{GL}(n, \mathbb{R})$ is any closed subgroup, then $gAg^{-1} \in \mathrm{Lie}(G)$ for all $g \in G$ and $A \in \mathrm{Lie}(G)$ (by the same argument as in Lemma 1.5.6). Thus we can define the adjoint representation Ad of G on the real vector space $\mathrm{Lie}(G)$ as for algebraic groups, and $\mathrm{Ad} : G \longrightarrow \mathrm{Aut}(\mathfrak{g})$ is a homomorphism from G to the group of automorphisms of $\mathrm{Lie}(G)$, and Theorem 1.5.7 holds for $\mathrm{Lie}(G)$.

1.5.4 Exercises

1. Let (π, V) be a rational representation of a linear algebraic group G.

   (a) Using equation (1.45) to define $d\pi(A)$, deduce from Proposition 1.4.8 (without using the exponential map) that $d\pi([A, B]) = [d\pi(A), d\pi(B)]$ for $A, B \in \mathfrak{g}$.

   (b) Prove (without using the exponential map) that equation (1.45) implies equation (1.46). (HINT: For fixed $g \in G$ consider the linear functional

   $$f \mapsto (X_{d\pi(A)} f)(\pi(g)) - X_A(f \circ \pi)(g) \quad \text{for } f \in \mathcal{O}[\mathrm{GL}(V)].$$

   This functional vanishes when $f = f_C$. Now apply Lemma 1.4.6.)

2. Give an algebraic proof of formula (1.47) that does not use the exponential map. (HINT: Assume $G \subset \mathrm{GL}(n, \mathbb{C})$, replace $\exp(tA)$ by the rational map $t \mapsto I + tA$ from $\mathbb{C}$ to $\mathrm{GL}(n, \mathbb{C})$, and use Theorem 1.5.2.)

3. Give an algebraic proof of formula (1.48) that does not use the exponential map. (HINT: Use the method of the previous exercise.)

4. (a) Let $A \in M_n(\mathbb{C})$ and $g \in \mathrm{GL}(n,\mathbb{C})$. Give an algebraic proof (without using the exponential map) that $R(g)X_A R(g^{-1}) = X_{gAg^{-1}}$.

   (b) Use the result of (a) to give an algebraic proof of Lemma 1.5.6. (HINT: If $f \in \mathfrak{I}_G$ then $R(g)f$ and $X_A f$ are also in $\mathfrak{I}_G$.)

5. Define $\varphi(A) = \begin{bmatrix} \det(A)^{-1} & 0 \\ 0 & A \end{bmatrix}$ for $A \in \mathrm{GL}(n,\mathbb{C})$. Show that the map $\varphi : \mathrm{GL}(n,\mathbb{C}) \longrightarrow \mathrm{SL}(n+1,\mathbb{C})$ is an injective regular homomorphism and that $d\varphi(X) = \begin{bmatrix} -\operatorname{tr}(X) & 0 \\ 0 & X \end{bmatrix}$ for $X \in \mathfrak{gl}(n,\mathbb{C})$.

1.6 Jordan Decomposition

1.6.1 Rational Representations of $\mathbb{C}$

Recall that we have given the additive group $\mathbb{C}$ the structure of a linear algebraic group by embedding it into $\mathrm{SL}(2,\mathbb{C})$ with the homomorphism

$$z \mapsto \varphi(z) = \begin{bmatrix} 1 & z \\ 0 & 1 \end{bmatrix} = I + ze_{12}.$$

The regular functions on $\mathbb{C}$ are the polynomials in z, and the Lie algebra of $\mathbb{C}$ is spanned by the matrix e_{12}, which satisfies $(e_{12})^2 = 0$. Thus $\varphi(z) = \exp(ze_{12})$. We now determine all the regular representations of $\mathbb{C}$.

A matrix $A \in M_n(\mathbb{C})$ is called *nilpotent* if $A^k = 0$ for some positive integer k. A nilpotent matrix has trace zero, since zero is its only eigenvalue. A matrix $u \in M_n(\mathbb{C})$ is called *unipotent* if $u - I$ is nilpotent. Note that a unipotent transformation is nonsingular and has determinant 1, since 1 is its only eigenvalue.

Let $A \in M_n(\mathbb{C})$ be nilpotent. Then $A^n = 0$ and for $t \in \mathbb{C}$ we have

$$\exp tA = I + Y, \quad \text{where} \quad Y = tA + \frac{t^2}{2!}A^2 + \cdots + \frac{t^{n-1}}{(n-1)!}A^{n-1}$$

is also nilpotent. Hence the matrix $\exp tA$ is unipotent and $t \mapsto \exp(tA)$ is a regular homomorphism from the additive group $\mathbb{C}$ to $\mathrm{GL}(n,\mathbb{C})$.

Conversely, if $u = I + Y \in \mathrm{GL}(n,\mathbb{C})$ is unipotent, then $Y^n = 0$ and we define

$$\log u = \sum_{k=1}^{n-1} (-1)^{k+1} \frac{1}{k} Y^k.$$

By the substitution principle for power series (as in Section 1.3.2), we have

$$\exp(\log(I + A)) = I + A$$

Thus the exponential function is a bijective polynomial map from the nilpotent elements in $M_n(\mathbb{C})$ onto the unipotent elements in $\mathrm{GL}(n,\mathbb{C})$, with polynomial inverse $u \mapsto \log u$.

Lemma 1.6.1 (Taylor's formula)**.** *Suppose $A \in M_n(\mathbb{C})$ is nilpotent and f is a regular function on $\mathrm{GL}(n,\mathbb{C})$. Then there exists an integer k so that $(X_A)^k f = 0$ and*

$$f(\exp A) = \sum_{m=0}^{k-1} \frac{1}{m!}(X_A)^m f(I). \tag{1.53}$$

Proof. Since $\det(\exp zA) = 1$, the function $\varphi(z) = f(\exp zA)$ is a polynomial in $z \in \mathbb{C}$. Hence there exists a positive integer k such that $(d/dz)^k \varphi(z) = 0$. Furthermore,

$$\varphi^{(m)}(0) = (X_A^m f)(I). \tag{1.54}$$

Equation (1.53) now follows from (1.54) by evaluating $\varphi(1)$ using the Taylor expansion centered at 0. ♦

Theorem 1.6.2. *Let $G \subset \mathrm{GL}(n,\mathbb{C})$ be a linear algebraic group with Lie algebra $\mathfrak{g}$.*

(1) *Let $A \in M_n(\mathbb{C})$ be nilpotent. Then $A \in \mathfrak{g}$ if and only if $\exp A \in G$.*

(2) *Suppose $A \in \mathfrak{g}$ is a nilpotent matrix and (ρ, V) is a regular representation of G. Then $d\rho(A)$ is a nilpotent transformation on V, and*

$$\rho(\exp A) = \exp d\rho(A). \tag{1.55}$$

Proof. (1): Take $f \in \mathcal{I}_G$. If $A \in \mathfrak{g}$, then $(X_A)^m f \in \mathcal{I}_G$ for all integers $m \geq 0$. Hence $(X_A)^m f(I) = 0$ for all m, and so by Taylor's formula (1.53) we have $f(\exp A) = 0$. Thus $\exp A \in G$. Conversely, if $\exp A \in G$, then the polynomial function $\varphi(z) = f(\exp zA)$ vanishes when z is an integer, so it must vanish for all $z \in \mathbb{C}$. Hence $X_A f(I) = 0$ for all $f \in \mathcal{I}_G$, and so by left invariance of X_A we have $X_A f(g) = 0$ for all $g \in G$. Thus $A \in \mathfrak{g}$.

(2): Apply Lemma 1.6.1 to the finite-dimensional space of functions f_B^ρ, where $B \in \mathrm{End}(V)$. This gives a positive integer k such that

$$0 = (X_A)^k f_B^\rho(I) = \mathrm{tr}_V(d\rho(A)^k B) \quad \text{for all } B \in \mathrm{End}(V).$$

Hence $(d\rho(A))^k = 0$. Applying Taylor's formula to the function f_B^ρ, we obtain

$$\begin{aligned} \mathrm{tr}_V(B\rho(\exp A)) &= \sum_{m=0}^{k-1} \frac{1}{m!} X_A^m f_B^\rho(I) = \sum_{m=0}^{k-1} \frac{1}{m!} \mathrm{tr}_V(d\rho(A)^m B) \\ &= \mathrm{tr}_V(B \exp d\rho(A)). \end{aligned}$$

This holds for all B, so we obtain (1.55). ♦

Corollary 1.6.3. *If (π, V) is a regular representation of the additive group $\mathbb{C}$, then there exists a unique nilpotent $A \in \mathrm{End}(V)$ so that $\pi(z) = \exp(zA)$ for all $z \in \mathbb{C}$.*

1.6.2 Rational Representations of $\mathbb{C}^\times$

The regular representations of $\mathbb{C}^\times = \mathrm{GL}(1,\mathbb{C})$ have the following form:

Lemma 1.6.4. *Let $(\varphi, \mathbb{C}^n)$ be a regular representation of $\mathbb{C}^\times$. For $p \in \mathbb{Z}$ define $E_p = \{v \in \mathbb{C}^n : \varphi(z)v = z^p v \text{ for all } z \in \mathbb{C}^\times\}$. Then*

$$\mathbb{C}^n = \bigoplus_{p\in\mathbb{Z}} E_p \tag{1.56}$$

and hence $\varphi(z)$ is a semisimple transformation. Conversely, given a direct sum decomposition (1.56) *of $\mathbb{C}^n$, define $\varphi(z)v = z^p v$ for $z \in \mathbb{C}^\times$, $v \in E_p$. Then φ is a regular representation of $\mathbb{C}^\times$ on $\mathbb{C}^n$ that is determined (up to equivalence) by the set of integers $\{\dim E_p \,:\, p \in \mathbb{Z}\}$.*

Proof. Since $\mathcal{O}[\mathbb{C}^\times] = \mathbb{C}[z, z^{-1}]$, the entries in the matrix $\varphi(z)$ are Laurent polynomials. Thus there is an expansion

$$\varphi(z) = \sum_{p\in\mathbb{Z}} z^p T_p\,, \tag{1.57}$$

where the coefficients $T_p \in M_n(\mathbb{C})$ and only a finite number of them are nonzero. Since $\varphi(z)\varphi(w) = \varphi(zw)$, we have

$$\sum_{p,q\in\mathbb{Z}} z^p w^q\, T_p T_q = \sum_{r\in\mathbb{Z}} z^r w^r\, T_r\,.$$

Equating coefficients of $z^p w^q$ yields the relations

$$T_p T_q = 0 \quad \text{for } p \neq q, \quad T_p^2 = T_p\,. \tag{1.58}$$

Furthermore, since $\varphi(1) = I_n$, one has

$$\sum_{p\in\mathbb{Z}} T_p = I_n\,.$$

Thus the family of matrices $\{T_p \,:\, p \in \mathbb{Z}\}$ consists of mutually commuting projections and gives a *resolution of the identity* on $\mathbb{C}^n$. If $v \in \mathbb{C}^n$ and $T_p v = v$, then

$$\varphi(z)v = \sum_{q\in\mathbb{Z}} z^q T_q T_p v = z^p v$$

by (1.58), so $\mathrm{Range}(T_p) \subset E_p$. The opposite inclusion is obvious from the uniqueness of the expansion (1.57). Thus $E_p = \mathrm{Range}(T_p)$, which proves (1.56).

Conversely, given a decomposition (1.56), we let T_p be the projection onto E_p defined by this decomposition, and we define $\varphi(z)$ by (1.57). Then φ is clearly a regular homomorphism from $\mathbb{C}^\times$ into $\mathrm{GL}(n,\mathbb{C})$. ⧫

1.6.3 Jordan – Chevalley Decomposition

A matrix $A \in M_n(\mathbb{C})$ has a unique *additive* Jordan decomposition $A = S + N$ with S semisimple, N nilpotent, and $SN = NS$. Likewise, $g \in \mathrm{GL}(n, \mathbb{C})$ has a unique *multiplicative* Jordan decomposition $g = su$ with s semisimple, u unipotent, and $su = us$ (see Sections B.1.2 and B.1.3).

Theorem 1.6.5. *Let $G \subset \mathrm{GL}(n, \mathbb{C})$ be an algebraic group with Lie algebra $\mathfrak{g}$.*

(1) *If $A \in \mathfrak{g}$ and $A = S + N$ is its additive Jordan decomposition, then $S, N \in \mathfrak{g}$.*

(2) *If $g \in G$ and $g = su$ is its multiplicative Jordan decomposition, then $s, u \in G$.*

Proof. For k a nonnegative integer let $\mathcal{P}^{(k)}(M_n(\mathbb{C}))$ be the space of homogeneous polynomials of degree k in the matrix entry functions $\{x_{ij} : 1 \leq i, j \leq n\}$. This space is invariant under the right translations $R(g)$ for $g \in \mathrm{GL}(n, \mathbb{C})$ and the vector fields X_A for $A \in M_n(\mathbb{C})$, by the formula for matrix multiplication and from (1.22). Set

$$W_m = \sum_{k,r=0}^{m} (\det)^{-r} \mathcal{P}^{(k)}(M_n(\mathbb{C})). \tag{1.59}$$

The space W_m is finite dimensional and invariant under $R(g)$ and X_A because $R(g)$ preserves products of functions, X_A is a derivation, and powers of the determinant transform by

$$R(g)(\det)^{-r} = (\det g)^{-r}(\det)^{-r} \quad \text{and} \quad X_A(\det)^{-r} = -r \operatorname{tr}(A)(\det)^{-r}.$$

Furthermore,

$$\mathcal{O}[\mathrm{GL}(n, \mathbb{C})] = \bigcup_{m \geq 0} W_m.$$

Suppose $S \in M_n(\mathbb{C})$ is semisimple. We claim that the restriction of X_S to W_m is a semisimple operator for all nonnegative integers m. To verify this, we may assume $S = \mathrm{diag}[\lambda_1, \ldots, \lambda_n]$. Then the action of X_S on the generators of $\mathcal{O}[\mathrm{GL}(n, \mathbb{C})]$ is

$$X_S f_{e_{ij}} = f_{Se_{ij}} = \lambda_i f_{e_{ij}}, \quad X_S(\det)^{-1} = -\operatorname{tr}(S)(\det)^{-1}.$$

Since X_S is a derivation, it follows that any product of the functions $f_{e_{ij}}$ and $\det^{-r}$ is an eigenvector for X_S. Because such products span W_m, we see that W_m has a basis consisting of eigenvectors for X_S.

Given a semisimple element $s \in \mathrm{GL}(n, \mathbb{C})$, we use a similar argument to show that the restriction of $R(s)$ to W_m is a semisimple operator. Namely, we may assume that $s = \mathrm{diag}[\sigma_1, \ldots, \sigma_n]$ with $\sigma_i \neq 0$. Then the action of $R(s)$ on the generators of $\mathcal{O}[\mathrm{GL}(n, \mathbb{C})]$ is

$$R(s) f_{e_{ij}} = f_{se_{ij}} = \sigma_i f_{e_{ij}}, \quad R(s)(\det)^{-1} = \det(s)^{-1}(\det)^{-1}.$$

Since $R(s)(f_1 f_2) = (R(s)f_1)(R(s)f_2)$ for $f_1, f_2 \in \mathcal{O}[\mathrm{GL}(n, \mathbb{C})]$, it follows that any product of the functions $f_{e_{ij}}$ and $\det^{-r}$ is an eigenvector for $R(s)$. Because

such products span W_m, we see that W_m has a basis consisting of eigenvectors for $R(s)$.

Let $N \in M_n(\mathbb{C})$ be nilpotent and let $u \in \mathrm{GL}(n,\mathbb{C})$ be unipotent. Then by Theorem 1.6.2 the vector field X_N acts nilpotently on W_m and the operator $R(u)$ is unipotent on W_m.

The multiplicative Jordan decomposition $g = su$ for $g \in \mathrm{GL}(n,\mathbb{C})$ gives the decomposition $R(g) = R(s)R(u)$, with commuting factors. From the argument above and the uniqueness of the Jordan decomposition we conclude that the restrictions of $R(s)$ and $R(u)$ to W_m provide the semisimple and unipotent factors for the restriction of $R(g)$. Starting with the additive Jordan decomposition $A = S+N$ in $M_n(\mathbb{C})$, we likewise see that the restrictions of X_S and X_N to W_m furnish the semisimple and nilpotent parts of the restriction of X_A.

With these properties of the Jordan decompositions established, we can complete the proof as follows. Given $f \in \mathcal{I}_G$, choose m large enough so that $f \in W_m$. The Jordan decompositions of $R(g)$ and X_A on W_m are

$$R(g)|_{W_m} = (R(s)|_{W_m})(R(u)|_{W_m}), \qquad X_A|_{W_m} = X_S|_{W_m} + X_N|_{W_m}.$$

Hence there exist polynomials $\varphi(z)$, $\psi(z)$ so that

$$R(s)f = \varphi(R(g))f, \quad X_S f = \psi(X_A) \quad \text{for all } f \in W_m.$$

Thus $R(s)f$ and $X_S f$ are in $\mathcal{I}_G$, which implies that $s \in G$ and $S \in \mathfrak{g}$. ♦

Theorem 1.6.6. *Let G be an algebraic group with Lie algebra $\mathfrak{g}$. Suppose (ρ, V) is a regular representation of G.*

(1) *If $A \in \mathfrak{g}$ and $A = S + N$ is its additive Jordan decomposition, then $d\rho(S)$ is semisimple, $d\rho(N)$ is nilpotent, and $d\rho(A) = d\rho(S) + d\rho(N)$ is the additive Jordan decomposition of $d\rho(A)$ in $\mathrm{End}(V)$.*

(2) *If $g \in G$ and $g = su$ is its multiplicative Jordan decomposition in G, then $\rho(s)$ is semisimple, $\rho(u)$ is unipotent, and $\rho(g) = \rho(s)\rho(u)$ is the multiplicative Jordan decomposition of $\rho(g)$ in $\mathrm{GL}(V)$.*

Proof. We know from Theorem 1.6.2 that $d\rho(N)$ is nilpotent and $\rho(u)$ is unipotent, and since $d\rho$ is a Lie algebra homomorphism, we have

$$[d\rho(N),\, d\rho(S)] = d\rho([N,S]) = 0.$$

Likewise, $\rho(u)\rho(s) = \rho(us) = \rho(su) = \rho(s)\rho(u)$. Thus by the uniqueness of the Jordan decomposition, it suffices to prove that $d\rho(S)$ and $\rho(s)$ are semisimple. Let

$$E^\rho = \{f^\rho_B \,:\, B \in \mathrm{End}(V)\} \subset \mathcal{O}[G]$$

be the space of representative functions for ρ. Assume that $G \subset \mathrm{GL}(n,\mathbb{C})$ as an algebraic subgroup. Let $W_m \subset \mathcal{O}[\mathrm{GL}(n,\mathbb{C})]$ be the space introduced in the proof of Theorem 1.6.5, and choose an integer m so that $E^\rho \subset W_m|_G$. We have shown

in Theorem 1.6.5 that $R(s)|_{W_m}$ is semisimple. Hence $R(s)$ acts semisimply on E^ρ. Thus there is a polynomial $\varphi(z)$ with distinct roots such that

$$\varphi(R(s))E^\rho = 0. \tag{1.60}$$

However, $R(s)^k f_B^\rho = f^\rho_{\rho(s)^k B}$ for all positive integers k. By the linearity of the trace and (1.60) we conclude that $\mathrm{tr}(\varphi(\rho(s))B) = 0$ for all $B \in \mathrm{End}(V)$. Hence $\varphi(\rho(s)) = 0$, which implies that $\rho(s)$ is semisimple. The same proof applies to $d\rho(S)$. ♦

From Theorems 1.6.5 and 1.6.6 we see that every element g of G has a *semisimple component* g_s and a *unipotent component* g_u such that $g = g_s g_u$. Furthermore, this factorization is independent of the choice of defining representation $G \subset \mathrm{GL}(V)$. Likewise, every element $Y \in \mathfrak{g}$ has a unique *semisimple component* Y_s and a unique *nilpotent component* Y_n such that $Y = Y_s + Y_n$.

We denote the set of all semisimple elements of G as G_s and the set of all unipotent elements as G_u . Likewise, we denote the set of all semisimple elements of $\mathfrak{g}$ as $\mathfrak{g}_s$ and the set of all nilpotent elements as $\mathfrak{g}_n$. Suppose $G \subset \mathrm{GL}(n, \mathbb{C})$ as an algebraic subgroup. Since $T \in M_n(\mathbb{C})$ is nilpotent if and only if $T^n = 0$, we have

$$\mathfrak{g}_n = \mathfrak{g} \cap \{T \in M_n(\mathbb{C}) \,:\, T^n = 0\}$$

$$G_u = G \cap \{g \in \mathrm{GL}(n, \mathbb{C}) \,:\, (I - g)^n = 0\}.$$

Thus $\mathfrak{g}_n$ is an algebraic subset of $M_n(\mathbb{C})$ and G_u is an algebraic subset of $\mathrm{GL}(n, \mathbb{C})$.

Corollary 1.6.7. *Suppose G and H are algebraic groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$. Let $\rho : G \longrightarrow H$ be a regular homomorphism such that $d\rho : \mathfrak{g} \longrightarrow \mathfrak{h}$ is surjective. Then $\rho(G_u) = H_u$.*

Proof. It follows from Theorem 1.6.2 that the map $N \mapsto \exp(N)$ from $\mathfrak{g}_n$ to G_u is a bijection, and by Theorem 1.6.6 we have

$$H_u = \exp(\mathfrak{h}_n) = \exp(d\rho(\mathfrak{g}_n)) = \rho(G_u).$$

♦

1.6.4 Exercises

1. Let $H, X \in M_n(\mathbb{C})$ be such that $[H, X] = 2X$. Show that X is nilpotent. (HINT: Show that $[H, X^k] = 2kX^k$. Then consider the eigenvalues of $\mathrm{ad}\, H$ on $M_n(\mathbb{C})$.)

2. Show that if $X \in M_n(\mathbb{C})$ is nilpotent then there exists $H \in M_n(\mathbb{C})$ such that $[H, X] = 2X$. (HINT: Use the Jordan canonical form to write $X = gJg^{-1}$ with $g \in \mathrm{GL}(n, \mathbb{C})$ and $J = \mathrm{diag}[J_1, \cdots, J_k]$ with each J_i either 0 or a

$p_i \times p_i$ matrix of the form

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Show that there exists $H_i \in M_{p_i}(\mathbb{C})$ such that $[H_i, J_i] = 2J_i$, and then take $H = g\,\mathrm{diag}[H_1, \cdots, H_k]g^{-1}$.)

3. Show that if $0 \neq X \in M_n(\mathbb{C})$ is nilpotent, then there exist $H, Y \in M_n(\mathbb{C})$ such that $[X, Y] = H$ and $[H, X] = 2X$, $[H, Y] = -2Y$. Conclude that $\mathbb{C}X + \mathbb{C}Y + \mathbb{C}H$ is a Lie subalgebra of $M_n(\mathbb{C})$ isomorphic with $\mathfrak{sl}(2, \mathbb{C})$.

4. Suppose V and W are finite-dimensional vector spaces over $\mathbb{C}$. Let $x \in \mathrm{GL}(V)$ and $y \in \mathrm{GL}(W)$ have multiplicative Jordan decompositions $x = x_s x_u$ and $y = y_s y_u$. Prove that the multiplicative Jordan decomposition of $x \otimes y$ in $\mathrm{GL}(V \otimes W)$ is $x \otimes y = (x_s \otimes y_s)(x_u \otimes y_u)$.

5. Suppose $\mathcal{A}$ is a finite-dimensional algebra over $\mathbb{C}$ (not necessarily associative). For example, $\mathcal{A}$ could be a Lie algebra. Let $g \in \mathrm{Aut}(\mathcal{A})$ have multiplicative Jordan decomposition $g = g_s g_u$ in $\mathrm{GL}(\mathcal{A})$. Show that g_s and g_u are also in $\mathrm{Aut}(\mathcal{A})$.

6. Suppose $g \in \mathrm{GL}(n, \mathbb{C})$ satisfies $g^k = I$ for some positive integer k. Prove that g is semisimple.

7. Let $G = \mathrm{SL}(2, \mathbb{C})$.

   (a) Show that $\{g \in G : \mathrm{tr}(g)^2 \neq 4\} \subset G_s$. (HINT: Show that the elements in this set have distinct eigenvalues.)

   (b) Let $u(t) = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$ and $v(t) = \begin{bmatrix} 1 & 0 \\ t & 1 \end{bmatrix}$ for $t \in \mathbb{C}$. Show that $u(r)v(t) \in G_s$ whenever $rt(4 + rt) \neq 0$ and that $u(r)v(t)u(r) \in G_s$ whenever $rt(2 + rt) \neq 0$.

   (c) Show that G_s and G_u are not subgroups of G.

8. Let $G = \{\exp(tA) : t \in \mathbb{C}\}$, where $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

   (a) Show that G is a closed subgroup of $\mathrm{GL}(2, \mathbb{C})$. (HINT: Calculate the matrix entries of $\exp(tA)$.)

   (b) Show that G is not an algebraic subgroup of $\mathrm{GL}(2, \mathbb{C})$. (HINT: If G were an algebraic group, then G would contain the semisimple and unipotent components of $\exp(tA)$. Show that this is a contradiction.)

   (c) Find the smallest algebraic subgroup $H \subset \mathrm{GL}(2, \mathbb{C})$ such that $G \subset H$. (HINT: Use the calculations from (b).)

1.7 Real Forms of Complex Algebraic Groups

In this section we introduce the notion of a *real form* of a complex linear algebraic group G and describe the real forms when G is a classical group (these Lie groups already appeared in Section 1.1). In this case we show that G has a compact real form.

1.7.1 Real Forms and Complex Conjugations

Let $G \subset \mathrm{GL}(n, \mathbb{C})$ be an algebraic subgroup.

Definition 1.7.1. G is *defined over* $\mathbb{R}$ if the ideal $\mathcal{I}_G$ is generated by

$$\mathcal{I}_{\mathbb{R},G} = \{f \in \mathcal{I}_G \,:\, f(\mathrm{GL}(n, \mathbb{R})) \subset \mathbb{R}\}.$$

If G is defined over $\mathbb{R}$ then we set $G_{\mathbb{R}} = G \cap \mathrm{GL}(n, \mathbb{R})$ and call $G_{\mathbb{R}}$ the *group of* $\mathbb{R}$*-rational points* of G.

Notice that this definition refers to a specific embedding of G as a subgroup of $\mathrm{GL}(n, \mathbb{C})$. We will obtain a more general notion of a *real form* of G later in this section.

Examples

1. The group $G = \mathrm{GL}(n, \mathbb{C})$ is defined over $\mathbb{R}$ (since $\mathcal{I}_G = 0$), and $G_{\mathbb{R}} = \mathrm{GL}(n, \mathbb{R})$.

2. The group $G = B_n$ of $n \times n$ invertible upper-triangular matrices is defined over $\mathbb{R}$, since $\mathcal{I}_G$ is generated by the matrix-entry functions $\{x_{ij} \,:\, n \geq i > j \geq 1\}$, which are real valued on $\mathrm{GL}(n, \mathbb{R})$. In this case $G_{\mathbb{R}}$ is the group of $n \times n$ real invertible upper-triangular matrices. ♦

For $g \in \mathrm{GL}(n, \mathbb{C})$ we set $\sigma(g) = \overline{g}$ (complex conjugation of matrix entries). Then σ is a involutive automorphism of $\mathrm{GL}(n, \mathbb{C})$ as a real Lie group (σ^2 is the identity) and $d\sigma(A) = \bar{A}$ for $A \in M_n(\mathbb{C})$.

If $f \in \mathcal{O}[\mathrm{GL}(n, \mathbb{C})]$ then we set

$$\bar{f}(g) = \overline{f(\sigma(g))}.$$

Here the overline on the right denotes complex conjugation. Since f is the product of $\det^{-k}$ (for some nonnegative integer k) and a polynomial φ in the matrix entry functions, we obtain the function $\bar{f}$ by conjugating the coefficients of φ. We can write $f = f_1 + if_2$, where $f_1 = (f + \bar{f})/2$, $f_2 = (f - \bar{f})/(2i)$, and $i = \sqrt{-1}$. The functions f_1 and f_2 are real valued on $\mathrm{GL}(n, \mathbb{R})$, and $\bar{f} = f_1 - if_2$. Thus $f(\mathrm{GL}(n, \mathbb{R})) \subset \mathbb{R}$ if and only if $\bar{f} = f$.

Lemma 1.7.2. *Let $G \subset \mathrm{GL}(n, \mathbb{C})$ be an algebraic subgroup. Then G is defined over $\mathbb{R}$ if and only if $\mathcal{I}_G$ is invariant under $f \mapsto \bar{f}$.*

Proof. Assume G is defined over $\mathbb{R}$. If $f_1 \in \mathcal{I}_{\mathbb{R},G}$ then $f_1 = \bar{f}_1$. Hence $f_1(\sigma(g)) = \overline{f_1(g)} = 0$ for $g \in G$. Since $\mathcal{I}_{\mathbb{R},G}$ is assumed to generate $\mathcal{I}_G$, it follows that $\sigma(g) \in G$ for all $g \in G$. Thus for any $f \in \mathcal{I}_G$ we have $\bar{f}(g) = 0$, and hence hence $\bar{f} \in \mathcal{I}_G$.

Conversely, if $\mathcal{I}_G$ is invariant under $f \mapsto \bar{f}$, then every $f \in \mathcal{I}_G$ is of the form $f_1 + if_2$ as above, where $f_j \in \mathcal{I}_{\mathbb{R},G}$. Thus $\mathcal{I}_{\mathbb{R},G}$ generates $\mathcal{I}_G$, and so G is defined over $\mathbb{R}$. ♦

Assume that $G \subset \mathrm{GL}(n,\mathbb{C})$ is an algebraic group defined over $\mathbb{R}$. Let $\mathfrak{g} \subset M_n(\mathbb{C})$ be the Lie algebra of G. Since $\mathcal{I}_{\mathbb{R},G}$ generates $\mathcal{I}_G$ and σ^2 is the identity map, this implies that $\sigma(G) = G$. Hence σ defines a Lie group automorphism of G and $d\sigma(A) = \bar{A} \in \mathfrak{g}$ for all $A \in \mathfrak{g}$. By definition, $G_{\mathbb{R}} = \{g \in G : \sigma(g) = g\}$. Hence $G_{\mathbb{R}}$ is a Lie subgroup of G and

$$\mathrm{Lie}(G_{\mathbb{R}}) = \{A \in \mathfrak{g} : \bar{A} = A\}.$$

If $A \in \mathfrak{g}$ then $A = A_1 + iA_2$, where $A_1 = (A + \bar{A})/2$ and $A_2 = (A - \bar{A})/2i$ are in $\mathrm{Lie}(G_{\mathbb{R}})$. Thus

$$\mathfrak{g} = \mathrm{Lie}(G_{\mathbb{R}}) \oplus i\,\mathrm{Lie}(G_{\mathbb{R}}) \tag{1.61}$$

as a real vector space, so $\dim_{\mathbb{R}} \mathrm{Lie}(G_{\mathbb{R}}) = \dim_{\mathbb{C}} \mathfrak{g}$. Therefore the dimension of the Lie group $G_{\mathbb{R}}$ is the same as the dimension of G as a linear algebraic group over $\mathbb{C}$ (see Appendix A.1.6).

Remark 1.7.3. If a linear algebraic group G is defined over $\mathbb{R}$, then there is a set $\mathcal{A}$ of polynomials with real coefficients such that G is the common zeros of these polynomials in $\mathrm{GL}(n,\mathbb{C})$. The converse assertion is more subtle, however, since the elements of $\mathcal{A}$ do not necessarily generate the ideal $\mathcal{I}_G$, as required by Definition 1.7.1. For example, the group B_n of upper-triangular $n \times n$ matrices is the zero set of the polynomials $\{x_{ij}^2 : n \geq i > j \geq 1\}$; these polynomials are real on $\mathrm{GL}(n,\mathbb{R})$ but do not generate $\mathcal{I}_{B_n}$ (of course, we already know that B_n is defined over $\mathbb{R}$).

By generalizing the notion of complex conjugation we now obtain a useful criterion (not involving a specific matrix form of G) for G to be isomorphic to a linear algebraic group defined over $\mathbb{R}$. This will also furnish the general notion of a *real form* of G.

Definition 1.7.4. Let G be a linear algebraic group and let τ be an automorphism of G as a real Lie group such that τ^2 is the identity. For $f \in \mathcal{O}[G]$ define f^τ by

$$f^\tau(g) = \overline{f(\tau(g))}$$

(with the overline denoting complex conjugation). Then τ is a *complex conjugation* on G if $f^\tau \in \mathcal{O}[G]$ for all $f \in \mathcal{O}[G]$.

When $G \subset \mathrm{GL}(n,\mathbb{C})$ is defined over $\mathbb{R}$, then the map $\sigma(g) = \overline{g}$ introduced previously is a complex conjugation. In Section 1.7.2 we shall give examples of complex conjugations when G is a classical group.

Theorem 1.7.5. *Let G be a linear algebraic group and let τ be a complex conjugation on G. Then there exists a linear algebraic group $H \subset \mathrm{GL}(n,\mathbb{C})$ defined over $\mathbb{R}$ and an isomorphism $\rho : G \longrightarrow H$ such that*

$$\rho(\tau(g)) = \sigma(\rho(g)),$$

where σ is the conjugation of $\mathrm{GL}(n,\mathbb{C})$ given by complex conjugation of matrix entries.

Proof. Fix a finite set $\{1, f_1, \ldots, f_m\}$ of regular functions on G that generate $\mathcal{O}[G]$ as an algebra over $\mathbb{C}$ (for example, the restrictions to G of the matrix entry functions and $\det^{-1}$ given by the defining representation of G). Set $C(f) = f^\tau$ for $f \in \mathcal{O}[G]$ and let

$$V = \mathrm{Span}_{\mathbb{C}}\{R(g)f_k,\ R(g)Cf_k\ :\ g \in G, k = 1, \ldots, m\}.$$

Then V is invariant under G and C, since $CR(g) = R(\tau(g))C$. Let $\rho(g) = R(g)|_V$. It follows from Proposition 1.5.1 that V is finite dimensional and (ρ, V) is a regular representation of G.

We note that if $g, g' \in G$ and $f_k(g) = f_k(g')$ for all k, then $f(g) = f(g')$ for all $f \in \mathcal{O}[G]$, since the set $\{1, f_1, \ldots, f_m\}$ generates $\mathcal{O}[G]$. Letting f run over the restrictions to G of the matrix entry functions (relative to some matrix form of G), we conclude that $g = g'$. Thus if $\rho(g)f = f$ for all $f \in V$, then $g = I$, proving that $\mathrm{Ker}(\rho) = \{I\}$.

Since C^2 is the identity map, we can decompose $V = V_+ \oplus V_-$ as a vector space over $\mathbb{R}$, where

$$V_+ = \{f \in V\ :\ C(f) = f\}, \quad V_- = \{f \in V\ :\ C(f) = -f\}.$$

Because $C(if) = -iC(f)$ we have $V_- = iV_+$. Choose a basis (over $\mathbb{R}$) of V_+, say $\{v_1, \ldots, v_n\}$. Then $\{v_1, \ldots, v_n\}$ is also a basis of V over $\mathbb{C}$. If we use this basis to identify V with $\mathbb{C}^n$ then C becomes complex conjugation. To simplify the notation we will also write $\rho(g)$ for the matrix of $\rho(g)$ relative to this basis.

We now have an injective regular homomorphism $\rho : G \longrightarrow \mathrm{GL}(n,\mathbb{C})$ such that $\rho(\tau(g)) = \sigma(\rho(g))$, where σ denotes complex conjugation of matrix entries. In Chapter 11 (Theorem 11.1.5) we will prove that the image of a linear algebraic group under a regular homomorphism is always a linear algebraic group (*i.e.* a closed subgroup in the *Zariski topology*). Assuming this result (whose proof does not depend on the current argument), we conclude that $H = \rho(G)$ is an algebraic subgroup of $\mathrm{GL}(n,\mathbb{C})$. Furthermore, if $\delta \in V^*$ is the linear functional $f \mapsto f(I)$, then

$$f(g) = R(g)f(I) = \langle \delta, R(g)f \rangle. \tag{1.62}$$

Hence $\rho^*(\mathcal{O}[H]) = \mathcal{O}[G]$ since by (1.62) the functions $f_1, \ldots, f_m$ are matrix entries of (ρ, V). This proves that ρ^{-1} is a regular map.

Finally, let $f \in \mathcal{I}_H$. Then for $h = \rho(g) \in H$ we have

$$\bar{f}(h) = \overline{f(\sigma(\rho(g)))} = \overline{f(\rho(\tau(g)))} = 0.$$

Hence $\bar{f} \in \mathcal{I}_H$, so from Lemma 1.7.2 we conclude that H is defined over $\mathbb{R}$. ♦

Definition 1.7.6. Let G be a linear algebraic group. A subgroup K of G is called a *real form* of G if there exists a complex conjugation τ on G such that

$$K = \{g \in G : \tau(g) = g\}.$$

Let K be a real form of G. Then K is a closed subgroup of G, and from Theorem 1.7.5 and (1.61) we see that the dimension of K as a real Lie group is equal to the dimension of G as a complex linear algebraic group, and

$$\mathfrak{g} = \mathrm{Lie}(K) \oplus i\,\mathrm{Lie}(K) \tag{1.63}$$

as a real vector space.

One of the motivations for introducing real forms is that we can study the representations of G using the real form and its Lie algebra. Let G be a linear algebraic group, and let G° be the connected component of the identity of G (as a real Lie group). Let K be a real form of G and set $\mathfrak{k} = \mathrm{Lie}(K)$.

Proposition 1.7.7. *Suppose (ρ, V) is a regular representation of G. Then a subspace $W \subset V$ is invariant under $d\rho(\mathfrak{k})$ if and only if it is invariant under G°. In particular, V is irreducible under $\mathfrak{k}$ if and only if it is irreducible under G°.*

Proof. Assume W is invariant under $\mathfrak{k}$. Since the map $X \mapsto d\rho(X)$ from $\mathfrak{g}$ to $\mathrm{End}(V)$ is complex linear, it follows from (1.63) that W is invariant under $\mathfrak{g}$. Let $W^\perp \subset V^*$ be the annihilator of W. Then $\langle w^*, (d\rho(X))^k w\rangle = 0$ for $w \in W$, $w^* \in W^\perp$, $X \in \mathfrak{g}$, and all integers k. Hence

$$\langle w^*,\, \rho(\exp X)w\rangle = \langle w^*,\, \exp(d\rho(X))w\rangle = \sum_{k=0}^{\infty} \frac{1}{k!}\langle w^*, d\rho(X)^k w\rangle = 0,$$

so $\rho(\exp X)W \subset W$. Since G° is generated by $\exp(\mathfrak{g})$, this proves that W is invariant under G°. To prove the converse we reverse this argument, replacing X by tX and differentiating at $t = 0$. ♦

1.7.2 Real Forms of the Classical Groups

We now describe the complex conjugations and real forms of the complex classical groups. We label the groups and their real forms using É. Cartan's classification. For each complex group there is one real form that is compact.

1. (Type AI) Let $G = \mathrm{GL}(n, \mathbb{C})$ (*resp.* $\mathrm{SL}(n, \mathbb{C})$) and define $\tau(g) = \overline{g}$ for $g \in G$. Then $f^\tau = \bar{f}$ for $f \in \mathbb{C}[G]$, and so τ is a complex conjugation on G. The associated real form is $\mathrm{GL}(n, \mathbb{R})$ (*resp.* $\mathrm{SL}(n, \mathbb{R})$).

2. (Type AII) Let $G = \mathrm{GL}(2n, \mathbb{C})$ (*resp.* $\mathrm{SL}(2n, \mathbb{C})$) and let J be the $2n \times 2n$ skew-symmetric matrix from Section 1.1.2. Define $\tau(g) = J\overline{g}J^t$ for $g \in G$. Since $J^2 = -I$, we see that τ^2 is the identity. Also if f is a regular function on G then $f^\tau(g) = \bar{f}(JgJ^t)$, and so f^τ is also a regular function on G. Hence τ is a complex

conjugation on G. The equation $\tau(g) = g$ can be written as $Jg = \overline{g}J$. Hence the associated real form of G is the group $\mathrm{GL}(n, \mathbb{H})$ (*resp.* $\mathrm{SL}(n, \mathbb{H})$) from Section 1.1.4), where we view $\mathbb{H}^n$ as a $2n$-dimensional vector space over $\mathbb{C}$.

3. (Type AIII) Let $G = \mathrm{GL}(n, \mathbb{C})$ (*resp.* $\mathrm{SL}(n, \mathbb{C})$) and let $p, q \in \mathbb{N}$ be such that $p + q = n$. Let $I_{p,q} = \mathrm{diag}[I_p, -I_q]$ as in Section 1.1.2 and define $\tau(g) = I_{p,q}(g^*)^{-1}I_{p,q}$ for $g \in G$. Since $I_{p,q}^2 = I_n$, we see that τ^2 is the identity. Also if f is a regular function on G then $f^\tau(g) = \bar{f}(I_{p,q}(g^t)^{-1}I_{p,q})$, and so f^τ is also a regular function on G. Hence τ is a complex conjugation on G. The equation $\tau(g) = g$ can be written as $g^* I_{p,q} g = I_{p,q}$, so the indefinite unitary group $\mathrm{U}(p, q)$ (*resp.* $\mathrm{SU}(p, q)$) defined in Section 1.1.3 is the real form of G defined by τ. The *unitary group* $\mathrm{U}(n, 0) = \mathrm{U}(n)$ (*resp.* $\mathrm{SU}(n)$) is a compact real form of G.

4. (Type BDI) Let G be $\mathrm{O}(n, \mathbb{C}) = \{g \in \mathrm{GL}(n, \mathbb{C}) : gg^t = 1\}$ (*resp.* $\mathrm{SO}(n, \mathbb{C})$) and let $p, q \in \mathbb{N}$ be such that $p + q = n$. Let the matrix $I_{p,q}$ be as in Type AIII. Define $\tau(g) = I_{p,q}\overline{g}I_{p,q}$ for $g \in G$. Since $(g^t)^{-1} = g$ for $g \in G$, τ is the restriction to G of the complex conjugation in Example 3. We leave it as an exercise to show that the corresponding real form is isomorphic to the group $\mathrm{O}(p, q)$ (*resp.* $\mathrm{SO}(p, q)$) defined in Section 1.1.2. When $p = n$ we obtain the compact real form real form $\mathrm{O}(n)$ (*resp.* $\mathrm{SO}(n)$).

5. (Type DIII) Let G be $\mathrm{SO}(2n, \mathbb{C})$ and let J be the $2n \times 2n$ skew-symmetric matrix as in Type AII. Define $\tau(g) = J\overline{g}J^t$ for $g \in G$. Just as in Type AII, we see that τ is a complex conjugation of G. The corresponding real form is the group $\mathrm{SO}^*(2n)$ defined in Section 1.1.4 (see Exercises 1.1.5, #12).

6. (Type CI) Let G be $\mathrm{Sp}(n, \mathbb{C}) \subset \mathrm{SL}(2n, \mathbb{C})$. The equation defining G is $g^t Jg = J$, where J is the skew-symmetric matrix in Type AII. Since J is real, we may define $\tau(g) = \overline{g}$ for $g \in G$ and obtain a complex conjugation on G. The associated real form is $\mathrm{Sp}(n, \mathbb{R})$.

7. (Type CII) Let $p, q \in \mathbb{N}$ be such that $p + q = n$ and let $K_{p,q} = \mathrm{diag}[I_{p,q}, I_{p,q}] \in M_{2n}(\mathbb{R})$ as in Section 1.1.4. Let Ω be the nondegenerate skew form on $\mathbb{C}^{2n}$ with matrix

$$K_{p,q}J = \begin{bmatrix} 0 & I_{p,q} \\ -I_{p,q} & 0 \end{bmatrix},$$

with J as in Type CI. Let $G = \mathrm{Sp}(\mathbb{C}^{2n}, \Omega)$ and define $\tau(g) = K_{p,q}(g^*)^{-1}K_{p,q}$ for $g \in G$. We leave it as an exercise to prove that τ is a complex conjugation of G. The corresponding real form is the group $\mathrm{Sp}(p, q)$ defined in Section 1.1.4. When $p = n$ we use the notation $\mathrm{Sp}(n) = \mathrm{Sp}(n, 0)$. Since $K_{n,0} = I_{2n}$, it follows that $\mathrm{Sp}(n) = \mathrm{SU}(2n) \cap \mathrm{Sp}(n, \mathbb{C})$. Hence $\mathrm{Sp}(n)$ is a compact real form of $\mathrm{Sp}(n, \mathbb{C})$.

Summary

We have shown that the classical groups (with the condition $\det(g) = 1$ included for conciseness) can be viewed either as

- the complex linear algebraic groups $\mathrm{SL}(n,\mathbb{C})$, $\mathrm{SO}(n,\mathbb{C})$, and $\mathrm{Sp}(n,\mathbb{C})$ together with their real forms, or alternatively as
- the special linear groups over the fields $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{H}$, together with the special isometry groups of non-degenerate forms (symmetric or skew-symmetric, Hermitian or skew-Hermitian) over these fields.

Thus we have the following families of classical groups.

Special linear groups: $\mathrm{SL}(n,\mathbb{R})$, $\mathrm{SL}(n,\mathbb{C})$, and $\mathrm{SL}(n,\mathbb{H})$. Of these, only $\mathrm{SL}(n,\mathbb{C})$ is an algebraic group over $\mathbb{C}$, whereas the other two are real forms of $\mathrm{SL}(n,\mathbb{C})$ (*respectively* $\mathrm{SL}(2n,\mathbb{C})$).

Automorphism groups of forms: On a real vector space, Hermitian and skew Hermitian are the same as symmetric and skew symmetric. On a complex vector space skew-Hermitian forms become Hermitian after multiplication by i, and *vice versa*, whereas on a quaternionic vector space there are no nonzero bilinear forms at all (by the noncommutativity of quaternionic multiplication), so the form must be either Hermitian or skew-Hermitian. Taking these restrictions into account, we see that the possibilities for isometry groups are those listed in Table 1.1.

Table 1.1: Isometry Groups of Forms

Group	Field	Form
$\mathrm{SO}(p,q)$	$\mathbb{R}$	Symmetric
$\mathrm{SO}(n,\mathbb{C})$	$\mathbb{C}$	Symmetric
$\mathrm{Sp}(n,\mathbb{R})$	$\mathbb{R}$	Skew-symmetric
$\mathrm{Sp}(n,\mathbb{C})$	$\mathbb{C}$	Skew-symmetric
$\mathrm{SU}(p,q)$	$\mathbb{C}$	Hermitian
$\mathrm{Sp}(p,q)$	$\mathbb{H}$	Hermitian
$\mathrm{SO}^*(2n)$	$\mathbb{H}$	Skew-Hermitian

Note that even though the field is $\mathbb{C}$, the group $\mathrm{SU}(p,q)$ is *not* an algebraic group over $\mathbb{C}$ (its defining equations involve complex conjugation). Likewise, the groups for the field $\mathbb{H}$ are not algebraic groups over $\mathbb{C}$, even though $\mathbb{C}$ is embedded in $\mathbb{H}$.

1.7.3 Exercises

1. On $G = \mathbb{C}^\times$ define the conjugation $\tau(z) = \bar{z}^{-1}$. Let $V \subset \mathcal{O}[G]$ be the subspace with basis $f_1(z) = z$ and $f_2(z) = z^{-1}$. Define $Cf(z) = \overline{f(\tau(z))}$ and $\rho(z)f(w) = f(wz)$ for $f \in V$ and $z \in G$, as in Theorem 1.7.5.

   (a) Find a basis $\{v_1, v_2\}$ for the real subspace $V_+ = \{f \in V : Cf = f\}$ so that in this basis $\rho(z) = \begin{bmatrix} (z+z^{-1})/2 & (z-z^{-1})/2i \\ -(z-z^{-1})/2i & (z+z^{-1})/2 \end{bmatrix}$.

(b) Let $K = \{z \in G : \tau(z) = z\}$. Use (a) to show that $G \cong \mathrm{SO}(2, \mathbb{C})$ as an algebraic group and $K \cong \mathrm{SO}(2)$ as a Lie group.

2. Show that $\mathrm{Sp}(1)$ is isomorphic with $\mathrm{SU}(2)$. (HINT: Consider the adjoint representation of $\mathrm{Sp}(1)$.)

3. Let ψ be the real linear transformation of $\mathbb{C}^{2n}$ defined by

$$\psi[z_1, \ldots, z_n, z_{n+1}, \ldots, z_{2n}] = [\overline{z}_{n+1}, \ldots, \overline{z}_{2n}, -\overline{z}_1, \ldots, -\overline{z}_n]$$

Define $\mathrm{SU}^*(2n) = \{g \in \mathrm{SL}(2n, \mathbb{C}) : g\psi = \psi g\}$. Show that $\mathrm{SU}^*(2n)$ is isomorphic with $\mathrm{SL}(n, \mathbb{H})$ as a Lie group.

4. Let $G = \mathrm{Sp}(\mathbb{C}^{2n}, \Omega)$ be the group for the real form of Type CII. Show that $g \mapsto (g^*)^{-1}$ defines an involutory automorphism of G as a real Lie group, and that $\tau(g) = K_{p,q}(g^*)^{-1}K_{p,q}$ defines a complex conjugation of G.

5. Let $G = \mathrm{O}(n, \mathbb{C})$ and let $\tau(g) = I_{p,q}\overline{g}I_{p,q}$ be the complex conjugation of Type BDI. Let $H = \{g \in G : \tau(g) = g\}$ be the associated real form. Define $J_{p,q} = \mathrm{diag}[I_p,\ iI_q]$ and set $\gamma(g) = J_{p,q}^{-1} g J_{p,q}$ for $g \in G$.

(a) Prove that $\gamma(\tau(g)) = \overline{\gamma(g)}$ for $g \in G$. Hence $\gamma(H) \subset \mathrm{GL}(n, \mathbb{R})$. (HINT: Note that $J_{p,q}^2 = I_{p,q}$ and $J_{p,q}^{-1} = \overline{J}_{p,q}$.)

(b) Prove that $\gamma(g)^t I_{p,q}\gamma(g) = I_{p,q}$ for $g \in G$. Together with the result from part (a) this shows that $\gamma(H) = \mathrm{O}(p, q)$.

1.8 Notes

Section 1.3 For a more complete introduction to Lie groups through matrix groups, see Rossmann [2002].

Section 1.4 Although Hermann Weyl seemed well aware that there could be a theory of algebraic groups (for example he calculated the ideal of the orthogonal groups in Weyl [1946]), he studied the classical groups as individuals with many similarities rather than as examples of linear algebraic groups. Chevalley considered algebraic groups to be a natural subclass of the class of Lie groups and devoted Volumes II and III of his *Theory of Lie Groups* to the development of their basic properties (Chevalley [1951], [1954]). The modern theory of linear algebraic groups has its genesis in the work of Borel ([1956] and [1991]) – see Borel [2001] for a detailed historical account. Additional books on algebraic groups are Humphreys [1975], Kraft [1985], Springer [1981], and Onishchik and Vinberg [1990].

Section 1.7 Proposition 1.7.7 is the Lie algebra version of Weyl's *unitary trick*. A detailed discussion of real forms of complex semisimple Lie groups and É. Cartan's classification can be found in Helgason [1978]. One can see from Helgason [1978, Ch. X, Table V] that the real forms of the classical groups contain a substantial portion of the connected simple Lie groups. The remaining simple Lie groups are the real forms of the five *exceptional* simple Lie groups (Cartan's types G2, F4, E6, E7, and E8 of dimension 14, 52, 78, 133, and 248 respectively).