

1 The Classical Groups

1.1 The groups

Let \mathbb{F} denote either the real numbers, \mathbb{R} , or the complex numbers, \mathbb{C} . In this section we will describe the main players in the rest of this book the Classical Groups as designated by Hermann Weyl. This section should be treated as a dictionary. The groups as named here will appear throughout the book.

1.1.1 The general linear group.

Let V be a finite dimensional vector space over \mathbb{F} . The set of all invertible transformations of V to V will be denoted $GL(V)$. This set has a group structure under composition of transformations with identity element the identity transformation $Id(x) = x$ for all $x \in V$. We will consider this family of groups in this subsection. We will be recalling some standard terminology related to linear transformations and their matrices.

Let V and W be finite dimensional vector spaces over \mathbb{F} with bases v_1, \dots, v_n and w_1, \dots, w_m respectively. If $T : V \rightarrow W$ is a linear map then we have

$$Tv_j = \sum_{i=1}^m a_{ij}w_i$$

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

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

is the *matrix* of T with respect to the two bases. If $S : W \rightarrow U$ is a linear map with U an l dimensional vector space with basis u_1, \dots, u_l and if B is the matrix of S with respect to the bases w_1, \dots, w_m and u_1, \dots, u_l then the matrix of $S \circ T$ with respect to v_1, \dots, v_n and u_1, \dots, u_l is given by BA the product being the usual product of matrices.

We denote by $M_n(\mathbb{F})$ the space of all $n \times n$ matrices over \mathbb{F} . Let V be an n dimensional vector space over \mathbb{F} with basis v_1, \dots, v_n . If $T : V \rightarrow V$ is a linear map we will (for the moment) write $\mu(T)$ for the matrix of T with respect to this basis. If $T, S \in GL(V)$ then the observations above imply that

$$\mu(S \circ T) = \mu(S)\mu(T).$$

Furthermore, if $T \in GL(V)$ then $\mu(T \circ T^{-1}) = \mu(T^{-1} \circ T) = \mu(Id) = I$ the identity matrix (the entries being $\delta_{ij} = 1$ if $i = j$ and 0 otherwise). A 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 $T : V \rightarrow V$ is in $GL(V)$ if and only

if $\mu(T)$ is invertible. We also recall that a matrix $A \in M_n(\mathbb{F})$ is invertible if and only if its determinant is non-zero.

We will use the notation $GL(n, \mathbb{F})$ for the set of $n \times n$ invertible matrices with coefficients in \mathbb{F} . $GL(n, \mathbb{F})$ is a group under matrix multiplication with identity the identity matrix. We note that if V is an n -dimensional vector space over \mathbb{F} with basis v_1, \dots, v_n then the map $\mu : GL(V) \rightarrow GL(n, \mathbb{F})$ corresponding to this basis is a group isomorphism. The group $GL(n, \mathbb{F})$ is called the *general linear group of rank n* . If w_1, \dots, w_n is another basis of V and if

$$w_j = \sum_i g_{ij} v_i$$

then

$$v_j = \sum_i h_{ij} w_i$$

with $[h_{ij}]$ the inverse matrix to $[g_{ij}]$. If T is a linear transformation of V to V and if $A = [a_{ij}]$ is the matrix of T with respect to v_1, \dots, v_n and if $B = [b_{ij}]$ is the matrix with respect to w_1, \dots, w_n then

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

Thus, if $g = [g_{ij}]$ then $B = g^{-1} A g$.

1.1.2 The special linear group.

The special linear group of $n \times n$ matrices 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 $GL(n, \mathbb{F})$ with will be denoted $SL(n, \mathbb{F})$.

We note that if V is an n -dimensional vector space with basis v_1, \dots, v_n and if $\mu : GL(V) \rightarrow GL(n, \mathbb{F})$ is as above then the set $\{T \in GL(V) \mid \det(\mu(T)) = 1\}$ is independent of the choice of basis. (See the change of basis formula in the previous subsection. You will be called upon to prove this in Exercise 1 at the end of this section). This subset us the subgroup $\mu^{-1}(SL(n, \mathbb{F}))$ and we will denote it by $SL(V)$.

1.1.3 Isometry groups of bilinear forms.

Let V be an n dimensional vector space over \mathbb{F} and let $B : V \times V \rightarrow \mathbb{F}$ be a bilinear map Then we denote by $O(B)$ the set of all $T \in GL(V)$ such that $B(Tv, Tw) = B(v, w)$ for all $v, w \in V$. We note that $O(B)$ is a subgroup of $GL(V)$. Let v_1, \dots, v_n be a basis of V relative to this basis we may write $B(v_i, v_j) = b_{ij}$. Assume that T has matrix $A = [a_{ij}]$ then we have

$$B(Tv_i, Tv_j) = \sum_{k,l} a_{ki} a_{lj} B(v_k, v_l) = \sum_{k,l} a_{ki} a_{lj} b_{kl}.$$

Thus if A^t is the matrix $[c_{ij}]$ with $c_{ij} = a_{ji}$ (the usual transpose of A) then the condition that $T \in O(B)$ is that

$$[b_{ij}] = A^t [b_{ij}] A.$$

If B is non-degenerate ($B(v, w) = 0$ for all w implies $v = 0$ and $B(v, w) = 0$ for all v implies $w = 0$) then we have $\det([b_{ij}]) \neq 0$. Hence using the formula above we see that if B is nondegenerate and if $T : V \rightarrow V$ is linear and satisfies $B(Tv, Tw) = B(v, w)$ for all $v, w \in V$ then $T \in O(B)$. The next two subsections will discuss the most important special cases of this class of groups.

1.1.4 Orthogonal groups.

We will start this section by first introducing the matrix groups and then identifying them with the corresponding isometry groups. Let $O(n, \mathbb{F})$ denote the set of all $g \in GL(n, \mathbb{F})$ such that $gg^t = I$. That is

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

We note that $(AB)^t = B^t A^t$ and if $A, B \in GL(n, \mathbb{F})$ then $(AB)^{-1} = B^{-1} A^{-1}$. It is therefore obvious that $O(n, \mathbb{F})$ is a subgroup of $GL(n, \mathbb{F})$. 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, $O(p, q)$, with $p + q = n$ with $p, q \in \mathbb{N}$. 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

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

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

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

that is the matrix with all entries zero except for ones on the skew diagonal (the $i, n + 1 - i$ entries are 1). Then $\sigma I_{p,q} \sigma^{-1} = \sigma I_{p,q} \sigma = \sigma I_{p,q} \sigma^t = -I_{q,p}$. Thus the map

$$\phi : O(p, q) \rightarrow GL(n, \mathbb{R})$$

given by $\phi(g) = \sigma g \sigma$ defines an isomorphism of $O(p, q)$ onto $O(q, p)$.

We will now describe these groups in terms of bilinear forms let V be an n -dimensional vector space over \mathbb{F} and let B be a symmetric non-degenerate bilinear form over \mathbb{F} (either \mathbb{R} or \mathbb{C}). We have

Proposition 1 If $\mathbb{F} = \mathbb{C}$ then there exists a basis v_1, \dots, v_n of V such that $B(v_i, v_j) = \delta_{ij}$. If $\mathbb{F} = \mathbb{R}$ then there exist $p, q \geq 0$ with $p + q = n$ and a basis v_1, \dots, 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 p, q are the same.

Remark 2 The number $p - q$ in the case when $\mathbb{F} = \mathbb{R}$ is called the signature of the form B . We will also say that (p, q) is the signature of B .

We preface the proof of this result with a simple lemma.

Lemma 3 Let M be a symmetric bilinear form on V . Then if $M(v, v) = 0$ for all $v \in V$ then $M = 0$.

Proof. Using the symmetry and bilinearity we have the identity

$$4B(v, w) = B(v + w, v + w) - B(v - w, v - w)$$

for all $v, w \in V$. This implies the lemma. ■

We will now prove the proposition. We will first show that there exists a basis w_1, \dots, w_n of V such that

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

and

$$B(w_i, w_i) \neq 0$$

by induction on n (such a basis is called an *orthogonal basis with respect to B*). Since B is non-degenerate the previous lemma implies that there exists a vector $w_n \in V$ with $B(w_n, w_n) \neq 0$. If $n = 1$ then this is what we set out to prove. Assume for $n - 1 \geq 1$. Set $V' = \{v \in V \mid B(w_n, v) = 0\}$. If $v \in V$ then $v' = v - \frac{B(v, w_n)}{B(w_n, w_n)} w_n \in V'$. This implies that $V = V' + \mathbb{F}w_n$. So, in particular, $\dim V' = n - 1$. We assert that the form $B' = B|_{V' \times V'}$ is non-degenerate on V' . Indeed if $v \in V'$ satisfies $B(v', w) = 0$ for all $w \in V'$ then since $B(v', w_n) = 0$ we see that $B(v', w) = 0$ for all $w \in V$ so $v' = 0$. We can now apply the inductive hypothesis to find a desired basis w_1, \dots, w_{n-1} for V' . Now it is clear that w_1, \dots, w_n has the desired properties.

If $\mathbb{F} = \mathbb{C}$ we can now complete the proof. Let w_1, \dots, w_n be an orthogonal basis of V with respect to B . Let z_i be a choice of square root of $B(w_i, w_i)$ then if $v_i = \frac{1}{z_i} w_i$ the $B(v_i, v_j) = \delta_{ij}$. At this point the reader should have noticed why we need the p when $\mathbb{F} = \mathbb{R}$.

We now consider the case when $\mathbb{F} = \mathbb{R}$. We rearrange (if necessary) so that $B(w_i, w_i) \geq B(w_{i+1}, w_{i+1})$ for $i = 1, \dots, n - 1$. Let $p = 0$ if $B(w_1, w_1) < 0$. Otherwise let $p = \max\{i \mid B(w_i, w_i) > 0\}$. Then $B(w_i, w_i) < 0$ for $i > p$. Define z_i to be a square root of $B(w_i, w_i)$ for $i \leq p$. Define z_i to be a square root of $-B(w_i, w_i)$ if $i > p$. Then if $v_i = \frac{1}{z_i} w_i$ we have $B(v_i, v_j) = \varepsilon_i \delta_{ij}$. We are left with proving that the p is intrinsic to B .

To complete the proof for $\mathbb{F} = \mathbb{R}$ we will use a terminology that will appear later in this opus.

Definition 4 Let V be a vector space over \mathbb{R} and let M be a symmetric bilinear form on V then M is said to be positive definite if $M(v, v) > 0$ for every $v \in V, v \neq 0$.

With this concept in hand we will complete the proof. Fix a basis v_1, \dots, 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_+ equal to the linear span of $\{v_1, \dots, v_p\}$ and V_- equal to the linear span of $\{v_{p+1}, \dots, v_n\}$. Then $V = V_+ + V_-$ direct sum. Let $\pi : V \rightarrow V_+$ be the projection onto the first factor. We note that $B|_{V_+}$ is positive definite. Let W be a subspace of V such that $B|_W$ is positive definite. Suppose that $w \in W$ and $\pi(w) = 0$. Then $w \in V_-$ thus $w = \sum_{i>p} a_i v_i$. Hence

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

Thus since $B|_W$ has been assumed to be positive definite $w = 0$. This implies that $\pi : W \rightarrow V_+$ is injective. Hence $\dim W \leq \dim V_+ = p$. If we had another basis satisfying the desired conditions yielding p' as above then this argument implies that $p' \leq p$ and $p \leq p'$ so $p = p'$. This completes the proof.

In the case when $\mathbb{F} = \mathbb{R}$ and B is a non-degenerate symmetric bilinear form on a finite dimensional vector space V we will call a basis v_1, \dots, v_n of V such that there exists $n > p > 0$ such that $B(v_i, v_j) = \varepsilon_i \delta_{ij}$ with $\varepsilon_i = 1$ for $i \leq p$ and $\varepsilon_i = -1$ for $i > p$ a pseudo orthonormal basis. We have seen that the number p is an invariant of B .

We now come to the crux of the matter. If B is a non-degenerate symmetric bilinear form on the n -dimensional vector space V over \mathbb{F} and if v_1, \dots, v_n is basis of V then we will use the notation $\mu(g)$ for the matrix of the linear operator g with respect to this basis.

The following result is now a direct consequence of the above and the discussion in subsection 1.1.3.

Proposition 5 Let the notation be as above. If v_1, \dots, v_n is an orthonormal basis for V with respect to B or $-B$ then $\mu : O(B) \rightarrow O(n, \mathbb{F})$ defines a group isomorphism. If $\mathbb{F} = \mathbb{R}$ and if v_1, \dots, v_n is a pseudo-orthonormal basis of V with signature $(p, n - p)$ then $\mu : O(B) \rightarrow O(p, n - p)$ is a group isomorphism.

The special orthogonal group over \mathbb{F} is the subgroup $SO(n, \mathbb{F}) = O(n, \mathbb{F}) \cap SL(n, \mathbb{F})$ of $O(n, \mathbb{F})$. The indefinite special orthogonal groups are the groups $SO(p, q) = O(p, q) \cap SL(p + q, \mathbb{R})$.

1.1.5 The 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

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

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

We will now look at the coordinate free version of these groups. For this we need

Proposition 6 *Let V be an m dimensional vector space over F and let B be a non-degenerate, skew symmetric ($B(v, w) = -B(w, v)$), bilinear form on V . Then m must be even. If $m = 2n$ then there exists a basis v_1, \dots, v_{2n} such that $[B(v_i, v_j)] = J$.*

Proof. Assume that B is non-degenerate and $\dim V = m > 0$. Let v be a non-zero element of V . Then there exists $w \in V$ with $B(v, w) \neq 0$. Replacing w with $\frac{1}{B(v, w)}w$ we may assume that $B(v, w) = 1$. Let $W = \{x \in V \mid B(v, x) = 0, B(w, x) = 0\}$. If $x \in V$ then set $x' = x - B(v, x)w - B(x, w)v$. Then

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

since $B(v, w) = 1$ and $B(v, v) = 0$. Similarly

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

since $B(w, v) = -1, B(w, w) = 0$. Thus $V = U + W$ where U is the span of v and w . We note that $B|_U$ is non-degenerate so $U \cap W = \{0\}$. This implies that $\dim W = m - 2$. We also note that $B|_W$ is non-degenerate. We will leave this to the reader (use the same technique as we used for the orthogonal groups see exercise 3 at the end of this section). We will now prove the first part of the proposition. We have seen that if $\dim V > 0$ then $\dim V \geq 2$. Thus the result is true for $\dim V \leq 2$. Assume that the result is true for $2 \leq \dim V < k$. We will prove the result for $\dim V = k$. If we argue as above we have $V = U + W$ a direct sum decomposition. With $\dim W = \dim V - 2$. Since $B|_W$ is non-degenerate the inductive hypothesis implies that $\dim W$ is even, Hence $\dim V$ is even.

We will now prove the second part of the result by induction on n where $m = 2n$. If $n = 1$ then we take $v_1 = v$ and $v_2 = w$ with v, w as above. Assume the result for $n - 1 \geq 1$. Then set $v_n = v, v_{2n} = w$. Then since $B|_W$ is non-degenerate we have a basis w_1, \dots, w_{2n-2} of W as in the second part of the assertion. Take $v_i = w_i$ and $v_{n+1-i} = w_{n-i}$ for $i \leq n - 1$. The basis v_1, \dots, v_{2n} is a desired basis of V . ■

Now let V be an m dimensional vector space over \mathbb{F} and let B be a non-degenerate, bilinear, skew symmetric form on V . Then we have seen that m is even $m = 2n$. Let v_1, \dots, v_{2n} be a basis of V as in the previous result and let $\mu(g)$ be the matrix of the linear transformation $g : V \rightarrow V$ with respect to this basis.

Proposition 7 *Let the notation be as above. Then $\mu : O(B) \rightarrow Sp(n, \mathbb{F})$ is a group isomorphism.*

1.1.6 The unitary groups.

In this subsection we will look at an important class of subgroups of $GL(n, \mathbb{C})$ the unitary groups and special unitary groups for definite and indefinite Hermitian forms. If $A \in M_n(\mathbb{C})$ then we will use the standard notation for its adjoint matrix

$$A^* = \overline{A}^t$$

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 $U(n) = \{A \in M_n(\mathbb{C}) | A^*A = I\}$. The *special unitary group* is $U(n) \cap SL(n, \mathbb{C})$. Let $I_{p,q}$ be as in 1.1.4. We define the indefinite unitary group of signature (p, q) to be

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

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

We will now consider a coordinate free description of these groups. Let V be an n -dimensional vector space over \mathbb{C} then a bilinear map of $B : V \times V \rightarrow \mathbb{C}$ as vector spaces over \mathbb{R} is said to be a *Hermitian form* if it satisfies the following two conditions:

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

A Hermitian form is said to be nondegenerate for $v \in V$ if $B(v, w) = 0$ for all $w \in V$ then $v = 0$. It is said to be positive definite if $B(v, v) > 0$ if $v \in V$ and $v \neq 0$ (note that if M is a Hermitian form then $M(v, v) \in \mathbb{R}$ for all $v \in V$). We define $U(B)$ to be the group of all elements, g , of $GL(V)$ such that $B(gv, gw) = B(v, w)$ for all $v, w \in V$. $U(B)$ is called the *unitary group* of B .

Proposition 8 *Let V be an n -dimensional vector space over \mathbb{C} and let B be a non-degenerate Hermitian form on V . Then there exists an integer $n \geq p \geq 0$ and a basis v_1, \dots, v_n of V over \mathbb{C} 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 depends on p and not on the choice of basis.

The proof of this Proposition is almost identical to that of Proposition 4 and will be left to the reader (see exercise 5 for a guide to a proof).

If V is an n -dimensional vector space over \mathbb{C} and B is a non-degenerate Hermitian form on V then a basis as in the above proposition will be called a *pseudo-orthonormal basis* (if $p = n$ then it will be called an *orthonormal basis*). 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.

Lemma 9 *Let V be a finite dimensional vector space over \mathbb{C} and let B be a non-degenerate Hermitian form on V of signature (p, q) and let v_1, \dots, v_{p+q} be a pseudo-orthonormal basis of V relative to B , Then if $\mu : GL(V) \rightarrow GL(n, \mathbb{C})$ is the assignment of a linear transformation of V to its matrix then $\mu(g) \in U(p, q)$ for all $g \in U(B)$ and $\mu : U(B) \rightarrow U(p, q)$ defines a group homomorphism.*

1.1.7 The quaternionic general linear group.

We will first recall some basic properties of the quaternions. We consider the vector space, \mathbb{H} , over \mathbb{R} given by

$$\mathbb{H} = \left\{ \begin{bmatrix} x & y \\ -\bar{y} & \bar{x} \end{bmatrix} \mid x, y \in \mathbb{C} \right\}.$$

One checks directly that \mathbb{H} is closed under multiplication in $M_n(\mathbb{C})$, that the identity matrix is in \mathbb{H} and that since

$$\begin{bmatrix} x & y \\ -\bar{y} & \bar{x} \end{bmatrix} \begin{bmatrix} \bar{x} & -y \\ \bar{y} & x \end{bmatrix} = (|x|^2 + |y|^2)I$$

every non-zero 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 $\mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k}$ and define a multiplication so that $\mathbf{1}$ is the identity and

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

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

and extended to \mathbb{H} by linearity. We note that this gives an isomorphic division algebra to the one described if we take

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

We define $GL(n, \mathbb{H})$ to be the group of all invertible $n \times n$ matrices over \mathbb{H} . Here the definition of matrix multiplication is the usual one but one must be careful about the order of multiplication.

If we look at the real vector space \mathbb{H}^n then we can define multiplication by \mathbb{H} by

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

We note that $\mathbf{1} \cdot u = u$ and $(ab) \cdot u = a \cdot (b \cdot u)$. We can therefore think of \mathbb{H}^n as a vector space over \mathbb{H} . We will think of \mathbb{H}^n as $n \times 1$ columns and define

$$Au$$

for $u \in \mathbb{H}^n$ and $A \in GL(n, \mathbb{H})$ by matrix multiplication. Then $A(a \cdot u) = aAu$ so A defines a quaternionic linear map. Now has 3 different structures as a vector space over \mathbb{C} . By looking at the subfields

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

Relative to each of these complex structures $GL(n, \mathbb{H})$ acts by complex linear transformations. We leave it to the reader to prove that the determinant of $A \in GL(n, \mathbb{H})$ as a complex linear transformation with respect to any of these three structures is the same (see exercise 10). We can thus define $SL(n, \mathbb{H})$ to be the elements of determinant 1 with respect to any of these structures. This group is also denoted classically as $SU^*(2n)$.

1.1.8 The quaternionic unitary groups.

If $w \in \mathbb{H}$, with

$$w = \begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix}$$

then we note that $w^*w = ww^* = |a|^2 + |b|^2$. If $X \in M_n(\mathbb{H})$ then set $X^* = [x_{ji}^*]$ if $X = [x_{ij}]$. The indefinite quaternionic unitary groups are the groups

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

We leave it to the reader to prove that this set is indeed a subgroup of $GL(p + q, \mathbb{H})$. One can define quaternionic Hermitian forms and prove a result analogous to Proposition 8 but we will have no need of such material in the rest of this book.

1.1.9 The group $SO^*(2n)$.

If we include covering groups (to be defined later) we will have described up all of the classical groups over \mathbb{C} and \mathbb{R} except for one infinite family that is related to the orthogonal groups. We will describe that family now.

Let

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

as usual with I the $n \times n$ identity matrix. We note that $J^2 = -I_{2n}$ the $2n \times 2n$ identity matrix. Thus the map $GL(2n, \mathbb{C})$ to itself given by $\theta(g) = -JgJ$ defines an automorphism and θ^2 is the identity automorphism. Our last classical group is

$$SO^*(2n) = \{g \in SO(2n, \mathbb{C}) | \theta(\bar{g}) = g\}$$

here, as usual, \bar{g} is the matrix whose entries are the complex conjugates of g .

1.1.10 Exercises.

1. Show that if v_1, \dots, v_n and w_1, \dots, w_n are bases of V , a vector space over \mathbb{F} , and if $T : V \rightarrow V$ is a linear map with matrices A and B respectively relative to these bases then $\det A = \det B$.

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

3. Let V be a vector space over \mathbb{F} and B a skew symmetric or symmetric non-degenerate 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 $W^\perp = \{v \in V \mid B(v, w) = 0 \text{ for all } w \in W\}$ is non-degenerate.

4. Let V denote the vector space of symmetric 2×2 matrices over $\mathbb{F} = \mathbb{R}$ or \mathbb{C} . If $x, y \in V$ define $B(x, y) = \frac{1}{2}(\det(x + y) - \det(x - y))$ show that $B(x, y) = B(y, x)$ for all $x, y \in V$. Show that if $\mathbb{F} = \mathbb{R}$ then the signature of the form B is $(2, 1)$. If $g \in SL(2, \mathbb{F})$ define $\varphi(g) \in GL(V)$ by $\varphi(g)(v) = gvgt$. Then show that $\varphi : SL(2, \mathbb{F}) \rightarrow O(V, B)$ is a group homomorphism with image $SO(V, B)$ and kernel $\{\pm I\}$.

5. The purpose of this exercise is to prove Proposition 8 by the method of proof of Proposition 1. First prove the analogue of Lemma 3. That is, if M is Hermitian form such that $M(v, v) = 0$ for all v then $M = 0$. To do this note that if the condition is satisfied then $M(v + tw, v + tw) = tM(w, v) + \bar{t}M(v, w)$. Now substitute values for t to see that $M(v, w) = 0$. The rest of the arguments are the same as for Proposition 4 once one observes that if M is Hermitian then $M(v, v) \in \mathbb{R}$.

6. Let V be a $2n$ -dimensional vector space over $\mathbb{F} = \mathbb{R}$ or \mathbb{C} . We consider the space $W = \bigwedge^n V$. Fix a basis ω of the one dimensional vector space $\bigwedge^n V$. Consider the bilinear form $B(u, v)$ defined by $u \wedge v = B(u, v)\omega$. Show that the form is nondegenerate and skew symmetric if n is odd and symmetric if n is even. If n is even and $\mathbb{F} = \mathbb{R}$ then determine the signature of B .

7. In the notation of problem 6. we consider the case when $V = \mathbb{F}^4$ and if e_1, e_2, e_3, e_4 is the standard basis then $\omega = e_1 \wedge e_2 \wedge e_3 \wedge e_4$. Consider $\varphi(g) = \bigwedge^2 g$ for $g \in SL(4, \mathbb{F})$. Show that $\varphi : SL(4, \mathbb{F}) \rightarrow O(\bigwedge^2 \mathbb{F}^4, B)$ is a group homomorphism with kernel $\{\pm I\}$. If $\mathbb{F} = \mathbb{C}$ and we choose an orthonormal basis then φ defines a group homomorphism of $SL(4, \mathbb{C})$ to $SO(6, \mathbb{C})$. If $\mathbb{F} = \mathbb{R}$ and we choose a pseudo-orthonormal basis show that φ defines a group homomorphism to $SO(3, 3)$.

8. In the notation of problem 7 let ϕ be the restriction of φ to $Sp(4, \mathbb{F})$. Let $\nu = e_1 \wedge e_3 + e_2 \wedge e_4$. Show that $\phi(g)\nu = \nu$. Show that $B(\nu, \nu) = -2$. Let $W = \{w \in \bigwedge^2 \mathbb{F}^4 \mid B(\nu, w) = 0\}$. Let $\rho(g) = \phi(g)|_W$. If \mathbb{R} show that the signature of the restriction of B to W is $(3, 2)$. Show that ρ is a group homomorphism from $Sp(4, \mathbb{F})$ to $SO(W, B|_W)$ with kernel $\{\pm 1\}$.

9. Let $G = SL(2, \mathbb{F}) \times SL(2, \mathbb{F})$ and define $\varphi : G \rightarrow GL(\mathbb{F}^2 \otimes \mathbb{F}^2)$ be given by $\varphi(a, b) = a \otimes b$. Let B be the bilinear form on $\mathbb{F}^2 \otimes \mathbb{F}^2$ given by $B(e_i \otimes e_j, e_k \otimes e_l) = \varepsilon_{ik} \varepsilon_{jl}$ with $\varepsilon_{ij} = -\varepsilon_{ji}$ for $i, j = 1, 2$ and $\varepsilon_{12} = 1$. Show that B defines a symmetric nondegenerate form on $\mathbb{F}^2 \otimes \mathbb{F}^2$ and calculate the signature of B if $\mathbb{F} = \mathbb{R}$. Show that if $g \in G$ then $\varphi(g) \in SO(\mathbb{F}^2 \otimes \mathbb{F}^2, B)$. What is the kernel of

φ ?

10. Show that if $X \in M_n(\mathbb{H})$ and we consider the three versions of \mathbb{C} in the quaternions as in 1.1.8. These give 3 ways of thinking of X as a $2n \times 2n$ matrix over \mathbb{C} . Show that the determinants of these three matrices are equal.

1.2 The Lie algebras.

Let V be a vector space over \mathbb{F} . Let $End(V)$ denote the algebra (under composition) of \mathbb{F} -linear maps of V to V . If $X, Y \in End(V)$ then we set $[X, Y] = XY - YX$. This defines a new product on $End(V)$ that satisfies two properties (see exercise 1 at the end of this section) :

1. $[X, Y] = -[Y, X]$ for all X, Y (skew symmetry).
2. For all X, Y, Z we have $[X, [Y, Z]] = [[X, Y], Z] + [Y, [X, Z]]$.

The latter condition is a substitute for the associative rule.

Definition 10 A pair of a vector space, \mathfrak{g} , over \mathbb{F} and a bilinear binary operation $[\dots, \dots]$ is said to be a Lie algebra if for each $X, Y, Z \in \mathfrak{g}$ the conditions 1. and 2. are satisfied.

Thus, in particular, we see that $End(V)$ is a Lie algebra under the binary operation $[X, Y] = XY - YX$.

We also note the obvious fact that 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 $[\dots, \dots]$. We will call \mathfrak{h} a sub-Lie algebra or a subalgebra.

Suppose that \mathfrak{g} and \mathfrak{h} are Lie algebras over \mathbb{F} then a Lie algebra homomorphism of \mathfrak{g} to \mathfrak{h} is an \mathbb{F} -linear map $T : \mathfrak{g} \rightarrow \mathfrak{h}$ such that $T[X, Y] = [TX, TY]$. We will say that a Lie algebra homomorphism is an isomorphism if it is bijective (you will be asked in Exercise 2, at the end of this section, to prove that the inverse of a bijective Lie algebra homomorphism is a Lie algebra homomorphism).

1.2.1 The general linear and special linear Lie algebras.

Let V be as in the previous section. We will use the notation $\mathfrak{gl}(V)$ for $End(V)$ looked upon as a Lie algebra under $[X, Y] = XY - YX$. We note that if we choose a basis for V and if V is n -dimensional then the matrices of the elements of $End(V)$ form a Lie algebra which we will denote by $\mathfrak{gl}(n, \mathbb{F})$.

If V is a finite dimensional vector space over \mathbb{F} and if v_1, \dots, v_n is a basis then each $T \in End(V)$ has a matrix, A , with respect to this basis. If $A \in M_n(\mathbb{F})$ and $A = [a_{ij}]$ then $tr(A) = \sum_i a_{ii}$. We note that

$$tr(AB) = tr(BA).$$

This implies that if A is the matrix of T with respect to v_1, \dots, v_n then $tr(A)$ is independent of the choice of basis. We will write $tr(T) = tr(A)$. This implies that we can define

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

Choosing a basis we may look upon this Lie algebra as

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

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

1.2.2 The Lie algebras associated with binary forms.

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

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

We leave it to the reader to check that $\mathfrak{so}(B)$ is a Lie subalgebra of $\mathfrak{gl}(V)$ by the obvious calculation (see Exercise 3 at the end of this section). If B is nondegenerate you will also be asked to prove in Exercise 3 that $\text{tr} X = 0$ for $X \in \mathfrak{so}(B)$. For this the following discussion should be useful.

Let the notation be as in section 1.1.3. That is we assume that V is finite dimensional and we choose a basis v_1, \dots, v_n of V . Then $b_{ij} = B(v_i, v_j)$. By a calculation completely analogous to that in 1.1.3 we see that if $T \in \mathfrak{so}(B)$ then its matrix, A , relative to v_1, \dots, v_n satisfies

$$A^t[b_{ij}] + [b_{ij}]A = 0.$$

1.2.3 The orthogonal Lie algebras.

We define

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

Since $\text{tr} X^t = \text{tr} X$ this implies that $\mathfrak{so}(n, \mathbb{F})$ is a subspace of $\mathfrak{sl}(n, \mathbb{F})$. It is also easily checked (by direct calculation) that $\mathfrak{so}(n, \mathbb{F})$ is a Lie subalgebra of $\mathfrak{sl}(n, \mathbb{F})$.

By analogy with the definition of $O(p, q)$ let $p, q \geq 0$ be integers such that $p + q = n$. Set (as in 1.1.4)

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

We set

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

As observed in the previous section $\mathfrak{so}(p, q) \subset \mathfrak{sl}(n, \mathbb{R})$ and it is easily checked to be a Lie subalgebra.

We can now apply the material in section 1.1.4. Let B denote a nondegenerate symmetric bilinear form on V a finite dimensional vector space over \mathbb{F} . If $\mathbb{F} = \mathbb{C}$ then we have seen in Proposition 1 (1.1.4) that there is a basis v_1, \dots, v_n of V such that $B(v_i, v_j) = \delta_{ij}$. If $\mu(T)$ is the matrix of $T \in \text{End}(V)$ relative to this matrix then μ defines a Lie algebra isomorphism of $\mathfrak{so}(B)$ onto $\mathfrak{so}(n, \mathbb{C})$. If $\mathbb{F} = \mathbb{R}$ and B has signature p, q and v_1, \dots, v_n is a corresponding pseudo-orthonormal basis then the analog of μ defines a Lie algebra isomorphism of $\mathfrak{so}(B)$ onto $\mathfrak{so}(p, q)$.

1.2.4 The symplectic Lie algebra.

In this case we take

$$J = \begin{bmatrix} 0 & I \\ -1 & 0 \end{bmatrix}$$

with I the $n \times n$ identity matrix. We define

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

Then this subspace of the $n \times n$ matrices is a Lie subalgebra which we call the *symplectic Lie algebra of rank n* .

Let V be a vector space over \mathbb{F} of dimension $m < \infty$ and B a skew symmetric bilinear form on V . Then m is even $m = 2n$ and applying Proposition 6 section 1.5 we see that there exists a basis v_1, \dots, v_{2n} of V such that $[B(v_i, v_j)] = J$. The map that assigns to an endomorphism of V its matrix relative to this basis defines an isomorphism of $\mathfrak{so}(B)$ onto $\mathfrak{sp}(n, \mathbb{F})$.

1.2.5 The unitary Lie algebras.

We will use the notation and results of section 1.1.6. Let p, q be integers with $p, q \geq 0$. We define

$$\mathfrak{u}(p, q) = \{X \in M_{p+q}(\mathbb{C}) \mid X^* I_{p,q} + I_{p,q} X = 0\}.$$

We note that this space is a *real* subspace of $M_{p+q}(\mathbb{C})$. One checks directly that it is a Lie subalgebra of $M_{p+q}(\mathbb{C})$ thought of as a Lie algebra over \mathbb{R} .

We define $\mathfrak{su}(p, q) = \mathfrak{u}(p, q) \cap \mathfrak{sl}(p+q, \mathbb{C})$.

Let $n = p + q$, let V be an n -dimensional vector space over \mathbb{C} and let B be a non-degenerate Hermitian form on V . Then we define

$$\mathfrak{u}(B) = \{T \in \text{End}_{\mathbb{C}}(V) \mid B(Tv, w) + B(v, Tw) = 0, v, w \in V\}.$$

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

1.2.6 The quaternion general linear Lie algebra.

In this subsection we will follow the notation of section 1.1.7. The Lie algebra in question consists of the $n \times n$ matrices over the quaternions, \mathbb{H} , with the usual matrix commutator. We will denote the Lie algebra $\mathfrak{gl}(n, \mathbb{H})$. This Lie algebra is to be thought of as a Lie algebra over \mathbb{R} (we have not defined Lie algebras over skew-fields). If we consider \mathbb{H}^n to be \mathbb{C}^{2n} relative to any one of the isomorphic copies of \mathbb{C} : $\mathbb{R}\mathbf{1} + \mathbb{R}\mathbf{i}$, $\mathbb{R}\mathbf{1} + \mathbb{R}\mathbf{j}$, $\mathbb{R}\mathbf{1} + \mathbb{R}\mathbf{k}$. Then we can take $\mathfrak{sl}(n, \mathbb{H})$ to be the elements of trace zero in any of these structures. This real Lie algebra is usually denoted $\mathfrak{su}^*(2n)$.

1.2.7 The quaternionic unitary Lie algebras.

Here we will use the notation of section 1.1.8 (especially the quaternionic adjoint). Then we define for $n = p + q$ with p, q non-negative integers

$$sp(p, q) = \{X \in \mathfrak{gl}(n, \mathbb{H}) \mid X^* I_{p,q} + I_{p,q} X = 0\}.$$

1.2.8 The Lie algebra $\mathfrak{so}^*(2n)$.

Here we will be using the notation of section 1.1.9 especially the definition of θ . The Lie algebra in question is defined to be space of all $X \in \mathfrak{so}(2n, \mathbb{C})$ such that $\theta(\overline{X}) = X$. This real vector subspace of $\mathfrak{so}(2n, \mathbb{C})$ defines a Lie subalgebra if we consider $\mathfrak{so}(2n, \mathbb{C})$ to be a Lie algebra over \mathbb{R} .

1.2.9 The list.

The Lie algebras described in the preceding sections constitute the list of *classical Lie algebras* over \mathbb{R} and \mathbb{C} . These Lie algebras will be the main subject of study throughout the remainder of this book. We will find that the choices of bases for the matrix forms of these Lie algebras are not the most convenient. An observant reader should have noticed by now that the Lie algebras that have a name that is "frakturized" form of a name in the previous section can be obtained by "differentiating" the definition as a group. That is we look at the rule for inclusion in the corresponding group, G . Let us call it \mathcal{R} . We will denote the Lie algebra with the "frakturized" designation \mathfrak{g} . Consider a curve $\sigma : (-\varepsilon, \varepsilon) \rightarrow GL(V)$ such that $\sigma(0) = I$ and σ is differentiable at 0. Then one should observe that if $\mathcal{R}(\sigma(t))$ is satisfied for all $t \in (-\varepsilon, \varepsilon)$ then $\sigma'(0) \in \mathfrak{g}$. For example, if G is the subgroup $O(B)$ of $GL(V)$ defined by a binary form B (i.e. $G = \{g \in GL(V) \mid B(gv, gw) = B(v, w) \text{ for all } v, w \in V\}$). Then the indicated conditions on the curve are $B(\sigma(t)v, \sigma(t)w) = B(v, w)$ for all $v, w \in V$ and $t \in (-\varepsilon, \varepsilon)$. If we differentiate these relations we have

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

Since $\sigma(0) = I$ we see that $\sigma'(0) \in \mathfrak{so}(B)$ as asserted. This is the reason why these Lie algebras are usually called the *infinitesimal form* of these groups.

1.2.10 Exercises

1. Prove the identities 1) and 2) in the beginning of this section.
2. Prove that the inverse of a bijective Lie algebra homomorphism is a Lie algebra homomorphism.
3. Prove that B is a bilinear form on V then $\mathfrak{so}(B)$ is a Lie subalgebra of $\mathfrak{gl}(V)$. Also show that if B is non-degenerate then $tr X = 0$ for all $X \in \mathfrak{so}(B)$.
4. Prove that the indicated spaces that were not proved to be Lie algebras in each of the previous sections is indeed a Lie subalgebra of the indicated general linear Lie algebra over \mathbb{R} or \mathbb{C} .

5. Let $X \in M_n(\mathbb{H})$ for each of the choices of a copy of \mathbb{C} in \mathbb{H} write out the corresponding matrix of X as an element of $M_{2n}(\mathbb{C})$. Use this formula to observe that the trace condition is indeed independent of the choice.

6. Show that $\mathfrak{sp}(p, q) \subset \mathfrak{sl}(p + q, \mathbb{H})$.

7. Carry out the infinitesimal calculation indicated in section 1.2.9 for each of the examples. For the case of $SL(n, \mathbb{F})$ you will need to prove that if $\sigma : (-\varepsilon, \varepsilon) \rightarrow GL(n, \mathbb{F})$ is differentiable and $\sigma(0) = I$ then

$$\frac{d}{dt} \det(\sigma(t))|_{t=0} = \text{tr}(\sigma'(0)).$$

1.3 Closed subgroups of $GL(n, \mathbb{R})$.

In this section we will give an introduction to some Lie theoretic ideas that will motivate the later developments in this book. We will begin with the definition of a topological group and then emphasize the topological groups that are closed subgroups of $GL(n, \mathbb{F})$. Our main tool will be the exponential map which we will deal with explicitly.

1.3.1 Topological Groups.

Let G be a group with a topology such that the maps

$$G \times G \rightarrow G, g, h \mapsto gh$$

and

$$G \rightarrow G, g \mapsto g^{-1}$$

are continuous is called a *topological group*. Here the set $G \times G$ is given the product topology above. We note that $GL(n, \mathbb{F})$ is a topological group when endowed with the topology of an open subset of $M_n(\mathbb{F})$ looked upon as \mathbb{F}^{n^2} by laying out the matrix entries as, say, $(x_{11}, x_{12}, \dots, x_{n1}, x_{21}, \dots, x_{2n}, \dots, x_{n1}, \dots, x_{nn})$. The multiplication is continuous and Cramer's rule implies that the inverse is continuous.

We note that if G is a topological group then we have for each $g \in G$ maps $L_g : G \rightarrow G$ and $R_g : G \rightarrow G$ 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 topological space then H is also a topological group called a topological subgroup. Thus all of the examples in section 1 are topological groups.

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.

In this section we will be studying closed subgroups of $GL(N, \mathbb{R})$ for all $N = 1, 2, \dots$. We first show how we will be looking at $GL(n, \mathbb{F})$ for $\mathbb{F} = \mathbb{C}$ or \mathbb{H} as

a closed subgroup of $GL(dn, \mathbb{R})$ with $d = \dim_{\mathbb{R}} \mathbb{F} = 2, 4$ for \mathbb{C} and \mathbb{H} respectively. If we consider \mathbb{C}^n to be \mathbb{R}^{2n} and multiplication by i to be the action of the matrix

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

with I the $n \times n$ identity matrix then $M_n(\mathbb{C})$ consists of the matrices in $M_{2n}(\mathbb{R})$ that commute with J and $GL(n, \mathbb{C})$ consists of the elements of $GL(2n, \mathbb{R})$ that commute with J . The case of the quaternionic groups is handled similarly where we look at the standard n -dimensional vector space over the quaternions, \mathbb{H}^n , as a \mathbb{R}^{4n} with three operators J_1, J_2, J_3 such that $J_k J_l = -J_l J_k$ for $k \neq l$, $J_i^2 = -I$ and $J_1 J_2 = J_3, J_1 J_3 = -J_2, J_2 J_3 = J_1$. Given as follows

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

Then $M_n(\mathbb{H})$ is the set of $X \in M_{4n}(\mathbb{R})$ such that $XJ_i = J_i X$ for $i = 1, 2, 3$. Furthermore, $GL(n, \mathbb{H})$ is with this identification the subgroup of all elements of $GL(4n, \mathbb{R})$ that commute with J_1, J_2 and J_3 .

Before we study our main examples we will prove a few general results about topological groups.

Lemma 11 *Let H be 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 since L_g is a homeomorphism $gH = L_g(H)$ is open. Hence every left coset is open. Since H the complement of the left cosets other than the identity coset is H , we see that the complement of H is open. Thus H is closed. ■

Lemma 12 *Let G be a topological group then the identity component of G (that is the connected subgroup that contains the identity element, e) is a normal subgroup.*

Proof. Let X be the identity component of G . If $g \in X$ then since $e \in X$ we see that $g \in L_g X$. Since L_g is a homeomorphism we see that $L_g X = X$. Hence X is closed under multiplication. Since $e \in L_g X$ we see that $g^{-1} \in X$. So X is a subgroup. If $g \in G$ then define $\tau(g)(x) = gxg^{-1}$ for $x \in G$. Then $\tau(g)$ is a homeomorphism of G and $\tau(g)(e) = e$. Thus $\tau(g)$ maps the identity component of G to itself. Hence X is a normal subgroup. ■

1.3.2 The exponential map.

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

$$\|X\| = \langle X, X \rangle^{\frac{1}{2}} = \left(\sum_{ij} x_{ij}^2 \right)^{\frac{1}{2}}$$

has the following properties.

1) It is a norm that is $\|X + Y\| \leq \|X\| + \|Y\|$, $\|cX\| = |c| \|X\|$ and $\|X\| = 0$ if and only if $X = 0$.

2) $\|XY\| \leq \|X\| \|Y\|$.

We will now show how to derive 2) from the usual Cauchy-Schwarz inequality.

$$\|XY\|^2 = \sum_{ij} \left(\sum_k x_{ik} y_{kj} \right)^2.$$

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

$$\|XY\|^2 \leq \sum_{ij} \left(\sum_k x_{ik}^2 \right) \left(\sum_k y_{kj}^2 \right) = \left(\sum_{ik} x_{ik}^2 \right) \left(\sum_{kj} y_{kj}^2 \right) = \|X\|^2 \|Y\|^2.$$

Taking the square root of both sides completes the proof.

If $A \in M_n(\mathbb{R})$ and $r > 0$ we will use the notation $B_r(A) = \{X \in M_n(\mathbb{R}) \mid \|X - A\| < r\}$.

We can use this norm in order to define some matrix valued analytic functions. We note that if we have a power series $\sum_{m=0}^{\infty} a_m z^m$ and if $r > 0$ is such that $\sum_{m=0}^{\infty} |a_m| r^m < \infty$ then the given series defines a function $f(z)$ on the disc of radius r . We note that for the same series if $k \geq l$ and $\|X\| < r$ then

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

This goes to 0 as $l \rightarrow \infty$. Thus the series $\sum_{m \geq 0} a_m X^m$ defines a function from the open ball of radius r in $M_n(\mathbb{R})$ to $M_n(\mathbb{R})$. The absolute convergence of this series implies that we can interchange the order of summation without changing the value. In this section we will be concentrating on two functions. The first is defined by the power series

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

and the second

$$\log(1+x) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n}.$$

The first converges on every disc and for the other the best choice is $r = 1$. We note that we have the formal identities

$$\begin{aligned} \exp(x+y) &= \exp(x) \exp(y), \\ \log(1 + (\exp(x) - 1)) &= x. \end{aligned}$$

We have therefore defined two matrix valued functions

$$\exp : M_n(\mathbb{R}) \rightarrow M_n(\mathbb{R})$$

and

$$\log : \{X \in M_n(\mathbb{R}) \mid \|X - I\| < 1\} \rightarrow M_n(\mathbb{R}).$$

Both of these functions are given by power series so are real analytic functions. We note that

$$\exp(0) = I$$

and

$$\exp(X + Y) = \exp(X) \exp(Y)$$

due to the formal identity. This implies that $\exp(X) \exp(-X) = I$ thus

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

Since \exp is, in particular, continuous there exists $r > 0$ such that if $X \in B_r(0) = \{X \in M_n(\mathbb{R}) \mid \|X\| < r\}$ then $\exp(X) \in B_1(I)$. We also note that the formal identity relating \exp and \log implies that

$$\log(I + (\exp X - I)) = X$$

for $X \in B_r(0)$.

We calculate the differential of \exp at 0 we we expand the series to order 1 and have

$$\exp(tv) = I + tv + O(t^2).$$

This implies that

$$d\exp_0(v) = v.$$

The inverse function theorem implies that there exists a ball $B_s(0)$ with $s > 0$ and $s \leq r$ (defined above) such that $\exp(B_s(0)) = U$ is open in $GL(n, \mathbb{R})$ and

$$\exp : B_s(0) \rightarrow U$$

is a homeomorphism with analytic inverse. We note that $U \subset B_1(I)$ so, in fact,

$$\left(\exp|_{B_s(0)}\right)^{-1} = \log|_U.$$

To give an idea of the power of what we have done so far we will prove

Theorem 13 *Let $\phi : \mathbb{R} \rightarrow GL(n, \mathbb{R})$ be a continuous group homomorphism of the additive group of real numbers then there exists $X \in M_n(\mathbb{R})$ such the $\phi(t) = \exp(tX)$ for all $t \in \mathbb{R}$. Furthermore, X is uniquely determined by ϕ .*

Proof. If we set $\phi_\varepsilon(t) = \phi(\varepsilon t)$ then ϕ_ε is also a continuous homomorphism of \mathbb{R} into $GL(n, \mathbb{R})$. Since ϕ is continuous we may assume that $\phi(t) \in \exp(B_{\frac{\varepsilon}{2}}(0))$ for $|t| \leq 2$. Thus $\phi(1) = \exp X$ for some $X \in B_{\frac{\varepsilon}{2}}(0)$. We also note that $\phi(\frac{1}{2}) = \exp Z$ with $Z \in B_{\frac{\varepsilon}{2}}(0)$. Thus $\phi(1) = \exp(2Z)$. Since $\exp : B_\varepsilon(0) \rightarrow U$ is injective this implies that $2Z = X$ so $Z = \frac{1}{2}X$. If we do the same argument with $\phi(\frac{1}{4}) = \exp(W)$ with $W \in B_{\frac{\varepsilon}{2}}(0)$ we have $\exp(2W) = \exp Z$ and since both $2W$ and Z are in $B_\varepsilon(0)$ we have $W = \frac{1}{2}Z = \frac{1}{4}X$. We can continue this argument and find that

$$\phi\left(\frac{1}{2^k}\right) = \exp\left(\frac{1}{2^k}X\right).$$

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

$$\begin{aligned} \phi(a) &= \lim_{k \rightarrow \infty} \phi\left(\frac{a_1}{2} + \frac{a_2}{2^2} + \dots + \frac{a_k}{2^k}\right) = \lim_{k \rightarrow \infty} \phi\left(\frac{1}{2}\right)^{a_1} \phi\left(\frac{1}{4}\right)^{a_2} \dots \phi\left(\frac{1}{2^k}\right)^{a_k} \\ &= \lim_{k \rightarrow \infty} \exp\left(\frac{1}{2}X\right)^{a_1} \dots \exp\left(\frac{1}{2^k}X\right)^{a_k} \\ &= \lim_{k \rightarrow \infty} \exp\left(\left(\frac{a_1}{2} + \frac{a_2}{2^2} + \dots + \frac{a_k}{2^k}\right)X\right) = \exp(aX). \end{aligned}$$

Now if $0 \leq a < 1$ then $\phi(-a) = \phi(a)^{-1} = \exp(aX)^{-1} = \exp(-aX)$. Finally, if $a \in \mathbb{R}$ there exists $k > 0$ with k an integer such that $|\frac{a}{k}| < 1$. Then $\phi(a) = \phi\left(\frac{a}{k}\right)^k = \exp\left(\frac{a}{k}X\right)^k = \exp(aX)$.

If $\exp(tX) = \exp(tY)$ for all t then $0 = \frac{\exp(tX) - \exp(tY)}{t} = X - Y + O(t)$ as $t \rightarrow 0$. Thus $X = Y$. ■

1.3.3 The Lie algebra of a closed subgroup of $GL(n, \mathbb{R})$.

Let G be a closed subgroup of $GL(n, \mathbb{R})$ then we will denote the set $\{X \in M_n(\mathbb{R}) \mid \exp(tX) \in G, t \in \mathbb{R}\}$ by $Lie(G)$. We will show that this set is a Lie subalgebra of $M_n(\mathbb{R})$ relative to the bracket $[X, Y] = XY - YX$. This will be the first of four apparently different definitions of Lie algebra of a class of groups that will be used in this book.

For this we need more information about the exponential and logarithm. We consider

$$\log(I + (\exp tX \exp tY - I))$$

we will expand the series in t to second order. We have

$$\begin{aligned}
& \log(I + (\exp tX \exp tY - I)) \\
&= \log\left(\left(I + tX + \frac{t^2}{2}X^2 + O(t^3)\right)\left(I + tY + \frac{t^2}{2}Y^2 + O(t^3)\right)\right) \\
&= \log\left(I + tX + tY + t^2XY + \frac{t^2}{2}X^2 + \frac{t^2}{2}Y^2 + O(t^3)\right) \\
&= t(X + Y) + t^2XY + \frac{t^2}{2}X^2 + \frac{t^2}{2}Y^2 - \frac{1}{2}(tX + tY + t^2XY + \frac{t^2}{2}X^2 + \frac{t^2}{2}Y^2)^2 + O(t^3) \\
&= t(X + Y) + t^2XY + \frac{t^2}{2}X^2 + \frac{t^2}{2}Y^2 - \frac{1}{2}(tX + tY)^2 + O(t^3)
\end{aligned}$$

since all the rest of the terms in the last expression involve at least t^3 . This last expression is

$$\begin{aligned}
& tX + tY + t^2XY + \frac{t^2}{2}X^2 + \frac{t^2}{2}Y^2 - \frac{1}{2}t^2(X^2 + XY + YX + Y^2) + O(t^3) \\
&= tX + tY + \frac{t^2}{2}[X, Y] + O(t^3).
\end{aligned}$$

If we do exactly the same argument with

$$\log(I + (\exp sX \exp tY - I))$$

neglecting any terms of order t^2 or s^2 then we have

$$\begin{aligned}
& \log(I + sX + tY + stXY + O(s^2) + O(t^2)) \\
&= sX + tY + stXY - \frac{1}{2}(sX + tY + stXY)^2 + O(s^2) + O(t^2) \\
&= sX + tY + stXY - \frac{1}{2}(stXY + stYX) + O(t^2) + O(s^2) \\
&= sX + tY + \frac{st}{2}[X, Y] + O(t^2) + O(s^2).
\end{aligned}$$

Lemma 14 *As $t \rightarrow 0$ we have*

$$\exp tX \exp tY = \exp\left(tX + tY + \frac{t^2}{2}[X, Y] + O(t^3)\right)$$

and as $t \rightarrow 0, s \rightarrow 0$ we have

$$\exp sX \exp tY = \exp\left(sX + tY + \frac{ts}{2}[X, Y] + O(t^2) + O(s^2)\right).$$

This result implies a few (standard) limit identities

Lemma 15 *Let $X, Y \in M_n(\mathbb{R})$ then*

- 1) $\lim_{k \rightarrow \infty} (\exp(\frac{1}{k}X) \exp(\frac{1}{k}Y))^k = \exp(X + Y)$.
- 2) $\lim_{k \rightarrow \infty} (\exp(\frac{1}{k}X) \exp(\frac{1}{k}Y) \exp(-\frac{1}{k}X) \exp(-\frac{1}{k}Y))^{k^2} = \exp([X, Y])$.

Proof. We have (using the first asymptotic result in the previous lemma

$$\left(\exp\left(\frac{1}{k}X\right)\exp\left(\frac{1}{k}Y\right)\right)^k = \exp\left(\frac{1}{k}X + \frac{1}{k}Y + O\left(\frac{1}{k^2}\right)\right)^k = \exp\left(X + Y + kO\left(\frac{1}{k^2}\right)\right).$$

This implies the first limit formula. As for the second we use the same asymptotic information and get

$$\left(\exp\left(\frac{1}{k}X\right)\exp\left(\frac{1}{k}Y\right)\right) = \exp\left(\frac{1}{k}(X + Y) + \frac{1}{2k^2}[X, Y] + O\left(\frac{1}{k^3}\right)\right)$$

hence

$$\begin{aligned} & \left(\exp\left(\frac{1}{k}X\right)\exp\left(\frac{1}{k}Y\right)\right)\left(\exp\left(-\frac{1}{k}X\right)\exp\left(-\frac{1}{k}Y\right)\right) \\ &= \exp\left(\frac{1}{k}(X + Y) + \frac{1}{2k^2}[X, Y] + O\left(\frac{1}{k^3}\right)\right)\exp\left(-\frac{1}{k}(X + Y) + \frac{1}{2k^2}[X, Y] + O\left(\frac{1}{k^3}\right)\right) \\ &= \exp\left(\frac{1}{k^2}[X, Y] + O\left(\frac{1}{k^3}\right)\right). \end{aligned}$$

We therefore have

$$\left(\exp\left(\frac{1}{k}X\right)\exp\left(\frac{1}{k}Y\right)\right)\left(\exp\left(-\frac{1}{k}X\right)\exp\left(-\frac{1}{k}Y\right)\right)^{k^2} = \left(\exp\left(\frac{1}{k^2}[X, Y] + O\left(\frac{1}{k^3}\right)\right)\right)^{k^2}.$$

This implies the second limit formula. ■

This result implies

Proposition 16 *If G is a closed subgroup of $GL(n, \mathbb{R})$ then $Lie(G)$ is a Lie subalgebra of $M_n(\mathbb{R})$.*

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

$$\exp(t(X + Y)) = \lim_{k \rightarrow \infty} \left(\exp\left(\frac{t}{k}X\right)\exp\left(\frac{t}{k}Y\right)\right)^k$$

is in G since G is a closed subgroup. Similarly

$$\exp(t[X, Y]) = \lim_{k \rightarrow \infty} \left(\exp\left(\frac{1}{k}X\right)\exp\left(\frac{1}{k}Y\right)\right)\left(\exp\left(-\frac{1}{k}X\right)\exp\left(-\frac{1}{k}Y\right)\right)^{k^2}$$

which is in G . ■

We will use the identification of the groups $GL(n, \mathbb{C})$ and $GL(n, \mathbb{H})$ with the corresponding subgroups of $GL(2n, \mathbb{R})$ and $GL(4n, \mathbb{R})$ respectively as in the end of subsection 1.3.1 and use the notation therein. With these identifications the Lie algebras in section 1.2 line up perfectly with the groups in section 1.1 in the sense that one replaces several of the capital letters in the group name with fraktur letters. We will do a few examples and leave the rest as exercises.

That $Lie(GL(n, \mathbb{R})) = M_n(\mathbb{R}) = \mathfrak{gl}(n, \mathbb{R})$ is obvious. We now look at $Lie(GL(n, \mathbb{C}))$. This Lie algebra is the space of all $X \in M_{2n}(\mathbb{R})$ such that $\exp(tX)J = J\exp(tX)$. This says that $J^{-1}\exp(tX)J = \exp(tX)$. Since $A^{-1}\exp(X)A = \exp(A^{-1}XA)$ we see that $X \in Lie(GL(n, \mathbb{C}))$ if and only

if $\exp(tJ^{-1}XJ) = \exp(tX)$ for all $t \in \mathbb{R}$. But the uniqueness in Theorem 11 now implies that $J^{-1}XJ = X$ so $X \in M_n(\mathbb{C}) = \mathfrak{gl}(n, \mathbb{C})$. As above $Lie(GL(n, \mathbb{H})) = \mathfrak{gl}(n, \mathbb{H})$.

We now look at $SL(n, \mathbb{R})$. We will first calculate $\det(\exp(tX))$ for $X \in M_n(\mathbb{R})$ this calculation uses a method that we will use in general in the next subsection. We note that we can apply Theorem 11 to the case of $n = 1$. Thus we have $\det(\exp(tX)) = \exp(t\mu(X))$ with $\mu(X) \in \mathbb{R}$. Now $\exp(t\mu(X)) = 1 + t\mu(X) + O(t^2)$. We will now give an expansion of the other side of the equation (here $I = [\delta_{ij}]$)

$$\begin{aligned} \det(\exp(tX)) &= \det(I + tX + O(t^2)) = \\ \det(I + tX) + O(t^2) &= \sum_{\sigma \in S_n} \text{sgn}(\sigma)(\delta_{\sigma 1,1} + tx_{\sigma 1,1})(\delta_{\sigma 2,2} + tx_{\sigma 2,2}) \cdots (\delta_{\sigma n,n} + tx_{\sigma n,n}) \end{aligned}$$

We note that if $\sigma i \neq i$ for some i there must be a $j \neq i$ with $\sigma j \pm j$ thus if σ is not the identity permutation then the corresponding term in the sum is of order at least t^2 . Thus the only terms contributing to order 0 and 1 come from the identity element. That term is $(1 + tx_{11}) \cdots (1 + tx_{nn}) = 1 + t(x_{11} + \dots + x_{nn}) + O(t^2)$. We conclude that $\mu(X) = \text{tr}X$. We therefore see that $Lie(SL(n, \mathbb{R})) = \{X \in M_n(\mathbb{R}) | \text{tr}X = 0\} = \mathfrak{sl}(n, \mathbb{R})$.

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

Lemma 17 *If $H \subset G \subset GL(n, \mathbb{R})$ are such that H is a closed subgroup of G and G is a closed subgroup of $GL(n, \mathbb{R})$ then H is a closed subgroup of $GL(n, \mathbb{R})$ and $Lie(H) = \{X \in Lie(G) | \exp(tX) \in H, t \in \mathbb{R}\}$.*

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

We consider $Lie(Sp(n, \mathbb{C}))$. We can consider $Sp(n, \mathbb{C}) \subset GL(2n, \mathbb{C}) \subset GL(2n, \mathbb{R})$. We can thus look upon $Lie(Sp(n, \mathbb{C}))$ as the set of $X \in M_n(\mathbb{C})$ such that $\exp tX \in Sp(n, \mathbb{C})$ for all $t \in \mathbb{R}$. The remarks in subsection 1.2.9 complete the argument to show that $Lie(Sp(n, \mathbb{C})) = \mathfrak{sp}(n, \mathbb{C})$.

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

$$Lie(G) = \{X \in M_n(\mathbb{C}) | \exp(tX)^* I_{p,q} \exp(tX) = I_{p,q}\}.$$

We note that $\exp(tX)^* = (I + tX + \frac{t^2}{2}X^2 + \dots)^* = I + tX^* + \frac{t^2}{2}(X^*)^2 + \dots$. Thus if we expand the two sides of

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

to first order in t we have

$$I_{p,q} + t(X^* I_{p,q} + I_{p,q} X) = I_{p,q}.$$

Hence $Lie(U(p, q))$ is contained in $\mathfrak{u}(p, q)$. If $X \in \mathfrak{u}(p, q)$ then $(X^*)^k I_{p,q} = (-1)^k I_{p,q} X^k$ thus

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

That is,

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

Hence $\exp(tX) \in U(p, q)$ for all $t \in \mathbb{R}$. We therefore see that $\text{Lie}(U(p, q)) = \mathfrak{u}(p, q)$.

1.3.4 Continuous homomorphisms, their differentials and a theorem of Von Neumann.

Let $G \subset GL(n, \mathbb{R})$ and let $H \subset GL(m, \mathbb{R})$ be closed subgroups. Let $\phi : H \rightarrow G$ be a continuous group homomorphism. If $X \in \text{Lie}(H)$ then $t \mapsto \phi(\exp tX)$ defines a continuous homomorphism of \mathbb{R} into $GL(n, \mathbb{R})$. Hence Theorem 11 implies that there exists $\mu(X) \in M_n(\mathbb{R})$ such that

$$\phi(\exp(tX)) = \exp(t\mu(X)).$$

We will now prove that $\mu : \text{Lie}(H) \rightarrow \text{Lie}(G)$ is a Lie algebra homomorphism. We will use Lemma 13. If $X, Y \in \text{Lie}(H)$ then by continuity we have

$$\begin{aligned} \phi(\exp(t(X+Y))) &= \phi(\lim_{k \rightarrow \infty} (\exp(\frac{t}{k}X) \exp(\frac{t}{k}Y))^k) \\ &= \lim_{k \rightarrow \infty} (\phi(\exp(\frac{t}{k}X)) \phi(\exp(\frac{t}{k}Y)))^k \\ &= \lim_{k \rightarrow \infty} (\exp(\frac{t}{k}\mu(X)) \phi(\exp(\frac{t}{k}\mu(Y))))^k \\ &= \exp(t\mu(X+Y)). \end{aligned}$$

Hence the uniqueness in Lemma 11 implies that $\mu(X+Y) = \mu(X) + \mu(Y)$. It is also obvious that $\mu(tX) = t\mu(X)$. Using the second limit formula in Lemma 13 in exactly the same way as we did the first we find that $\mu([X, Y]) = [\mu(X), \mu(Y)]$.

We will use the notation $d\phi = \mu$. This will be our preliminary definition of differential of a continuous homomorphism of closed subgroups of general linear groups. By the end of this subsection we will have made a link to the material in appendix D on Lie groups.

We will now record what we have proved up to this point in this subsection.

Proposition 18 *Let $G \subset GL(n, \mathbb{R})$ and $H \subset GL(m, \mathbb{R})$ be closed subgroups and let $\varphi : H \rightarrow G$ be a continuous homomorphism. Then there exists a unique Lie algebra homomorphism $\mu : \text{Lie}(H) \rightarrow \text{Lie}(G)$ such that $\varphi(\exp(X)) = \exp(\mu(X))$. We will use the notation $d\varphi$ for μ .*

We will now study in more detail the relationship between the Lie algebra of a closed subgroup, G , of $GL(n, \mathbb{R})$ the group structure. We first note that since the map $t \mapsto \exp tX$ from \mathbb{R} to G the Lie algebra of G is the same as the Lie algebra of its identity component. It is therefore reasonable to confine our attention to connected groups in this discussion. The key result in this direction is due to Von Neumann.

Theorem 19 *Let G be a closed subgroup of $GL(n, \mathbb{R})$ then there exists an open neighborhood, V , of 0 in $Lie(G)$ so that $\exp(V)$ is open in G and $\exp : V \rightarrow \exp(V)$ is a homeomorphism.*

Proof. Let $W = \{X \in M_n(\mathbb{R}) | \text{tr} X^t Lie(G) = \{0\}\}$. Then $M_n(\mathbb{R}) = Lie(G) \oplus W$ and orthogonal direct sum. If we take orthonormal basis of $M_n(\mathbb{R})$ that starts with a basis of $Lie(G)$ then continues with W then we can think of $M_n(\mathbb{R})$ as \mathbb{R}^{n^2} and of $Lie(G)$ as the subspace consisting of those vectors with the last $w = \dim W$ entries 0 and W as those with the first $d = \dim Lie(G)$ entries 0. We consider the map $\varphi : M_n(\mathbb{R}) \rightarrow GL(n, \mathbb{R})$ given by $\varphi(X) = \exp(X_1) \exp(X_2)$ with $X = X_1 + X_2$ and $X_1 \in Lie(G)$, $X_2 \in W$. We note that $\varphi(tX) = (I + tX_1 + O(t^2))(I + tX_2 + O(t^2)) = I + t(X_1 + X_2) + O(t^2)$. We therefore see that the differential of φ at 0 is the identity map. We may therefore apply the inverse function theorem and find there exists $s_1 > 0$ such that $\varphi : B_{s_1}(0) \rightarrow GL(n, \mathbb{R})$ has an open image U_1 with the map $\varphi : B_{s_1}(0) \rightarrow U_1$ a homeomorphism. With these preliminaries we can begin the argument. Suppose that for any $\varepsilon > 0$ with $s_1 \geq \varepsilon$ the set $\varphi(B_\varepsilon(0)) \cap G$ contains an element that outside of $\exp(Lie(G))$. Thus there exists for each $n \geq \frac{1}{s_1}$ an element in $Z_n \in B_{\frac{1}{n}}(0)$ such that $\exp(Z_n) \in G$ but $Z_n \notin Lie(G)$. We write $Z_n = X_n + Y_n$ with $X_n \in Lie(G)$ and $Y_n \in W$. Then

$$\varphi(Z_n) = \exp(X_n) \exp(Y_n).$$

Since $\exp(X_n) \in G$ we see that $\exp(Y_n) \in G$. We also observe that $\|Y_n\| \leq \frac{1}{n}$. Let $\varepsilon_n = \|Y_n\|$ then $\varepsilon_n \leq \frac{1}{n} \leq s_1$. There exists a positive integer m_n such that $s_1 \leq m_n \varepsilon_n < 2s_1$. Hence

$$s_1 \leq \|m_n Y_n\| < 2s_1.$$

Now $\exp(m_n Y_n) = \exp(Y_n)^{m_n} \in G$. Since the sequence $m_n Y_n$ is bounded we can replace it with a subsequence that converges. We may therefore assume that there exists $Y \in W$ with $\lim_{n \rightarrow \infty} m_n Y_n = Y$. Since G is closed $\exp(Y) \in G$. We will now use (variant of) the classic argument of Von Neumann to show that if $t \in \mathbb{R}$ then $\exp(tY) \in G$. Indeed, let $m_n t = a_n + b_n$ with $a_n \in \mathbb{Z}$ and $b_n \in \mathbb{R}$, $0 \leq b_n < 1$ (i.e. a_n is the integer part of $m_n t$). Then

$$tm_n Y_n = a_n Y_n + b_n Y_n.$$

Hence

$$\exp(tm_n Y_n) = \exp(Y_n)^{a_n} \exp(b_n Y_n).$$

Now $0 \leq b_n < 1$ so since $\lim_{n \rightarrow \infty} Y_n = 0$ we see that $\lim_{n \rightarrow \infty} \exp(b_n Y_n) = I$. We therefore see that

$$\exp(tY) = \lim_{n \rightarrow \infty} \exp(tm_n Y_n) = \lim_{n \rightarrow \infty} \exp(Y_n)^{a_n} \in G.$$

this completes the proof that $\exp(tY) \in G$ for all t . We have thus come upon the contradiction $Y \in Lie(G) \cap W = \{0\}$ since $Y \neq 0$. We have therefore shown that there exists an $\varepsilon > 0$ such that $\varphi(B_\varepsilon(0)) \cap G \subset \exp(Lie(G))$. ■

This result implies that a closed subgroup of $GL(n, \mathbb{R})$ has an open neighborhood of I that is homeomorphic with an open ball in \mathbb{R}^m with $m = \dim Lie(G)$. Now Theorem A.D.1 implies that it has a Lie group structure.

We also note that Proposition 16 be reformulated as

Proposition 20 *Let $G \subset GL(n, \mathbb{R})$ and $H \subset GL(m, \mathbb{R})$ be closed subgroups and let $\varphi : H \rightarrow G$ be a continuous homomorphism. Then φ is of class C^∞ .*

We will now relate the two notions of Lie algebra and differential that have now arisen.

Let x_{ij} be the standard coordinates of $M_n(\mathbb{R})$, that is we write $X \in M_n(\mathbb{R})$ as $X = [x_{ij}]$. In other words we take the standard basis E_{ij} (the matrix with exactly one non-zero entry that is a 1 in the i, j position. Then the x_{ij} are the coordinates of X with respect to this basis. We first note that if $f \in C^\infty(U)$ with U an open neighborhood of I in $M_n(\mathbb{R})$ then

$$\frac{\partial}{\partial x_{ij}} f(I) = \frac{d}{dt} f(I + tE_{ij})|_{t=0} = \frac{d}{dt} f(\exp(tE_{ij}))|_{t=0}.$$

Using the chain rule we see that

$$\frac{d}{dt} f(X(I + tE_{ij}))|_{t=0} = D_{ij} f(X)$$

with $D_{ij} = \sum_k x_{ki} \frac{\partial}{\partial x_{kj}}$. We therefore see that if $A \in M_n(\mathbb{R})$ then

$$\frac{d}{dt} f(X(I + tA))|_{t=0} = \sum_{ij} a_{ij} D_{ij} f(X).$$

Our observations about the expansion of $\exp(tA)$ in powers of t imply that

$$\frac{d}{dt} f(X \exp(tA))|_{t=0} = \sum_{ij} a_{ij} D_{ij} f(X).$$

We will write X_A for $\sum_{ij} a_{ij} D_{ij}$. Then $(X_A)_I = \sum_{ij} a_{ij} \frac{\partial}{\partial x_{ij}} \in T_I(M_n(\mathbb{R}))$. We note that a direct calculation implies that $[X_A, X_B] = X_{[A, B]}$.

Now assume that G is a closed subgroup of $GL(n, \mathbb{R})$. Then we have seen that G is a Lie subgroup. Let $i_G : G \rightarrow GL(n, \mathbb{R})$ be the injection of G into $GL(n, \mathbb{R})$. We assert that (the $Lie(G)$ in the statement is in the sense of this section)

Lemma 21 *We have $(di_G)_I(T_I(G)) = \{(X_A)_I | A \in Lie(G)\}$.*

Proof. Let $A \in Lie(G)$ then the map of \mathbb{R} to G given by $t \rightarrow \exp(tA)$ is a C^∞ map. If $f \in C^\infty(G)$ then we define

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

an element of $T_I(G)$. Now $(di_G)_I(v_A)f = (X_A)_I$. We therefore see that $(di_G)_I(T_I(G)) \supset \{(X_A)_I | A \in Lie(G)\}$. Since $\dim Lie(A) = \dim T_I(G)$ we see that the two spaces are the same. ■

We also note that if $A \in Lie(G)$ then we can define

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

Then by the very definition and the observation that the map $\mathbb{R} \times G \rightarrow G, t, g \mapsto g \exp tA$ is smooth we see that X_A^G is a left invariant vector field. We note

Lemma 22 *We have if $A, B \in Lie(G)$ then $[X_A^G, X_B^G] = X_{[A, B]}^G$.*

Proof. If $A, B \in Lie(G)$ then

$$\begin{aligned} f(\exp(sA) \exp(tB)) &= f(\exp sA) + tX_B^G f(\exp(sA)) + O(t^2) \\ &= f(I) + sX_A f(I) + O(s^2) + tX_B^G f(I) + stX_A^G X_B^G f(I) + O(t^2). \end{aligned}$$

We therefore see that since

$$f(\exp(sA) \exp(tB)) - f(\exp(tB) \exp(sA)) = st(X_A^G X_B^G f(I) - X_B^G X_A^G f(I)) + O(s^2) + O(t^2)$$

we have $[X_A^G, X_B^G]f(I) = \frac{\partial^2}{\partial s \partial t} (f(\exp(sA) \exp(tB)) - f(\exp(tB) \exp(sA)))|_{s=0, t=0}$.

On the other hand if $f \in C^\infty(GL(n, \mathbb{R}))$ then the same argument implies (since $[X_A, X_B] = X_{[A, B]}$ and $X_A^{GL(n, \mathbb{R})} = X_A$)

$$\frac{\partial^2}{\partial s \partial t} (f(\exp(sA) \exp(tB)) - f(\exp(tB) \exp(sA)))|_{s=0, t=0} = X_{[A, B]} f(I).$$

This implies that $(di_G)_I([X_A^G, X_B^G]_I) = (X_{[A, B]})_I = (di_G)_I(X_{[A, B]}^G)_I$. Hence, since $[X_A^G, X_B^G]$ is completely determined by $[X_A^G, X_B^G]_I$ we see that $[X_A^G, X_B^G] = X_{[A, B]}^G$. ■

We have already observed that if $G \subset GL(n, \mathbb{R})$ and $H \subset GL(m, \mathbb{R})$ are closed subgroups and if $\varphi : H \rightarrow G$ is a continuous homomorphism then φ is of class C^∞ . We will now calculate $d\varphi_I : T_I(H) \rightarrow T_I(G)$ we have (using the notation in the proof of the preceding lemma) the map $Lie(H) \rightarrow T_I(G), A \mapsto v_A$. Now if $f \in C^\infty(G)$ then

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

But $\varphi(\exp(tA)) = \exp(t\mu(A))$ (as in Proposition 16). Thus $d\varphi_I(v_A) = v_{\mu(A)}$. From this we see that if we define the vector field X_A^H on H by $X_A^H f(h) = \frac{d}{dt} f(h \exp(tA))|_{t=0}$ then we find that $d\varphi_h(X_A^H)_h = (X_{\mu(A)}^G)_{\varphi(h)}$.

We therefore see that the abstract definition of Lie algebra as left invariant vector fields and the concrete notion here of a space of matrices yield the same Lie algebras under the correspondence $A \mapsto X_A^G$. Furthermore the definition of differential in Proposition 16 yields the same differential as the abstract one under this correspondence.

1.3.5 Covering and quotient groups.

1.3.6 Exercises.

1. Complete the proof that the Lie algebras of the groups described in section 1.1 “line up” correctly with the Lie algebras described in section 1.2.

2. We will use the notation of Exercise 4 in 1.1.10. Observe 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 in $SO(V, B)$ and hence closed/

3. In this exercise we will use the notation of Exercise 7 of 1.1.10. Observe that φ is continuous and prove that $d\varphi$ is a Lie algebra isomorphism use this to prove that the image of φ is open and closed in the corresponding orthogonal group.

4. In the notation of exercises 8 and 9 of 1.1.10 prove that the appropriate differentials are Lie algebra isomorphisms.

2 Addendum for Appendix D.

In this addendum we add the following result which is for the most part proved in the end of the proof of Theorem D.2.8.

Theorem A.D.1. *Let G be a topological group and assume that there is an open neighborhood U of e (the identity of G) such that if $u \in U$ then $u^{-1} \in U$ and a surjective homeomorphism $\Phi : U \rightarrow B_r(0) \subset R^m$ for some $r > 0$, an s with $0 < s < r$ such that if $V = \Phi^{-1}(B_s(0))$ then if $u, v \in V$ then $uv \in U$ and a C^∞ map $F : B_s(0) \times B_s(0) \rightarrow B_r(0)$ such that*

- 1) $\Phi(u^{-1}) = -\Phi(u)$ for $u \in U$,
- 2) $\Phi(uv) = F(\Phi(u), \Phi(v))$ for $u, v \in V$.

Then there exists a structure of a Lie group on G compatible with its structure as a topological group.

We also observe that this combined with the following discussion gives a simple proof of the fact that a connected covering space of a connected Lie group is a Lie group.

If G is a connected and locally arcwise connected topological group we will now show that if we have a covering space $\pi : H \rightarrow G$ and we choose $e_o \in \pi^{-1}(e)$ then H has a structure of a topological group with identity e_o such that π is a group homomorphism. We will now begin that task.

Let $L_g : G \rightarrow G$ be given (as usual) by $L_g(x) = gx$. Then for each $h \in H$ there exists a unique homeomorphism $\tilde{L}_h : H \rightarrow H$ such that $\tilde{L}_h(e_o) = h$ and $\pi(\tilde{L}_h(x)) = L_{\pi(h)}(\pi(x))$. We assert that if we set $m(u, v) = \tilde{L}_u(v)$ then with this multiplication H is a group. The identity map has the same property assumed for \tilde{L}_{e_o} thus $m(e_o, u) = u$ and by definition $m(u, e_o) = u$ thus e_o is and identity element for the multiplication m . If $u \in H$ then there exists a unique $v \in H$ such that $\tilde{L}_u(v) = e_o$. Thus if we show that the associative rule is satisfied we will see that H with multiplication m and identity e_o is a group.

We note $L_x \circ L_y = L_{xy}$ and $\pi(m(x, y)) = xy$ and since $m(x, y) = \tilde{L}_{m(x, y)}(e_o) = \tilde{L}_x \circ \tilde{L}_y(e_o)$ we have $\tilde{L}_{m(x, y)} = \tilde{L}_x \circ \tilde{L}_y$. This is the content of the associative rule.

We must now prove that $m : H \times H \rightarrow H$ is continuous. We note that if we set $\mu(x, y) = xy$ for $x, y \in G$ then there is a unique lift $\tilde{\mu} : H \times H \rightarrow H$ of μ such that $\tilde{\mu}(e_o, e_o) = e_o$. Since m is another such lift we see that $m = \tilde{\mu}$.

It also implies an easy proof of the fact that if $H \subset G$ is a closed subgroup of a Lie group G then G/H has the structure of a Lie group.