Chapter 1
General Linear Groups

A. Groups

Before we can discuss matrix groups we need to talk a little about groups in general. If $X$ and $Y$ are sets, their Cartesian product $X \times Y$ is defined to be the set of all ordered pairs $(x,y)$ with $x \in X$ and $y \in Y$. A convenient notation for describing this set of all ordered pairs is

$$X \times Y = \{(x,y) | x \in X \text{ and } y \in Y\},$$

the curly brackets being read as "the set of all" and the vertical bar as "such that."

By a binary operation $\circ$ on a set $S$ we mean a function

$$\circ : S \times S \to S,$$

i.e., for an ordered pair $(s_1,s_2)$ of elements of $S$, $\circ$ assigns another element of $S$ which we write as $\circ(s_1,s_2)$. For example, the set $N = \{1,2,3,\ldots\}$ of natural numbers has two well-known binary operations on it. Addition sends the ordered pair $(a,b)$ of natural numbers to the natural number $a+b$. Multiplication sends the ordered pair $(a,b)$ to $ab$. 
Definition: A group \( G \) is a set \( G \) along with a binary operation

\[ \phi : G \times G \to G \]

satisfying certain properties. To state these properties it is convenient to adopt a simple notation--for \( \phi(a,b) \) we just write \( ab \).

Required properties of the operation:

(i) The operation is **associative**. This means that for any \( a, b, c \in G \) we have

\[ (ab)c = a(bc) \]

we had maintained the \( \phi(a,b) \) notation this would read

\[ (a,b,c) = \phi(\phi(a,b),c) \]

(ii) There exists an **identity** element \( e \) of \( G \). This means that for any \( a \in G \) we have \( ea = ae = a \).

(iii) **Inverses** exist. This means that for any \( a \in G \) there is element \( a^{-1} \in G \) such that \( aa^{-1} = a^{-1}a = e \).

Note that properties (ii) and (iii) leave open the possibilities that there may be more than one identity element and that an element have more than one inverse. But neither of these can happen.

Proposition 1: A group \( G \) has exactly one identity element and \( \forall a \in G \) has exactly one inverse.

Proof: Suppose \( e \) and \( f \) are identity elements of \( G \). Then

\[ fe = e \text{ since } f \text{ is an identity element, and} \]

\[ fe = f \text{ since } e \text{ is an identity element.} \]

Suppose both \( b \) and \( c \) are inverses of \( a \). Then

\[ b = eb = (ca)b = c(ab) = ce = c . \]

Examples

(1) The set \( \mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\} \) of integers is a group under addition. \( 0 \) is the identity and the inverse of \( a \) is \( -a \).

(2) \( \mathbb{Z} \) is not a group under multiplication. The operation is associative and \( 1 \) is the identity. But, for example, there is no inverse for \( 2 \).

(3) The set \( \mathbb{Q} \) of rational numbers is a group under addition.

(4) The set \( \mathbb{Q} - \{0\} \) (i.e., all nonzero rationals) is a group under multiplication.

(5) \( \mathbb{R}^+ = \{x \in \mathbb{R} | x > 0\} \) is the set of all positive real numbers. It forms a group under multiplication.

(6) \( \mathbb{R}^n \) = the set of all ordered n-tuples of real numbers is a group under the following operation: if

\[ x = (x_1, x_2, \ldots, x_n) \text{ and} \]

\[ y = (y_1, y_2, \ldots, y_n), \text{ then} \]

\[ x + y = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) . \]

The identity is

\[ x + 0 = (x_1, x_2, \ldots, x_n) \text{ and} \]

\[ 0 + x = (0, 0, \ldots, 0) . \]
\[ a = (0,0,\ldots,0) \]

the inverse of \( x \) is \((-x_1,-x_2,\ldots,-x_n)\).

Let \( S = \{a,b,c\} \); i.e., \( S \) is a set with three elements which denote by \( a,b,c \). Let \( G \) be the set of all one-to-one maps (functions) of \( S \) onto \( S \). For example \( f:S \to S \) given by
\[
f(a) = b, \quad f(b) = c, \quad f(c) = a
\]
is one element of \( G \). We define an operation on \( G \) as follows: if \( f,g \in G \) we let
\[
f \circ g : S \to S
\]
defined by \((f \circ g)(a) = f(g(a))\), \((f \circ g)(b) = f(g(b))\), \((f \circ g)(c) = f(g(c))\), i.e., \( f \circ g \) means first apply \( g \) to \( S \) and then apply \( f \).

Let \( i:S \to S \) be the identity element \((i(a) = a, i(b) = b, i(c) = c)\). In this is the identity element for \( G \) for this operation. Then usual inverse of \( f \in G \) is the inverse for \( f \) relative to this operation. Thus \( G \) is a group. It is called the symmetric group on \( \{a,b,c\} \) (or just the symmetric group on three elements).

**Definition:** A group \( G \) is **abelian** if for every \( a,b \in G \) we have \( ab = ba \).

In the examples above, (1), (3), (4), (5), and (6) are abelian groups, but the symmetric group on three elements is not abelian. (Exercise.)

The kind of functions (mapping one group to another) of interest to us are those which "preserve" the operations--these are called **homomorphisms**.

**Definition:** Let \( G \) and \( H \) be groups. A function \( \sigma:G \to H \) is a **homomorphism** if for every \( a,b \) in \( G \) we have
\[
\sigma(ab) = \sigma(a)\sigma(b).
\]

What this means is that we can first multiply \( a \) and \( b \) (using the operation in \( G \)) and then map the result by \( \sigma \), or we can map \( a \) and \( b \) into \( H \) by \( \sigma \) and multiply there--with the same result.

**Proposition 2:** A homomorphism \( \sigma:G \to H \) sends identity to identity and inverses to inverses.

**Proof:** Let \( e,e' \) be the identities in \( G,H \). We have
\[
\sigma(e) = \sigma(ee) = \sigma(e)\sigma(e) \quad \text{and} \quad \sigma(e) \text{ has an inverse, call it } h \text{, in } H.
\]
So
\[
e' = h\sigma(e) = h\sigma(e)\sigma(e) = \sigma(e).
\]
For \( a \in G \) we have
\[
\sigma(a)\sigma(a^{-1}) = \sigma(aa^{-1}) = \sigma(e) = e',
\]
showing that \( \sigma(a^{-1}) = (\sigma(a))^{-1} \).

A homomorphism is **surjective** (or **onto**) if \( \sigma(G) = H \). If we define \( \sigma:R \to \mathbb{R}^2 \) (\( R = \) additive group of reals, \( \mathbb{R}^2 \) as in example (6)) by \( \sigma(x) = (x,x) \), then \( \sigma \) is a homomorphism but is **not surjective** because \( \sigma(0) = (0,0) \) is just the diagonal line in \( \mathbb{R}^2 \). But \( \sigma: \mathbb{R} \to \mathbb{R} \) defined by \( \sigma(x,y) = x \) is a surjective homomorphism.

A homomorphism \( \sigma:G \to H \) is **injective** if \( \sigma(a) = \sigma(b) \) always implies \( a = b \); i.e., no two elements go to the same place. Sometimes this is called one-to-one-into, but we won't do that. For example, the map \( \sigma:R \to R \) \( \sigma(x) = (x,x) \) is injective, and the map
\( \sigma: \mathbb{R} \rightarrow \mathbb{R}^+ \) \( (x,y) = x \) is not injective.

A homomorphism which is both injective and surjective is called an isomorphism. From an abstract point of view, two groups which are isomorphic are "really" the same group—even if they were defined in seemingly different manners. There is a classic example of this.

Let \( \mathbb{R} \) be the additive group of all real numbers and let \( \mathbb{R}^+ \) (Example 5) be the multiplicative group of all positive real numbers. Let \( a \) be any real number greater than 1. Define

\[
\sigma: \mathbb{R} \rightarrow \mathbb{R}^+
\]

\[
\sigma(x) = a^x.
\]

Then \( \sigma \) is a homomorphism

\[
\sigma(x + y) = a^{x+y} = a^x a^y = \sigma(x) \sigma(y).
\]

So, \( \sigma \) is injective. For, suppose \( \sigma(x) = \sigma(y) \). This means \( a^x = a^y \) and so \( a^{-y} a^x = a^{-y} a^y = 1 \) and \( a^{x-y} = 1 \) which implies \( y = 0 \) or \( x = y \). Also, \( \sigma \) is surjective. For, if \( y \) is any positive real number \( x = \log_a y \) has the property that \( a^x = y \).

These two groups are isomorphic—not only that, but there are lots of isomorphisms.

We conclude this section with a simple, but important, remark. A priori it looks difficult to see if a homomorphism \( \sigma: G \rightarrow H \) is injective. Do we really have to check all pairs \( a, b \) in \( G \) to see \( \sigma(a) = \sigma(b) \)? Fortunately not.

\( \sigma \) is injective \( \iff \sigma^{-1}(\sigma'(e)) = e \).

\[
\sigma(a) = \sigma(b) = \sigma(a)^{-1} = e' = \sigma(ab^{-1}) = e'
\]

and

\[
ab^{-1} = e = a = b.
\]

\section{Fields, Quaternions}

\textbf{Definition:} A field \( k \) is a set that has operations of addition and multiplication satisfying certain requirements:

(i) multiplication distributes over addition;

\[
a(b + c) = ab + ac;
\]

(ii) \( k \) is an abelian group, with identity written as \( 0 \), under addition.

(iii) \( k - \{0\} \) is an abelian group under multiplication.

\textbf{Examples.} The rationals \( \mathbb{Q} \) and the reals \( \mathbb{R} \) are fields. We can make \( \mathbb{R}^2 \) into a field \( \mathbb{C} \) (the complex numbers) as follows. If \( (x_1, x_2) \) and \( (y_1, y_2) \) are two ordered pairs of real numbers, we define \( (x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2) \) and we have seen that this operation makes \( \mathbb{R}^2 \) into an abelian group. Suppose for multiplication we try

\[
(x_1, x_2)(y_1, y_2) = (x_1 y_1, x_2 y_2)
\]

(surely the most obvious thing). Then we would have

\[
(1,0)(0,1) = (0,0).
\]
(0,0) is the additive identity or "zero" and we would have two zero elements of \( \mathbb{R}^2 \) with a zero product. The result could not be a field because:

**Proposition 3:** In a field \( k \) if \( a \neq 0 \) and \( b \neq 0 \), then \( ab \neq 0 \).

**Proof:** If \( a \neq 0 \) then \( a \in k - \{0\} \) which by (iii) is required to be a group under multiplication. Thus there is an \( a^{-1} \) in \( k - \{0\} \) such that \( a^{-1}a = 1 \) (the multiplicative identity). Thus if \( ab = 0 \) we have

\[
a^{-1}(ab) = (a^{-1})(0) = 0
\]

so \( b = 0 \).

The statement of Proposition 3 is equivalent to the statement that a field has "no divisors of zero."

So how do we make \( \mathbb{R}^2 \) into a field? Our most naive attempt led flat. Well, what turns out to work is

\[(a,b)(c,d) = (ac - bd, ad + bc) .\]

must first verify that this distributes over addition.

\[
(a,b)((c,d) + (e,f)) = (a,b)((c + e, d + f))
\]

= \((a(c + e) - b(d + f), a(d + f) + b(c + e) .)\)

should equal \((a,b)(c,d) + (a,b)(e,f) \). This latter equals (a, b) \((c, d) + (a, b) \((e, f) \). and we easily check that these are equal. Next we need to see that if \((a, b) \neq (0, 0) \) then it has a multiplicative inverse. Well, \((a, b) \neq (0, 0) \) if \( a^2 + b^2 \neq 0 \); in which case, we need to find a multiplicative inverse for \((a, b) \). The multiplicative identity clearly is \((1, 0) \) and

\[
(a,b)(\frac{a}{a^2 + b^2}, -\frac{b}{a^2 + b^2}) = (1, 0),
\]

as you can readily verify. Thus we have made \( \mathbb{R}^2 \) into a field which we denote by \( \mathbb{C} \) and call the complex numbers.

You may know that there is a simple mnemonic device for remembering multiplication in \( \mathbb{C} \). Write \((a, b) = a + ib \) or \( a + bi \) and treat these as polynomials in \( i \) with the side condition that \( i^2 = -1 \). Thus

\[
(a + ib)(c + id) = ac + aid + ibc + ibid
\]

\[
= (ac - bd) + i(ad + bc) .
\]

We can consider \( \mathbb{R} \) to be a subfield of \( \mathbb{C} \) (i.e., a subset which becomes a field using the operations in the larger set) by letting

\[
x \in \mathbb{R} \text{ be } x + 10 .
\]

Then if \( x, y \in \mathbb{R} \) we have

\[
x + y = x + 10 + y + 10 = (x + y) + 10
\]

\[
xy = (x + 10)(y + 10) = (xy) + 10 .
\]

So we have taken the field \( \mathbb{R} \) as all \((x, 0)\) in \( \mathbb{R}^2 \) and extended...
The operations in \( \mathbb{R} \) to \( \mathbb{R}^2 \) to get a field.

This strongly suggests that we try to extend the field on \( \mathbb{R}^2 \) to a field on \( \mathbb{R}^3 \). Now for the bad news.

**Proposition 4:** The operations on \( \mathbb{R} \) cannot be extended to make into a field.

**Proof:** Take basis vectors 1, i, j so that any element of \( \mathbb{R}^3 \) can be written uniquely as \( a + ib + jc \) with \( a, b, c \in \mathbb{R} \).

If we are to have a multiplication extending that of \( \mathbb{R} \) we must have \( ij = a + ib + jc \) for some three real numbers \( a, b, c \). But then

\[
ij = 1 \implies 1 = a + ib + jc
\]

so

\[
j = 1a - b + ijc
\]

\[
-1 = a - b + (a + ib + jc)c
\]

\[
-j = (a - b) + i(a + b) + jc^2.
\]

This implies \( c^2 = -1 \), contradicting \( c \in \mathbb{R} \).

The main thrust of this proof is that if we insist that the product \( ij \) be in \( \mathbb{R}^3 \) we get into trouble. Maybe if we had one more dimension it would work. This is almost true; we can define a multiplication on \( \mathbb{R}^4 \) which satisfies conditions (i) and (ii) for a field but (iii) must be replaced by (iii)' which makes it not an abelian group. We will just describe how this can be done. You may be interested in reading "Hamilton's discovery of the quaternions" by B. L. van der Waerden in the Mathematics Magazine (vol. 49; #5, (1976)). We take a basis 1, i, j, k for \( \mathbb{R}^4 \) and define

\[
\begin{array}{cccc}
1 & i & j & k \\
1 & l & i & j \\
l & i & -1 & k \\
j & j & -k & l \\
k & k & j & -1 \\
\end{array}
\]

Thus \( l \) acts as identity, \( ij = k \), \( ji = -k \), etc.

This tells us how to multiply quadruples of real numbers:

\[
(a + ib + jc + kd)(x + iy + jz + kw) = (ax - by - cz - dw) + i(ay + bx + cw - dz) + j(az + cx + dy - bw) + k(aw + bx + bz - cy),
\]

with this multiplication is called the quaternions. It is easy to verify that this does extend the multiplication in \( \mathbb{R} \) by taking \( c = 0 = d \) and \( z = 0 = w \) in the formula above. The modified field axioms (i), (ii), (iii)' are readily verified except for showing that every nonzero quaternion has an inverse. But if

\[
q = a + ib + jc + kd
\]

is not the zero \( (0 + 0i + 0j + 0k) \) then \( a^2 + b^2 + c^2 + d^2 \neq 0 \) and we set

\[
q^{-1} = \frac{a - ib - jc - kd}{a^2 + b^2 + c^2 + d^2},
\]

and readily verify that \( qq^{-1} = 1 = q^{-1}q \).

There are certain constructions we want to make for \( \mathbb{R} \) and \( \mathbb{C} \) and the quaternions (which we denote by \( \mathbb{H} \)), so we will write
k \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}.

**Vectors and Matrices**

For \( k \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\} \) let \( k^n \) be the set of all ordered \( n \)-tuples of elements of \( k \). Define addition on \( k^n \) by
\[
x = (x_1, \ldots, x_n) \quad y = (y_1, \ldots, y_n)
\quad x + y = (x_1 + y_1, \ldots, x_n + y_n).
\]

This makes \( k^n \) into an abelian group with identity \( \mathbf{0} = (0, \ldots, 0) \).

If \( c \in k \) we define
\[
 cx = (cx_1, \ldots, cx_n)
\]

This makes \( k^n \) into a vector space over \( k \) (for \( k = \mathbb{H} \) we must take the usual definition which insists that \( k \) be a field).

**Definition:** A map \( k^n \rightarrow k^n \) is linear if it respects linear combinations; i.e., if \( c, d \in k \) and \( x, y \in k^n \) then
\[
(\ast) \quad \varphi(cx + dy) = c\varphi(x) + d\varphi(y).
\]

In particular, \( \varphi(x + y) = \varphi(x) + \varphi(y) \), so that a linear map is a homomorphism of the additive group of \( k^n \). Also
\[
 \varphi(cx) = c\varphi(x),
\]
these two conditions together are equivalent to \( (\ast) \).

**Proposition 5:** If \( k^n \rightarrow k^n \rightarrow k^n \) are both linear, then
\[
 \varphi \circ \psi.
\]

**Proof:**
\[
 (\varphi \circ \psi)(cx + dy) = \varphi(c \varphi(x) + d \varphi(y)) = c(\varphi \circ \psi)(x) + d(\varphi \circ \psi)(y).
\]

**Definition:** \( M_n(k) \) is the set of all \( n \times n \) matrices with elements from \( k \).

If \( M \in M_n(k) \), \( M = (m_{ij}) \) \( (m_{ij} \in k) \), we can define a linear map \( \varphi(M) \) by
\[
 \varphi(M)(x_1, \ldots, x_n) = (x_1, \ldots, x_n)(m_{ij})
\]

where matrix multiplication is indicated on the right; i.e., we are multiplying a \( 1 \times n \) matrix by an \( n \times n \) matrix to give a \( 1 \times n \) matrix. This is easily seen to be linear.

\[
 \varphi(M)(cx + dy) = (cx + dy)(m_{ij}) = c(x_1, \ldots, x_n)(m_{ij}) + d(y_1, \ldots, y_n)(m_{ij}).
\]

We use row vectors instead of column vectors because we have no longer have a choice when \( k = \mathbb{H} \). We made \( \mathbb{H}^n \) into a vector space by defining scalar multiplication on the left,
\[
 c(x_1, \ldots, x_n) = (cx_1, \ldots, cx_n)
\]
and this is not the same as \( (x_1, \ldots, x_n)c \) in general. If we use column vectors and multiply by matrices on the left we do not always get linear maps. For \( c, c, d \in \mathbb{H} \) and \( x, y \in \mathbb{H}^n \) consider
\[
\begin{pmatrix}
q_1 & \circ \\
\vdots & \vdots \\
q_n & \circ \\
\end{pmatrix}
\begin{pmatrix}
cx_1 + dy_1 \\
\vdots \\
cx_n + dy_n \\
\end{pmatrix} =
\begin{pmatrix}
qcx_1 + qdy_1 \\
\vdots \\
qcx_n + qdy_n \\
\end{pmatrix}
\]

and we certainly can't expect this to equal

\[
c \begin{pmatrix}
x_1 \\
\vdots \\
x_n \\
\end{pmatrix} +
\begin{pmatrix}
y_1 \\
\vdots \\
y_n \\
\end{pmatrix} =
\begin{pmatrix}
x_1 + y_1 \\
\vdots \\
x_n + y_n \\
\end{pmatrix}
\]

Take \( n = 1, x = 1, y = 1, d = 0, c = 1 \) and \( q = j \).

Conversely, given a linear map \( \varphi : k^n \to k^n \), it is easy to find an \( n \times n \) matrix \( M \) such that \( \varphi = \varphi(M) \) (and it will clearly be unique). The first row of \( M \) is the \( n \)-tuple \( \varphi(1,0,...,0) \), the second row of \( M \) is \( \varphi(0,1,0,...,0) \), etc.

Note that if the matrix \( A \) gives the linear map \( \varphi \) and the matrix \( B \) gives the linear map \( \psi \) then \( AB \) gives \( \varphi \circ \psi \). A linear map \( \varphi \) is an isomorphism if it is injective and surjective (same definitions as for group homomorphisms). Then \( \varphi^{-1} \) is also a linear map and \( \varphi \circ \varphi^{-1} = \text{identity map} = \varphi^{-1} \circ \varphi \). For the corresponding matrices this means that \( M(\varphi^{-1})M(\varphi) = I = M(\varphi)M(\varphi^{-1}) \) so that \( M(\varphi^{-1}) \) is a 2-sided inverse for \( M(\varphi) \). So if \( A^{-1} \) is a left inverse for \( A \), then it is also a right inverse for \( A \).

We make the set \( M_n(k) \) into a vector space in a fairly obvious way:

(1) If \( A = (a_{ij}) \) and \( B = (b_{ij}) \), then

\[ A + B = (a_{ij} + b_{ij}) \]
multiplication and $U \subset G$ is the set of units in $G$, then $U$ is a group under multiplication.

**Proof:** The operation is associative, there is an identity element 1 and every element has an inverse.

**Definition:** The group of units in the algebra $M_n(F)$ is denoted by $GL(n,F)$, in $M_n(C)$ by $GL(n,C)$ and in $M_n(H)$ by $GL(n,H)$. These are the **general linear groups**.

Note that: $A \in M_n(k)$ is a unit $\iff A$ represents an isomorphism of $k^n$.

**Definition:** If $G$ is a group and $H$ is a subset of $G$, then $H$ is a **subgroup** of $G$ if the operation on $G$ makes $H$ into a group.

**Proposition 7:** $H$ is a subgroup of the group $G$ if \((H \subset G \text{ and})\)

(i) $x,y \in H \Rightarrow xy \in H$,

(ii) id. el. is in $H$,

(iii) $x \in H \Rightarrow x^{-1} \in H$.

**Proof:** (Exercise.) The subject of this course is the study of subgroups of these general linear groups.

A $1 \times 1$ matrix over $k$ is just an element of $k$ and matrix multiplication of two is just multiplication in $k$. So we see that

- $GL(1,F) = \{a \mid a \neq 0\}$
- $GL(1,C) = \{c \mid c \neq 0\}$
- $GL(1,H) = \{h \mid h \neq 0\}$

because all nonzero elements are units. $GL(2,F)$ is the set of units in the vector space $M_2(F)$ of dimension 4. So

$$GL(2,F) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mid ad - bc \neq 0 \, ,$$

i.e., all points in 4-space not on the set where $ad = bc$.

For $F$ and $C$ we have determinants defined on $M_n(F)$ and $M_n(C)$ and from linear algebra we know that

$$GL(n,F) = \{A \in M_n(F) \mid \det A \neq 0\}$$

$$GL(n,C) = \{A \in M_n(C) \mid \det A \neq 0\} \, .$$

Suppose we define a "determinant" on $M_2(H)$ by

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc \, .$$

Then $\det \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = k - (-k) - 2k \neq 0\, , \text{ but this matrix cannot be a unit}$

of the corresponding linear map would be an isomorphism, whereas

$$\begin{bmatrix} i & j \\ i & j \end{bmatrix} \begin{bmatrix} i & 1 \\ 1 & j \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

and the map is not injective. Similar definitions give similar problems, but we can define a complex-valued determinant with the desired property: namely, $A \in M_n(H)$ has an inverse if and only if this determinant is nonzero.

**Proposition 8:** Let $\phi: G \rightarrow H$ be a homomorphism of groups. Then $\phi(G)$ is a subgroup of $H$.

**Proof:** $\phi(id) = id$ so that $\phi(G)$ contains the identity element of $H$. If $x,y \in \phi(G)$ there exist $a,b \in G$ such that $\phi(a) = x$, $\phi(b) = y$. Then
\[ xy = \varphi(a)\varphi(b) = \varphi(ab) \in \varphi(G). \]

Finally, suppose \( x \in \varphi(G) \). Then \( x = \varphi(a) \) and so \( x^{-1} = \varphi(a^{-1}) \in \varphi(G) \). So \( \varphi(G) \) is a subgroup of \( H \).

If \( \varphi: G \to H \) is an injective homomorphism, then \( \varphi \) is an isomorphism of \( G \) onto the subgroup \( \varphi(G) \) of \( H \), so we can then consider \( G \) as a subgroup of \( H \). We are going to construct an injective homomorphism

\[ \psi: \text{GL}(n, \mathbb{H}) \to \text{GL}(2n, \mathbb{C}), \]

and then for \( A \in \text{GL}(n, \mathbb{H}) \) we will assign as the determinant of \( A \) the determinant of \( \psi(A) \).

We begin with

\[ \psi: H \to \text{M}_2(\mathbb{C}) \]

defined by

\[ \psi(x+iy+jz+kd) = \begin{pmatrix} x+iy & -z-iw \\ z-iw & x-iy \end{pmatrix}. \]

**Lemma 9:**

(i) \( \psi(a+b) = \psi(a) + \psi(b) \)

(ii) \( \psi(ab) = \psi(a)\psi(b) \)

(iii) \( \psi \) is injective.

**Proof:** (i) is trivial and (ii) is a routine, but somewhat tedious, computation, and (iii) is trivial.

Next, for \( A \in \text{M}_n(\mathbb{H}) \) we set

\[ \gamma(A) = (\psi(a_{ij})). \]

i.e. \( \gamma(A) \) is the complex \( 2n \times 2n \) matrix whose \( 2 \times 2 \) block in the \( ij \) position is \( \psi(a_{ij}) \).

**Lemma 10:** \( \gamma(AB) = \gamma(A)\gamma(B) \).

**Proof:** Let \( A = (a_{ij}) \), \( B = (b_{ij}) \). Then

\[ (AB)_{ij} = a_{1i}b_{ij} + \cdots + a_{ni}b_{nj}. \]

By Lemma 9

\[ (\gamma(AB))_{ij} = \psi(a_{ij})\psi(b_{ij}). \]

and this is just the \( ij \) entry in \( \gamma(A)\gamma(B) \).

Now let \( A \in \text{GL}(n, \mathbb{H}) \) so that there exists \( A^{-1} \in \text{GL}(n, \mathbb{H}) \) with \( AA^{-1} = I = A^{-1}A \). Then \( \gamma(A) \) has \( \gamma(A^{-1}) = (\gamma(A))^{-1} \) so that \( \gamma(A) \) is nonsingular and thus \( \det \gamma(A) \neq 0 \).

Conversely, suppose \( \det \gamma(A) \neq 0 \). Then \( (\gamma(A))^{-1} \) exists and since \( \psi(\text{GL}(n, \mathbb{H})) \) is a subgroup of \( \text{GL}(2n, \mathbb{C}) \) we have that \( (\gamma(A))^{-1} \in \psi(\text{GL}(n, \mathbb{H})) \). Thus \( A^{-1} \in \text{GL}(n, \mathbb{H}) \) such that \( \gamma(A^{-1}) = (\gamma(A))^{-1} \).

Thus \( \gamma(AB) = I \) and \( \gamma \) is injective so \( AA^{-1} = I \). Thus \( A \) is nonsingular.

**Exercises**

1. Let \( \varphi: G \to H \) be a homomorphism of groups. The kernel of \( \varphi \) is defined to be

\[ \ker(\varphi) = \{ a \in G | \varphi(a) = e_H \}, \]
\[ \ker \varphi = \{ x \in G \mid \varphi(x) = \text{identity of } H \}. \]

Show that \( \ker \varphi \) is a subgroup of \( G \).

2. A subgroup \( W \) of a group \( G \) is normal if for each \( x \in G \) we have

\[ xWx^{-1} = W. \]

Show that \( \ker \varphi \) (Exercise 1) is a normal subgroup of \( G \).

3. The center \( C \) of a group \( G \) is defined by

\[ C = \{ y \in G \mid xy = yx \text{ for all } x \in G \}. \]

Show that \( C \) is a normal subgroup of \( G \).

4. Let \( S \) be a nonempty subset in a group \( G \). Define the center \( G(S) \) of \( S \) by

\[ G(S) = \{ x \in G \mid xs = sx \text{ for all } x \in S \}. \]

Show that \( G(S) \) is a subgroup of \( G \).

5. Let \( S \) be a nonempty set in a group \( G \). Define

\[ N(S) = \{ x \in G \mid xSx^{-1} = S \} \]

and call \( N(S) \) the normalizer of \( S \). Show that \( G(S) \subseteq N(S) \) and that \( N(S) \) is a subgroup of \( G \). Show that if \( S \) is a subgroup of \( G \), then \( S \subseteq N(S) \) and \( S \) is a normal subgroup of \( N(S) \).

6. Show that if \( \{ H_\alpha \mid \alpha \in A \} \) is any collection of subgroups of \( G \), then their intersection is also a subgroup of \( G \). If \( W \) is any subset of \( G \), by the subgroup generated by \( W \) we mean the intersection of all subgroups of \( G \) which contain \( W \). Show that this is the smallest subgroup of \( G \) which contains \( W \).

7. Consider two specific elements of \( G = \text{GL}(n,2) \)

\[ A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}. \]

Let \( H \) be the subgroup of \( G \) generated by \( A \) and \( K \) be the subgroup of \( G \) generated by \( B \). Prove that \( H = \{ \ldots, A^{-2}, A^{-1}, I, A, A^2, \ldots \} \), and similarly for \( K \).

8. Continuing with exercise 7, show that the product set

\[ HK = \{ hk \mid h \in H, k \in K \} \]

is not a subgroup of \( G \). (Show that \( ABAB \) is not of the form \( A^* B \).

9. We say that a subgroup \( K \) of \( G \) normalizes a subgroup \( H \) of \( G \) if for each \( k \in K \) we have \( kHk^{-1} = H \). Prove that if \( K \) normalizes \( H \), then \( KH \) is a subgroup of \( G \).

10. We can define an injective map

\[ \varphi : G \to M_2(\mathbb{R}) \]

as follows: represent \( a \in G \) as \( a = e^{i\theta} \) with \( \theta \geq 0 \) and set

\[ \varphi(a) = \sqrt{e} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}. \]

Show that \( \varphi(ab) = \varphi(a) \varphi(b) \) but that \( \varphi(a+b) \) need not equal \( \varphi(a)+\varphi(b) \).
11. Let $G$ be the multiplicative group of complex numbers of unit length. We say that $a \in G$ is a primitive $n^{th}$ root of unity if $a^n = 1$, but none of $a, a^2, \ldots, a^{n-1}$ are equal to one. Show that an isomorphism of $G$ onto itself must send primitive $n^{th}$ roots of unity to primitive $n^{th}$ roots of unity for each $n$. For each $n$, how many primitive $n^{th}$ roots of unity are there in $G$?

12. Let $\alpha = (a_1, a_2, a_3)$ and $\beta = (b_1, b_2, b_3)$ be two elements in $\mathbb{R}^3$. Take the two "purely imaginary" quaternions

$$
\alpha' = a_1 i + a_2 j + a_3 k,
$$

$$
\beta' = b_1 i + b_2 j + b_3 k.
$$

Show that if $\alpha'$ and $\beta'$ are multiplied as quaternions, then

$$
\alpha' \beta' = \text{real part}(\alpha' \beta')
$$

is just the usual cross product of vectors in $\mathbb{R}^3$.

Chapter 2
Orthogonal Groups

A. Inner products

We have a consistent notion of conjugation for $R \subseteq \mathbb{C} \subseteq \mathbb{H}$.

Namely,

$$
\bar{x} = x, \quad \bar{\bar{x}} = x.
$$

For $a = x + iy \in \mathbb{C}$, $\bar{a} = x - iy$.

For $q = x + iy + jz + kw \in \mathbb{H}$, $\bar{q} = x - iy - jz + kw$.

We clearly have $\bar{\bar{a}} = a$ in all cases and

$$
\bar{a + \beta} = \bar{a} + \bar{\beta}.
$$

It is an exercise to prove that

$$
\bar{a \beta} = \bar{a} \bar{\beta}.
$$

Of course for $R$ or $\mathbb{C}$ this is the same as

$$
\bar{a \beta} = \bar{a} \bar{\beta}.
$$

Let $k \in \{R, C, H\}$ and define an inner product $\langle , \rangle$ on $k^n$ by
\begin{align*}
\langle x, y \rangle &= x_1 \bar{y}_1 + \cdots + x_n \bar{y}_n .
\end{align*}

**Proposition 1:** \(( , )\) has the following properties:

(i) \(\langle x+y, z \rangle = \langle x, z \rangle + \langle y, z \rangle\)

(ii) \(\langle x, y+z \rangle = \langle x, z \rangle + \langle y, z \rangle\)

(iii) \(a \langle x, y \rangle = \langle ax, y \rangle, \ (x, ay) = \langle x, y \rangle a\)

(iv) \(\overline{\langle x, y \rangle} = \langle y, x \rangle\)

(v) \(\langle x, x \rangle\) is always a real number \(\geq 0\) and \(\langle x, x \rangle = 0 \iff x = (0, \ldots, 0)\).

(vi) If \(e_1, \ldots, e_n\) is the standard basis for \(k^n\) \(e_i = (0, \ldots, 0, 1, 0, \ldots, 0)\), then
\[
\langle e_i, e_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}
\]

(vii) The inner product is nondegenerate; i.e.,
- If \(\langle x, y \rangle = 0\) for all \(y\), then \(x = (0, \ldots, 0)\);
- If \(\langle x, y \rangle = 0\) for all \(x\), then \(y = (0, \ldots, 0)\).

**Proof:** Exercise.

**Definition:** The length \(|x|\) of \(x \in k^n\) is
\[
|x| = \sqrt{\langle x, x \rangle}.
\]

Recall that if \(A \in M_n(k)\), its conjugate \(\bar{A}\) is obtained by replacing each \(a_{ij}\) by \(\bar{a}_{ij}\); its transpose \(^tA\) is obtained by replacing each \(a_{ij}\) by \(a_{ji}\). These two operations commute so that

\begin{align*}
\langle x, y \rangle &= x_1 \bar{y}_1 + \cdots + x_n \bar{y}_n .
\end{align*}

Recall that for \(M_n\) we must operate on the right (since we defined (scalar)(vector) on the left). So we do the same for \(k^n\) and \(\mathbb{C}^n\). Thus we use row vectors.

**Proposition 2:** For any \(x, y \in k^n\) and \(A \in M_n(k)\) we have
\[
\langle xA, y \rangle = \langle x, y^t \bar{A} \rangle.
\]

**Proof:** Let \(A = (a_{ij})\).
\[
xA = (x_1 a_{11} \cdots + x_n a_{n1}, \ldots, x_1 a_{1n} \cdots + x_n a_{nn})
\]
\[
y^t \bar{A} = (y_1 \bar{a}_{11} \cdots + y_n \bar{a}_{n1}, \ldots, y_1 \bar{a}_{1n} \cdots + y_n \bar{a}_{nn})
\]
Thus the left hand side \(\langle xA, y \rangle\) equals
\[
(x_1 a_{11} \cdots + x_n a_{n1}) \bar{y}_1 + \cdots + (x_1 a_{1n} \cdots + x_n a_{nn}) \bar{y}_n,
\]
and the right hand side \(\langle x, y^t \bar{A} \rangle\) equals
\[
x_1 (a_{11} \bar{y}_1 + \cdots + a_{1n} \bar{y}_n) + \cdots + x_n (a_{n1} \bar{y}_1 + \cdots + a_{nn} \bar{y}_n).
\]
It is easy to see that these contain exactly the same terms.

**A. Orthogonal groups**

Again let \(k \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}\).
Definition: \( s(n,k) = \{ A \in \mathbb{M}_n(k) \mid \langle xA, yA \rangle = \langle x, y \rangle \text{ for all } x, y \in k^n \} \).

Proposition 3: \( s(n,k) \) is a group.

Proof: If \( A, B \in s(n,k) \), then
\[
\langle xAB, yAB \rangle = \langle xA, yA \rangle = \langle x, y \rangle
\]
so that
\[
AB \in s(n,k).
\]

Clearly the identity matrix \( I \) is in \( s(n,k) \).

If \( A \in s(n,k) \) we have
\[
\langle e_i A, e_j A \rangle = \langle e_i, e_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}.
\]
Now \( e_i A \) is just the \( i \)th row of \( A \) and we see that \( \langle e_i A, e_j A \rangle \) is just the \( ij \) entry in the product
\[
A^T A.
\]
Thus \( A^T A = I \). But then \( tA A \) is also the identity since
\[
tA A = t(\langle A A \rangle) = A^T A = I.
\]
Thus \( tA = A^{-1} \), a left hand and right inverse for \( A \). (More generally, we saw in section C of chapter I that for matrices a left inverse was automatically a right inverse.)

Finally,
\[
\langle xA^{-1}, yA^{-1} \rangle = \langle xA^{-1} A, yA^{-1} A \rangle = \langle x, y \rangle,
\]
showing that \( A^{-1} \in s(n,k) \). q.e.d.

Definition: For \( k = \mathbb{R} \) we write \( \sigma(n,k) \) as \( \sigma(n) \) and call it the orthogonal group. For \( k = \mathbb{C} \) we write it as \( U(n) \) and call it the unitary group. For \( k = \mathbb{H} \) we write it as \( Sp(n) \) and call it the symplectic group.

Proposition 4: Let \( A \in \mathbb{M}_n(k) \). Then the following conditions are equivalent:

(i) \( A \in s(n,k) \)

(ii) \( \langle e_i A, e_j A \rangle = \delta_{ij} \)

(iii) \( A \) sends orthonormal bases to orthonormal bases

(iv) The rows of \( A \) form an orthonormal basis

(v) The columns of \( A \) form an orthonormal basis

(vi) \( tA = A^{-1} \).

Proof: Exercise.

Proposition 5: Let \( A \in \mathbb{M}_n(k) \). Then \( A \in s(n) \) if and only if \( A \) preserves lengths.

Proof: \( A \) preserves lengths if \( \langle xA, xA \rangle = \langle x, x \rangle \) for all \( x \). So \( \sigma \) is trivial. Conversely, we have
\[
\langle (x+y)A, (x+y)A \rangle = \langle x+y, x+y \rangle = \langle x, x \rangle + \langle y, y \rangle + \langle x, y \rangle + \langle y, x \rangle
\]
\[
= \langle xA, xA \rangle + \langle xA, yA \rangle + \langle yA, xA \rangle + \langle yA, yA \rangle.
\]
Thus given \( \langle x, y \rangle + \langle y, x \rangle = \langle xA, yA \rangle + \langle yA, xA \rangle \) and since \( \langle , \rangle \) is symmetric, this proves...
Proposition 5 bis: Proposition 5 also holds for \( \mathfrak{g} \) and \( \mathfrak{h} \).

Proof: Calculate \( \langle e_i + e_j, A, (e_i + e_j) A \rangle \) just as above to get

\[
\langle e_i A, e_j A \rangle + \langle e_j A, e_i A \rangle = 0.
\]

Then consider \( x = x_i e_i + x_j e_j \) and calculate \( \langle x A, x A \rangle \). We get

\[
x_i \overline{x}_j \langle e_i A, e_j A \rangle + x_j \overline{x}_i \langle e_j A, e_i A \rangle = 0
\]

and thus

\[
\langle e_i A, e_j A \rangle (x_i \overline{x}_j - x_j \overline{x}_i) = 0
\]

and this forces \( \langle e_i A, e_j A \rangle = 0 \). q.e.d.

Let us look at \( \mathfrak{o}(n) \), \( \mathfrak{u}(n) \) and \( \mathfrak{sp}(n) \) for small \( n \). \( \mathfrak{o}(1) \)
is the set of all real numbers of length one, so \( \mathfrak{o}(1) = \{1, -1\} \).
\( \mathfrak{u}(1) \) is just the set of all complex numbers of length one. This is
the circle group \( S^1 \). \( \mathfrak{sp}(1) \) is the group of all quaternions of
unit length. If we define

\[
S^{k-1} = \{ x \in \mathbb{R}^k \mid |x| = 1 \}
\]
to be the unit \((k-1)\)-sphere we see that

\( \mathfrak{o}(1) = S^0 \), \( \mathfrak{u}(1) = S^1 \), \( \mathfrak{sp}(1) = S^3 \).

It is an interesting fact that these are the only spheres which can
be groups.

Proposition 6: If \( k \in \{\mathbb{R}, \mathbb{C}\} \) and \( A \in \mathfrak{o}(n,k) \), then

\[
\det A \det \overline{A} = 1.
\]

Proof: \( \overline{A} \mathbf{1} = \mathbf{1} - (\det A)(\det \overline{A}) = 1 \), and clearly

\[
\det \overline{A} = \det A = \det \overline{A}.
\]

q.e.d.

Thus if \( A \in \mathfrak{o}(n) \) \( (-\mathfrak{o}(n)) \), then \( \det A \in \{1, -1\} \). We
define

\[
\mathfrak{so}(n) = \{ A \in \mathfrak{o}(n) \mid \det A = 1 \}
\]

called the special orthogonal group (also called the rotation group). Similarly, we define

\[
\mathfrak{su}(n) = \{ A \in \mathfrak{u}(n) \mid \det A = 1 \}
\]

called this the special unitary group.

An example of an element of \( \mathfrak{o}(2) - \mathfrak{so}(2) \) is \( \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \). This
sends \( e_1 = (1,0) \) to \( e_1 \) and sends \( e_2 = (0,1) \) to \( -e_2 \). It is
not the reflection in the first axis, and has determinant equal to

the isomorphism question

At the end of Chapter I we showed that two groups which were

\overset{\text{exchanged}}{\text{isomorphic}}

were quite differently were isomorphic. We have now defined several

\overset{\text{classes}}{\text{of groups}} \( \mathfrak{gl}(n,k) \) for \( n = 1,2,\ldots \) and \( k \in \{\mathbb{R}, \mathbb{C}, \mathbb{H} \} \) and
\( \mathfrak{so}(n,k) \) for \( n = 1,2,\ldots \) and \( k \in \{\mathbb{R}, \mathbb{C}, \mathbb{H} \} \) and our major goal is to

\overset{\text{identify}}{\text{determine}}

which of these are isomorphic. The basic idea will be to

\overset{\text{identify}}{\text{identify}}

invariants of matrix groups (dimension, rank, etc.), i.e., two

groups which are isomorphic must have the same invariants. This will

\overset{\text{be}}{\text{not}}

possible to say that certain groups are not isomorphic. But

\overset{\text{nevertheless}}{\text{two}}

differently defined groups are indeed isomorphic, an
isomorphism may be hard to find. This is why we will work so hard to develop invariants—to reduce as much as possible the cases where we must look for isomorphisms. In this section we will give one isomorphism.

Suppose you suspect that $\text{Sp}(1)$ and $\text{SU}(2)$ are isomorphic. How would you try to find an isomorphism? $\text{Sp}(1)$ is the set of all quaternions of unit length and $\text{SU}(2)$ is the set of all complex $2 \times 2$ matrices $A$ such that $A^* A = I$ and $\det A = 1$. The operation in $\text{Sp}(1)$ is multiplication of quaternions, in $\text{SU}(2)$ it is matrix multiplication.

**Proposition 7:** The map $\phi : M_n(\mathbb{H}) \to M_{2n}(\mathbb{C})$ defined in §D of Chapter I induces an isomorphism

$$\phi : \text{Sp}(1) \to \text{SU}(2).$$

**Proof:** We have seen that $\phi$ induces an injective homomorphism of $\text{GL}(n,\mathbb{H})$ into $\text{GL}(2n,\mathbb{C})$, so restriction of $\phi$ to $\text{Sp}(1)$ is still an injective homomorphism. So we just need to show that

(i) $A \in \text{Sp}(1) = \phi(A) \in \text{SU}(2)$ and

(ii) every $B \in \text{SU}(2)$ is some $\phi(A)$ with $A \in \text{Sp}(1)$.

If $A = a + ib + jc + kd$ then $\phi(A) = \begin{pmatrix} a+ib & -c+id \\ c-id & a-ib \end{pmatrix}$ so that

$$\phi(A) \phi(A) = \begin{pmatrix} a+ib & -c+id \\ c-id & a-ib \end{pmatrix} \begin{pmatrix} a+ib & -c+id \\ c-id & a-ib \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

since $a^2 + b^2 + c^2 + d^2 = 1$. Also $\det \phi(A) = 1$ so $\phi(A) \in \text{SU}(2)$.

Let $B = \begin{pmatrix} a \\ b \end{pmatrix} \in \text{SU}(2)$. Using $\det B = 1$ and the fact that the rows are orthogonal unit vectors, we find that

$$\phi(A) = \begin{pmatrix} a+ib & -c+id \\ c-id & a-ib \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \cdot a + b \cdot b \\ c \cdot a + d \cdot b \end{pmatrix}.$$
relative to our standard basis \( e_1, \ldots, e_n \) the reflection \( \sigma \) is given by

\[
A = \begin{pmatrix}
-1 & 0 \\
0 & 1
\end{pmatrix}
\]

Conversely we see that such a matrix represents a reflection in the orthogonal complement of the vector \( e_1 \).

In \( \mathbb{R}^2 \) let the unit vector \( u \) be written as

\[
u = (\cos \alpha, \sin \alpha).
\]

Then \((-\sin \alpha, \cos \alpha)\) is a unit vector in \( u^\perp \). The matrix \( A \) sending \( e_1 \) to \( u \) and \( e_2 \) to \((-\sin \alpha, \cos \alpha)\) must satisfy

\[
\begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix}
\begin{pmatrix}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{pmatrix}
= (\cos \alpha, \sin \alpha)
\]

\[
\begin{pmatrix}
0 & 1 \\
1 & 0
\end{pmatrix}
\begin{pmatrix}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{pmatrix}
= (-\sin \alpha, \cos \alpha)
\]

so

\[
A = \begin{pmatrix}
\cos \alpha & \sin \alpha \\
-\sin \alpha & \cos \alpha
\end{pmatrix}
\]

Thus the matrix giving reflection in \( u^\perp \) is

\[
\sigma = \begin{pmatrix}
\cos \alpha & -\sin \alpha \\
\sin \alpha & \cos \alpha
\end{pmatrix}
\begin{pmatrix}
-1 & 0 \\
0 & 1
\end{pmatrix}
\begin{pmatrix}
\cos \alpha & \sin \alpha \\
-\sin \alpha & \cos \alpha
\end{pmatrix}
\]

\[
\sigma = \begin{pmatrix}
\cos 2\alpha & \sin 2\alpha \\
-\sin 2\alpha & \cos 2\alpha
\end{pmatrix}
\]

The matrix \( A \) is easily seen to be a rotation of \( \mathbb{R}^2 \) through an angle \( \alpha \).

Later in this course we will prove that \( \mathfrak{S}(n) \) is generated by reflections—that is, any element of \( \mathfrak{S}(n) \) may be obtained by a finite sequence of reflections.

\section{Exercises}

1. Prove Proposition 1.


3. Let \( A \) be any element of \( \mathfrak{S}(n) \) with \( \det A = -1 \). Show that

\[
\mathfrak{S}(n) = \text{SO}(n) = \{ BA \mid B \in \text{SO}(n) \}.
\]

4. Show that any element of \( \text{SO}(2) \) can be written as

\[
\begin{pmatrix}
\cos \theta & \sin \theta \\
-\sin \theta & \cos \theta
\end{pmatrix}
\]

5. If \( A \in \text{U}(n) \) and \( \lambda \in \mathbb{C} \) has length one, show that \( A \in \text{U}(n) \).

6. Let \( L_1 \) and \( L_2 \) be lines through the origin in \( \mathbb{R}^2 \). Show that reflection in \( L_1 \) followed by reflection in \( L_2 \) equals a reflection through twice the angle between \( L_1 \) and \( L_2 \).

7. A matrix \( A \in \text{M}_n(\mathbb{R}) \) is said to be idempotent if \( AA = A \).

Show that the image of \( \mathbb{R}^n \) under \( A \) is precisely the fixed-point set of \( A \). Such a map is called a projection of \( \mathbb{R}^n \) onto its image.
What can you say about the determinant of a idempotent matrix? What is the image of $S^2$ under $A = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$.

8. A matrix $A$ is nilpotent if some power of it is the zero matrix. For example, $A = \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix}$ has its third power zero. Prove that a nilpotent matrix is singular. Prove that any $A = (a_{ij})$ with $a_{ij} = 0$ whenever $i > j$ is nilpotent. Find two nilpotent matrices $A$ and $B$ whose product $AB$ is not nilpotent.

9. Let $U$ be the set of all matrices $A = (a_{ij})$ with all diagonal elements equal to one and $a_{ij} = 0$ whenever $i > j$. Prove that $U$ is a group under matrix multiplication (the group of unipotent matrices in $M_n(\mathbb{R})$).

Chapter 3
Homomorphisms

A. Curves in a vector space

We are going to define our first invariant of a matrix group, its dimension. Matrix groups whose dimensions are different can't be isomorphic. The dimension of a matrix group is going to be the dimension of its space of tangent vectors (a vector space), so we first define these.

Let $V$ be a finite-dimensional real vector space. By a curve $\gamma: (a,b) \to V$ we mean a continuous function $\gamma: (a,b) \to V$ where $(a,b)$ is an open interval in $\mathbb{R}$.

When we say $\gamma$ is differentiable at $c$ if

$$\lim_{h \to 0} \frac{\gamma(c+h) - \gamma(c)}{h}$$

exists. When this limit exists, it is a vector in $V$. We denote it $\gamma'(c)$ and call it the tangent vector to $\gamma$ at $\gamma(c)$.
It is a standard result from calculus that if we choose a basis for \( V \) and thus represent \( \gamma \) as \((\gamma_1, \ldots, \gamma_n)\) (\(\gamma_i\)'s being real valued), then \( \gamma'(c) \) exists if and only if each \( \gamma_i'(c) \) exists and
\[
\gamma'(c) = (\gamma_1'(c), \ldots, \gamma_n'(c)).
\]

Now \( M_n(k), M_n(R), M_n(H) \) can all be considered to be real vector spaces (of dimensions \( n^2, 2n^2 \) and \( 4n^2 \)). If \( G \) is a matrix group in \( M_n(k) \) then a curve in \( G \) is a curve in \( M_n(k) \) with all values \( \gamma(u) \) for \( u \in (a,b) \) lying in \( G \).

\[
\begin{array}{c}
M_n(k) \\
zero \\
\gamma \\
I \\
\end{array} \quad \begin{array}{c}
G \\
r \in \mathbb{R}
\end{array}
\]

Suppose we have curves \( \gamma, \sigma : (a,b) \to G \). Then we can define a new curve, the product curve, by
\[
(\gamma \sigma)(u) = \gamma(u)\sigma(u).
\]

**Proposition 1:** Let \( \gamma, \sigma : (a,b) \to G \) be curves, both of which are differentiable at \( c \in (a,b) \). Then the product curve \( \gamma \sigma \) is differentiable at \( c \) and
\[
(\gamma \sigma)'(c) = \gamma(c)\sigma'(c) + \gamma'(c)\sigma(c).
\]

**Proof:** Let \( \gamma(u) = (\gamma_{ij}(u)) \), \( \sigma(u) = (\sigma_{ij}(u)) \). Then
\[
(\gamma \sigma)(u) = \left( \sum_k \gamma_{ik}(u)\sigma_{kj}(u) \right),
\]
so that
\[
(\gamma \sigma)'(u) = \sum_k \gamma_{1k}(u)\sigma_{kj}'(u) + \gamma_{ik}'(u)\sigma_{kj}(u) = \gamma'(u)\sigma(u) + \gamma(u)\sigma'(u).
\]

**Proposition 2:** Let \( G \) be a matrix group in \( M_n(k) \). Let \( T \) be the set of all tangent vectors \( \gamma'(0) \) to curves \( \gamma : (a,b) \to G \), \( \gamma(0) = I \) \( (0 \in (a,b)) \). Then \( T \) is a subspace of \( M_n(k) \).

**Proof:** If \( \gamma'(0) \) and \( \sigma'(0) \) are in \( T \), then
\[
(\gamma \sigma)(0) = \gamma(0)\sigma(0) = II = I
\]
and
\[
(\gamma \sigma)'(0) = \gamma'(0)\sigma(0) + \gamma(0)\sigma'(0) = \gamma'(0) + \sigma'(0).
\]

Thus \( T \) is closed under vector addition.

\( T \) is also closed under scalar multiplication, for if \( \gamma'(0) \in T \) and \( r \in \mathbb{R} \), let
\[
\sigma(u) = \gamma(ru).
\]

Then \( \sigma(0) = \gamma(0) = I \), \( \sigma \) is differentiable and \( \sigma'(0) = r\gamma'(0) \).

Since \( M_n(k) \) is a finite dimensional vector space, so is \( T \).

**Definition:** If \( G \) is a matrix group, its dimension is the dimension of the vector space \( T \) (of tangent vectors to \( G \) at \( I \)).

**Example 1:** \( U(1) \) has dimension 1.
Example 2: \( \text{dim } \text{Sp}(1) = 3 \).

Let \( \gamma: (a, b) \to \text{Sp}(1) \) be a smooth curve with \( \gamma(0) = 1 \). Then \( \gamma'(0) \) will be an element of \( H = \mathbb{R}^4 \). We first show \( \gamma'(0) \) is in the span of \( i, j, k \); i.e. it is a quaternion with zero real part. Let

\[
\gamma(t) = x(t) + iy(t) + jz(t) + kw(t)
\]

with \( x(0) = 1 \) and \( y(0) = 0 \), \( z(0) = 0 \), \( w(0) = 0 \). We note that \( x(0) \) is a maximum for the function \( x \) so that \( \gamma'(0) = 0 + iy'(0) + jz'(0) + kw'(0) \), as asserted.

Conversely, let \( q = iu + jv + kw \) be any quaternion with zero real part. We claim that there exists a smooth curve \( \gamma \) in \( \text{Sp}(1) \) such that \( \gamma'(0) = q \). Indeed,

\[
\gamma(t) = \sqrt{1 - \sin^2 ut - \sin^2 vt - \sin^2 wt + i \sin vt + j \sin wt + k \sin wt}
\]

can be readily verified to be such a curve (which is defined on some interval \( [0, \epsilon] \), i.e. for \( t \) small).

Example 3: \( \text{dim } \text{GL}(n, \mathbb{R}) = n^2 \).

The determinant function \( \det: \text{M}_n(\mathbb{R}) \to \mathbb{R} \) is continuous and \( \det(I) = 1 \). So there is some \( \epsilon \)-ball about \( I \) in \( \text{M}_n(\mathbb{R}) \) such that for each \( A \) in this ball \( \det A \neq 0 \); i.e.,

\( A \in \text{GL}(n, \mathbb{R}) \). If \( v \) is any vector in \( \text{M}_n(\mathbb{R}) \) define a curve \( \sigma \) in \( \text{M}_n(\mathbb{R}) \) by

\[
\sigma(t) = tv + I.
\]

Then \( \sigma(0) = I \) and \( \sigma'(0) = v \) and for small \( t \), \( \sigma(t) \) is in \( \text{GL}(n, \mathbb{R}) \).

Hence the tangent space \( T \) is all of \( \text{M}_n(\mathbb{R}) \) which has dimension \( n^2 \).

A similar argument shows that \( \text{dim } \text{GL}(n, \mathbb{C}) = 2n^2 \).

We will now get upper bounds for the dimensions of \( \text{O}(n) \), \( \text{U}(n) \) and \( \text{Sp}(n) \) after a few preliminaries.

Definition: \( A \in \text{M}_n(\mathbb{R}) \) is said to be **skew-symmetric** if

\[
A + \overline{A} = 0.
\]

I.e. if \( a_{ij} = -a_{ji} \) for each \( i,j \). In particular, the diagonal elements must all be zero.

Proposition 3: Let \( \text{so}(n) \) denote the set of all skew-symmetric matrices in \( \text{M}_n(\mathbb{R}) \). Then \( \text{so}(n) \) is a linear subspace of \( \text{M}_n(\mathbb{R}) \), and its dimension is \( \frac{n(n-1)}{2} \).

Proof: The zero matrix is in \( \text{so}(n) \), and if \( A, B \) belong to \( \text{so}(n) \), then

\[
(A + B) + \overline{(A + B)} = A + \overline{A} + B + \overline{B} = 0,
\]

so that \( \text{so}(n) \) is closed under vector addition. It is also closed under scalar multiplication, for if \( A \in \text{so}(n) \) and \( r \in \mathbb{R} \), then

\[
(rA) + \overline{(rA)} = r(A + \overline{A}) = r(A + \overline{A}) = 0.
\]

To check the dimension of \( \text{so}(n) \) we get a basis. Let \( E_{ij} \) denote the matrix whose entries are all zero except the \( ij \) entry, which is 1, and the \( ji \) entry, which is -1. If we define these
E^i_j only for i < j, it is easy to see that they form a basis for so(n), and it is easy to count that there are
\[(n-1) + (n-2) + \ldots + 1 = \frac{n(n-1)}{2}\]
of them.

**Definition:** A matrix \( B \in M_n(c) \) is **skew-Hermitian** if
\[B^T \bar{c} = 0.\]

Thus if \( b_{jk} = c + i d \), then \( \bar{b}_{kj} = -b_{jk} = -c - i d \) and \( b_{kj} = -c + i d \). In particular if \( j = k \) we have \( c + i d = -c + i d \), so that the diagonal terms of a skew-Hermitian matrix are purely imaginary.

Let \( su(n) \) be the set of skew-Hermitian matrices in \( M_n(c) \). By the observation just made we see that \( su(n) \) is not a vector space over \( c \).

**Proposition 4:** \( su(n) \subset M_n(c) \) is a real vector space of dimension
\[n + 2 \frac{n(n-1)}{2} = n^2.\]

**Proof:** Exercise.

We make a similar definition for matrices in \( M_n(H) \), and call \( C \in M_n(H) \) **skew-symplectic** if
\[C + C^T = 0.\]

In the exercises one shows that the set \( sp(n) \) of such matrices is a real vector space of dimension
\[3n + 4 \frac{n(n-1)}{2} = n(2n+1).\]

**Proposition 5:** If \( \gamma \) is a curve through the identity \( (\gamma(0) = I) \), then \( \gamma'(0) \) is skew-symplectic.

**Proof:** In each case we have that the product curve is constant
\[\gamma(u) \gamma(u) = I.\]

Thus its derivative is zero, and the result follows from Proposition 1.

**Corollary:**
\[
\begin{align*}
\dim \text{O}(n) & \leq \frac{n(n-1)}{2} \\
\dim \text{U}(n) & \leq n^2 \\
\dim \text{Sp}(n) & \leq n(2n+1)
\end{align*}
\]

Later we will show that these are equalities.

**Smooth homomorphisms**

Let \( \phi : G \to H \) be a homomorphism of matrix groups. Since \( G \) and \( H \) are in vector spaces, it is clear what it means for \( \phi \) to be continuous. From now on homomorphism always means continuous homomorphism. This being so, a curve
\[\phi : (a, b) \to G\]

gives a curve \( \phi \circ : (a, b) \to H \) by \( (\phi \circ)(u) = \phi(\alpha(u)) \) in \( H \).

**Definition:** A homomorphism \( \phi : G \to H \) of matrix groups is smooth
if for every differentiable curve \( \gamma \) in \( G \), \( \phi \circ \gamma \) is differentiable.

**Definition:** Let \( \varphi : G \to H \) be a smooth homomorphism of matrix groups. If \( \gamma'(0) \) is a tangent vector to \( G \) at \( I \) we define a tangent vector \( d\varphi(\gamma'(0)) \) to \( H \) at \( I \) by

\[
d\varphi(\gamma'(0)) = (\phi \circ \gamma)'(0) .
\]

The resulting map \( d\varphi : T_G \to T_H \) is called the **differential** of \( \varphi \).

**Proposition 6:** \( d\varphi : T_G \to T_H \) is a linear map.

**Proof:** If \( \varphi'(0) \) and \( \gamma'(0) \) are in \( T_G \), consider

\[
d\varphi(a\varphi'(0) + b\gamma'(0))
\]

with \( a,b \in \mathbb{R} \). By definition this equals

\[
[\varphi(a\varphi' + b\gamma')]'(0) = [a(\varphi \circ \varphi') + b(\varphi \circ \gamma')]'(0)
\]

\[
= a(\varphi \circ \varphi)'(0) + b(\varphi \circ \gamma)'(0) = a d\varphi(\varphi'(0)) + b d\varphi(\gamma'(0)) ,
\]

proving that \( d\varphi \) is linear.

**Proposition 7:** If \( G \) \( H \ \mathbb{R} \) are smooth homomorphisms, then so is \( \gamma \circ \varphi \) and

\[
d(\gamma \circ \varphi) = d\gamma \circ d\varphi .
\]

**Proof:** The first part is obvious. For the second, let \( \gamma'(0) \)
be a tangent vector of \( G \). Then

\[
d(\gamma \circ \varphi)(\gamma'(0)) = (\gamma \circ \varphi)'(0) = d\gamma(\varphi \circ \gamma)'(0) = d\gamma \circ d\varphi(\gamma'(0)) .
\]

**Corollary:** If \( \varphi : G \to H \) is a smooth isomorphism, then

\[ d\varphi : T_G \to T_H \text{ is a linear isomorphism and } \dim G = \dim H . \]

**Proof:** \( \varphi^{-1} \circ \varphi \) is the identity, so \( d\varphi^{-1} \circ d\varphi : T_G \to T_G \) is the identity. Thus \( d\varphi \) is injective and \( d\varphi^{-1} \) is surjective. \( \varphi^{-1} \circ \varphi \) is the identity, so \( d\varphi \circ d\varphi^{-1} : T_H \to T_H \) is the identity. Thus \( d\varphi^{-1} \) is injective and \( d\varphi \) is surjective.

\( \text{q.e.d.} \)

### Exercises

1. Let \( \gamma : (-1,1) \to N_3(\mathbb{R}) \) be given by

\[
(\cos t \ \sin t \ 0) \\
(-\sin t \ \cos t \ 0) \\
0 \ 0 \ 1
\]

Show that \( \gamma \) is a curve in \( SO(3) \) and find \( \gamma'(0) \). Show that

\[
(\gamma')'(0) = 2\gamma'(0) .
\]

7. Let \( \sigma : (-1,1) \to N_3(\mathbb{R}) \) be given by

\[
(0 \ 0 \ 0) \\
(0 \ 0 \ 0) \\
0 \ -\sin t \ \cos t
\]

Show that \( \sigma' \) is a curve in \( SO(3) \). Write the matrix for \( \gamma(t) \circ \sigma(t) \) and verify that

\[
(\gamma \circ \sigma)'(0) = \gamma'(0) + \sigma'(0) .
\]
3. Let $\phi : (-1,1) \to M_3(\mathbb{R})$ be given by

$$\phi(t) = \begin{pmatrix} e^{it} & 0 & 0 \\ 0 & e^{\frac{it}{2}} & 0 \\ 0 & 0 & e^{\frac{it}{2}} \end{pmatrix}.$$ 

Show that $\phi$ is a curve in $U(3)$. Calculate $\phi'(0)$.

4. Let $\alpha : (-1,1) \to \mathbb{H}$ be defined by

$$\alpha(t) = (\cos t)j + (\sin t)k.$$ 

Show that $\alpha$ is in $Sp(1)$ and calculate $\alpha'(t)$.

5. Let $H$ be a subgroup of a matrix group $G$. Show that $T_H$ is a linear subspace of $T_G$ so that $\dim H \leq \dim G$.

6. Show that the set $sp(n)$ of $n \times n$ skew-symmetric matrices is a real vector space and calculate its dimension.

7. Let $T$ be the set of upper triangular matrices in $M_n(\mathbb{R})$. That is, $A = (a_{ij}) \in T$ if and only if $a_{ij} = 0$ whenever $i > j$. Show that $T$ is a linear subspace of $M_n(\mathbb{R})$ and calculate its dimension. Show that $T$ is a subalgebra of $M_n(\mathbb{R})$ (i.e. show that $T$ is closed under multiplication). Show that $A \in T$ is nonsingular (i.e. is a unit) if and only if each $a_{ii} \neq 0$. (Note that the group $U$ defined in Exercise 9 of Chapter 2 is a subgroup of the group of units in the algebra $T$.)

Chapter 4
Exponential and Logarithm

4.1 Exponential of a matrix

Given a matrix group $G$ we have defined a vector space $T$ -- the tangent space to $G$ at $I$. In this chapter we develop maps to send $Y \in T$ and $G$ to $T$ and study their properties. We will work with real matrices -- developments for $C$ and $H$ are quite analogous. (We used these maps to determine dimensions of some of our matrix groups.)

**Definition:** Let $A$ be a real $n \times n$ matrix and set

$$e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \ldots$$

Where $A^n$ means the matrix product $AA$, etc. We say that this sequence converges if each of the $n^2$ real-number sequences

$$I_{ij} + (A)_{ij} + \left(\frac{A^2}{2!}\right)_{ij} + \left(\frac{A^3}{3!}\right)_{ij} + \ldots$$

converges.

**Proposition 1:** For any real $n \times n$ matrix $A$, the sequence

$$I + A + \frac{A^2}{2!} + \ldots$$
converges.

Proof: Let m be the largest |a_{ij}| in A. Then:

The biggest element in the first term is 1.
The biggest element in the second term is m.
The biggest element in the third term is \( \frac{m^2}{2!} \).
The biggest element in the fourth term is \( \frac{n^3}{3!} \), etc.

Any ij sequence is dominated by 1, m, \( \frac{m^2}{2!} \), \( \frac{m^3}{3!} \), ..., \( \frac{k-2}{k-1} \frac{k}{k!} \), ..., etc.

Applying the ratio test to this maximal sequence gives

\[
\lim_{k \to \infty} \frac{n^{k-1} \frac{m^k}{k!}}{n^{k-2} \frac{m^{k-1}}{(k-1)!}} = \frac{m}{k}.
\]

Since n and m are fixed, the ratio goes to 0 as k \( \to \infty \), proving (absolute) convergence.

This exponential behaves somewhat like the familiar \( e^x \) (x \( \in \mathbb{R} \))

For if 0 is the zero matrix, we have

\[ e^0 = I. \]

Also:

Proposition 2: If the matrices A and B commute, then

\[ e^{A+B} = e^A e^B. \]

Proof: We will just indicate a proof by looking at the first few terms.

\[ e^{A+B} = I + A + B + \frac{A^2}{2} + AB + \frac{B^2}{2} + \frac{A^3}{6} + \frac{A^2B}{2} + AB^2 + \frac{B^3}{6} + \ldots \]

\[ e^A e^B = (I + A + \frac{A^2}{2} + \frac{A^3}{6} + \ldots)(I + B + \frac{B^2}{2} + \frac{B^3}{6} + \ldots) \]

\[ = I + A + B + \frac{A^2}{2} + AB + \frac{B^2}{2} + \frac{A^3}{6} + \frac{A^2B}{2} + AB^2 + \frac{B^3}{6} + \ldots \]

Corollary: For any real n \( \times \) n matrix A, \( e^A \) is nonsingular.

Proof: A and \(-A\) commute, so \( I = e^0 = e^{A-A} = e^A e^{-A} \) and

\[ I = (\det e^A)(\det e^{-A}) \text{ and } \det e^A \neq 0. \]

From this corollary we see that the map \( \exp : M_n(\mathbb{R}) \rightarrow M_n(\mathbb{R}) \)

\( \exp (A) = e^A \), actually maps \( M_n(\mathbb{R}) \) into \( GL(n,\mathbb{R}) \).

Proposition 3: If A is a real skew-symmetric matrix, then

\( e^A \) is orthogonal.

Proof: We have \( I = e^0 = e^{A+e} = e^A e^e = (e^A)^t(e^e) \), proving

\( e^A \) is orthogonal.

So, if \( so(n) \subset M_n(\mathbb{R}) \) is the subspace of skew-symmetric matrices,

we see that

\[ \exp : so(n) \rightarrow s(n). \]

It is important to note two things which Proposition 3 does not say:

(i) It does not say that every orthogonal matrix is some \( e^A \)

with A skew-symmetric (i.e., it does not say \( \exp : so(n) \rightarrow s(n) \) is

injective), and (ii) it does not say that \( e^A \) orthogonal implies

A is skew-symmetric. It is instructive to examine the case \( n = 2 \)

in some detail.

The general 2 \( \times \) 2 real skew-symmetric matrix is of the form
\[
\alpha = \begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix}, \quad x \in \mathbb{R}.
\]

To calculate \( e^\alpha \), we calculate the powers of \( \alpha \).
\[
\alpha^2 = \begin{pmatrix} -x^2 & 0 \\ 0 & -x^2 \end{pmatrix}, \quad \alpha^3 = \begin{pmatrix} 0 & -x^3 \\ x^3 & 0 \end{pmatrix}, \quad \alpha^4 = \begin{pmatrix} x^4 & 0 \\ 0 & x^4 \end{pmatrix}, \quad \alpha^5 = \begin{pmatrix} 0 & x^5 \\ -x^5 & 0 \end{pmatrix}, \quad \text{etc.}
\]

Then
\[
e^\alpha = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{x^2}{2!} \begin{pmatrix} -x^2 & 0 \\ 0 & -x^2 \end{pmatrix} + \frac{x^4}{4!} \begin{pmatrix} 0 & -x^3 \\ x^3 & 0 \end{pmatrix} + \frac{x^6}{6!} \begin{pmatrix} x^4 & 0 \\ 0 & x^4 \end{pmatrix} + \ldots.
\]

From the \(1,1\) position we get
\[
1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots = \cos x,
\]

and
\[
\frac{x^2}{2!} - \frac{x^4}{4!} + \frac{x^6}{6!} + \ldots = \sin x,
\]

We find that
\[
e^\alpha = \begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix},
\]

which is a plane rotation of \( x \) radians. Thus for any real \( 2 \times 2 \) skew-symmetric matrix \( \alpha \) we have
\[
\det e^\alpha = 1, \quad \text{i.e.,} \quad e^\alpha \in SO(2).
\]

Thus, for example, the reflection \( \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \in SO(2) \) could never be obtained this way.

Note also that \( e^x = I \) does not imply that \( \alpha = \begin{pmatrix} 0 & 2\pi \\ -2\pi & 0 \end{pmatrix} \) has \( e^\alpha = I \).

We will see later that these results also hold for larger \( n \).

We conclude this section with a simple observation which is sometimes quite useful in computations.

**Proposition 4:** If \( A, B \) are \( n \times n \) matrices over \( k \in \{ \mathbb{R}, \mathbb{C}, \mathbb{H} \} \) and \( B \) is nonsingular, then
\[
e^{BAB^{-1}} = Be^{A}B^{-1}.
\]

**Proof:** \( (BAB^{-1})^n = (BAB^{-1})(BAB^{-1})\ldots(BAB^{-1}) = Bab^{-1} \), and \( \log CH = CB^{-1} + DB^{-1} \); these and the definition of the exponential of a matrix yield the result.

**Logarithm**

Just as \( e^x \) is defined for all \( x \in \mathbb{R} \) and \( \log x \) is defined only for \( x > 0 \), the logarithm of a matrix will be defined only for matrices near to the identity matrix \( I \).

Let \( X \) be a real \( n \times n \) matrix and set
\[
\log X = (X-I) - \frac{(X-I)^2}{2} + \frac{(X-I)^3}{3} - \frac{(X-I)^4}{4} + \ldots.
\]

**Proposition 5:** For \( X \) near \( I \) this series converges.

**Proof:** Let \( Y = X - I \) and \( Y = (y_{ij}) \), and suppose each \( y_{ij} \) is small. Then
\[
|y_{ij}|^2 \leq \varepsilon, \quad |\frac{y_{ij}^2}{2}|^3 \leq \frac{\varepsilon^2}{2},
\]

\[
|\frac{y_{ij}^3}{2}|^3 \leq \frac{\varepsilon^3}{3}, \quad |\frac{y_{ij}^k}{k}|^k \leq \frac{\varepsilon^k}{k}.
\]

The test gives
\[
\frac{\varepsilon k}{k+1} \leq \frac{k}{k+1} \varepsilon \leq \frac{\varepsilon}{k+1} n \varepsilon \rightarrow n \varepsilon.
\]

The series converges for any \( X \) such that each entry of \( X - I \) is small in magnitude.

**Proposition 6:** In \( M_n(\mathbb{R}) \) let \( U \) be a neighborhood of \( I \) on
which \( \log \) is defined and let \( V \) be a neighborhood of \( 0 \) such that \( \exp(V) \subset U \). Then

(i) for \( X \in U \), \( e^{\log X} = X \)

(ii) for \( A \in V \), \( \log e^A = A \).

**Proof:** We do (ii) first. \( A \in U = e^V \in U \) is defined (i.e., the series converges). \( e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \ldots \). So

\[
\log e^A = (A + \frac{A^2}{2!} + \ldots) - \frac{1}{2}(A + \frac{A^2}{2!} + \ldots)^2 + \frac{1}{3}(A + \frac{A^2}{2!} + \ldots)^3 + \ldots
\]

\[
= A + \left[ \frac{A^2}{2!} - \frac{A^2}{2} \right] + \left[ \frac{A^3}{3!} - \frac{A^3}{2} + \frac{A^3}{3} \right] + \ldots = A.
\]

(i) is similar. \( \log X = (X-I) - \frac{(X-I)^2}{2} + \frac{(X-I)^3}{3} - \ldots \).

\[
e^{\log X} = (I + (X-I) - \frac{(X-I)^2}{2} + \ldots) + \frac{1}{2!} \left( (X-I) - \frac{(X-I)^2}{2} + \ldots \right)^2
\]

\[
+ \frac{1}{3!} \left( (X-I) - \frac{(X-I)^2}{2} + \ldots \right)^3 + \ldots
\]

\[
= X - \frac{(X-I)^2}{2} + \frac{(X-I)^3}{2} - \frac{(X-I)^3}{3} - \frac{(X-I)^3}{2} + \frac{(X-I)^3}{6} + \ldots
\]

\[
= X.
\]

**Proposition 7:** If \( X \) and \( Y \) are near \( I \) and \( \log X \) and \( \log Y \) commute, then

\[
\log (XY) = \log X + \log Y.
\]

So if \( X \) is near \( I \) and orthogonal, \( \log X \) is skew-symmetric.

**Proof:** \( e^{\log XY} = XY = e^{\log X} e^{\log Y} = e^{\log X + \log Y} \), and \( e \) is one-to-one near \( 0 \).

Next \( X \) and \( e^X \) commute so that \( \log X \) and \( e^{\log X} \) commute. If \( X \) is orthogonal...
Example: Let $k \in \{R, C, H\}$ and $A \in M_n(k)$. Then

$$\gamma(u) = e^{uA} = I + uA + \frac{u^2}{2!}A^2 + \ldots$$

is a one-parameter subgroup of $GL(n,k)$ and $\gamma'(0) = A$.

Proposition 8: Let $\gamma$ be a one-parameter subgroup of $GL(n,k)$. Then $\exists A \in M_n(k)$ such that

$$\gamma(u) = e^{uA}.$$

Proof: Let $\sigma(u) = \log \gamma(u)$. Then $\sigma$ is a curve in $M_n(k)$ with

$$\gamma(u) = e^{\sigma(u)}.$$

Let $\sigma'(0) = A$. We just need to show that $\sigma(u)$ is a line through $0$ in $M_n(k)$, for then $\sigma(u) = uA$. Hold $u$ fixed.

$$\sigma'(u) = \lim_{v \to 0} \frac{\sigma(u+v) - \sigma(u)}{v} = \lim_{v \to 0} \frac{\log \gamma(u+v) - \log \gamma(u)}{v} = \lim_{v \to 0} \log(\gamma(u)\gamma(v))/\gamma(v) - \log(\gamma(u))/v.$$

Now $u + v = u + v$ and $\gamma$ is a one-parameter subgroup so that $\gamma(u)$ and $\gamma(v)$ commute. Thus

$$\log(\gamma(u)\gamma(v)) = \log(\gamma(u)) + \log(\gamma(v)).$$

So

$$\sigma'(u) = \lim_{v \to 0} \frac{\log \gamma(v)}{v} = \sigma'(0).$$

This proves that $\sigma'(u)$ is independent of $u$ so $\sigma(u)$ is indeed a line through $0$ in $M_n(k)$.

So any tangent vector to $GL(n,k)$ is the derivative at $0$ of some one-parameter subgroup. We will see now that this is also true for the orthogonal groups $O(n,k)$.

Proposition 9: Let $A$ be a tangent vector to $O(n,k)$. Then there exists a unique one-parameter subgroup $\gamma$ in $O(n,k)$ such that

$$A = \gamma'(0).$$

Proof: By definition $A = \rho'(0)$ where $\rho$ is a curve in $O(n,k)$. Then

$$\sigma(u)\frac{\sigma(u)}{\sigma(u)} = I$$

so that

$$\sigma'(0) + \frac{\sigma'(0)}{\sigma'(0)} = 0$$

i.e.,

$$A + \frac{A}{A} = 0.$$

$$\gamma(u) = e^{uA}$$

is a one-parameter subgroup of $GL(n,k)$, but it is in $O(n,k)$ because

$$\gamma(u)\frac{\gamma(u)}{\gamma(u)} = e^{uA}e^{\frac{A}{A}} = e^{u(A+\frac{A}{A})} = I.$$

This proves the proposition. So we have (for $GL(n,k)$ and $O(n,k)$) a one-to-one correspondence between tangent vectors and one-parameter subgroups.

Taking $k = R$ we have that the tangent space to $O(n) = O(n,R)$ in $o(n)$, the vector space of all skew-symmetric $n \times n$ matrices.

$$\dim o(n) = \dim so(n) = \frac{n(n-1)}{2}.$$n

Taking $k = C$ we have that the tangent space to $U(n) = O(n,C)$
is su(n), the vector space of all skew-Hermitian n x n complex matrices. Thus
\[ \dim U(n) - \dim su(n) = n^2. \]

Taking k = H we get
\[ \dim Sp(n) = n(2n+1). \]

What about the dimensions of SO(n) and SU(n)? We will see in Chapter VI (Proposition 3) that the tangent space to SO(n) is again so(n), so the dimension of SO(n) is also \( \frac{n(n-1)}{2} \). But the dimension of SU(n) is one less than the dimension of U(n). The proof of this must also be deferred to a later chapter, but we will indicate here the result on which it is based.

**Definition:** The **trace** of a matrix \( A = (a_{ij}) \) is the sum of the diagonal terms;
\[ \text{Tr}(A) = a_{11} + a_{22} + \ldots + a_{nn}. \]

We clearly have

1. \( \text{Tr}(A+B) = \text{Tr}(A) + \text{Tr}(B) \), and \( \text{Tr}(aA) = a \text{Tr}(A) \) (so \( \text{Tr} \) is linear).
2. \( \text{Tr}(AB) = \text{Tr}(BA) \).

Now suppose \( A = (a_{ij}) \) is real or complex. Then

To prove (ii) we just write it out. The sum of the diagonal terms in \( AB \) is
\[ (a_{11} b_{11} + \ldots + a_{1n} b_{n1}) + (a_{21} b_{12} + \ldots + a_{2n} b_{n2}) + \ldots + (a_{n1} b_{1n} + \ldots + a_{nn} b_{nn}). \]

and the sum of the diagonal terms in \( BA \) is
\[ (b_{11} a_{11} + \ldots + b_{1n} a_{n1}) + (b_{21} a_{12} + \ldots + b_{2n} a_{n2}) + \ldots + (b_{n1} a_{1n} + \ldots + b_{nn} a_{nn}). \]

Since \( a \) and \( c \) are commutative one easily checks that these are equal.

Clearly

(iii) \( \text{Tr}(I) = n \).

(iv) If \( B \) is nonsingular, then
\[ \text{Tr}(BAB^{-1}) = \text{Tr}(A). \]

**Proof:** By (ii), \( \text{Tr}(B(AB^{-1})) = \text{Tr}((AB^{-1})B) = \text{Tr}(AB^{-1}) = \text{Tr}A \).

Now we come to the crucial relation.

**Theorem:** If \( A \) is a real or complex matrix, then
\[ e^{\text{Tr}(A)} = \det (e^A). \]

We will prove this later, but a few comments are in order here.

Note off. (iv) looks wrong because the left hand side depends only on the diagonal elements of \( A \), and it is not immediately clear that it is true for the right-hand side. The point is that \( \det \) and \( e^{\text{Tr}} \) are also invariant under conjugation just as (iv) for \( \text{Tr} \); so if \( B \) is nonsingular
\[ \det (e^{BAB^{-1}}) = \det (Be^A B^{-1}) = \det (e^A). \]

We will prove (iv) once we have found how to put matrices in simpler form by conjugation (in Chapter VIII).
Suppose we know (+). The linear map

$$\text{Tr} : u(n) \to \mathbb{C}$$

actually maps into $i\mathbb{R} \subset \mathbb{C}$ since all diagonal terms in a skew-Hermitian matrix are purely imaginary. It is easy to see that $\text{Tr}(u(n))$ all of $i\mathbb{R}$. From the rank theorem in linear algebra we know that (all vector spaces being over $\mathbb{R}$ now)

$$\dim u(n) = \dim \text{Tr}(u(n)) + \dim \text{Tr}^{-1}(0).$$

Thus the dimension of $\text{Tr}^{-1}(0)$ is just one less than $\dim u(n)$ i.e. $n^2 - 1$. But $su(n) = \text{Tr}^{-1}(0)$ is just the tangent space of $SU(n)$, since

$$\text{Tr} (C) = 0 \implies 1 = e^{\text{Tr}(C)} = \det e^{C} = e^{C} \in SU(n).$$

D. Lie algebras

It is easy to see that $so(n)$, $su(n)$ and $sp(n)$ are not closed under matrix multiplication. For example, if

$$a = \begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix} \quad \text{then} \quad a^2 = \begin{pmatrix} -x^2 & 0 \\ 0 & -x^2 \end{pmatrix},$$

which is not skew-symmetric.

Proposition 10: For $k \in \{ \mathbb{R}, \mathbb{C}, \mathbb{H} \}$ and $A,B \in M_n(k)$ we define

$$[A,B] = AB - BA.$$
\[ [x,y] = x[1,1] = xy[1,1] = 0 \quad \text{(by (i))}. \]

So we have the trivial product (which obviously satisfies (i)...(iv)).

Consider \( \mathbb{R}^2 \) with basis \( e_1, e_2 \).

We must have

\[ [e_1, e_1] = 0, \quad [e_2, e_2] = 0 \quad \text{and} \quad [e_1, e_2] = -[e_2, e_1]. \]

Let \( [e_1, e_2] = ae_1 + be_2 \). Then, for example,

\[
[e_1, [e_1, e_2]] = [e_1, (ae_1 + be_2)] = a[e_1, e_1] + b(e_1, e_2) = b(ae_1 + be_2).
\]

By the Jacobi identity

\[
[e_1, [e_1, e_2]] + [e_2, [e_2, e_1]] + [e_2, [e_1, e_1]] = 0,
\]

so

\[ b(ae_1 + be_2) + [e_1, (-ae_1 - be_2)] = 0 \]

which is true with no conditions on \( a, b \). If we take \( a = 0 = b \) we get the trivial Lie algebra. For any other choice we get a nontrivial Lie algebra. In the exercises one shows that these nontrivial 2-dimensional Lie algebras are all "essentially the same."

We will not try to find out all nontrivial 3-dimensional Lie algebras, but will simply look at two which arise quite naturally.

\[
\text{so}(3) = \left\{ \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix} \mid a, b, c \in \mathbb{R} \right\}
\]

clearly has dimension three. Also using basis

\[ [i,j] = k \]
\[ [j,k] = i \]
\[ [k,i] = j \]

gives a 3-dimensional Lie algebra.

8. Exercises

1. Let \( A \) be a \( 3 \times 3 \) skew-symmetric matrix. Show that \( A^2 \) is skew-symmetric, but show by example that \( A^3 \) could be neither symmetric nor skew-symmetric.

2. Let \( B \in \text{SU}(3) \). Show that the series for \( \log B \) does not converge.

3. Prove the Jacobi identity for \( [A,B] = AB - BA \).

5. a. Prove that any two nontrivial Lie algebras on \( \mathbb{R}^2 \) are isomorphic as Lie algebras.

b. Show that the two 3-dimensional Lie algebras defined above are isomorphic.
A. The homomorphism $\phi : S^3 \to SO(3)$

We have seen that $Sp(1)$, which is all quaternions of unit length, is just the unit 3-sphere in $\mathbb{H}$ ($= \mathbb{H}$). Also we have seen that $\dim SO(3) = \frac{3 \cdot 2}{2} = 3$. So dimension won't distinguish $S^3$ from $SO(3)$, and, for all we know now, they might be isomorphic. In this section we define and study an "almost isomorphism" between them.

**Proposition 1:** If $q \in S^3$, then the "left translation"

$$L_q : \mathbb{H} \to \mathbb{H}$$

given by $L_q(q') = qq'$ is an orthogonal map of $\mathbb{H}$ to $\mathbb{H}$.

**Proof:** As vector spaces over $\mathbb{R}$, $\mathbb{H}$ and $\mathbb{H}$ are the same. So $L_q$ is surely a linear map of $\mathbb{H}$, for if $a,b \in \mathbb{R}$ and $c,d \in \mathbb{H}$ we have

$$L_q(a + bi + cj + dk) = q(a + bi + cj + dk) = aq + bq + c\bar{q} + dq = aL_q(q) + bL_q(q).$$

To see that $L_q$ is orthogonal, it suffices to show that $L_q$ preserves the perpendicularity (using $\langle \cdot, \cdot \rangle$ for $\mathbb{H}$) of the four unit vectors $1, i, j, k$. For example, let $q = a + ib + jc + kd$ and calculate $\langle L_q(1), L_q(i) \rangle$ (using $\langle \cdot, \cdot \rangle$ for $\mathbb{H}$). We get

$$a + bc - bd - ca = 0.$$ For $(L_q(1), L_q(i))$ we get

$$a + ib + jc + kd, ai - b - kc + jd = -ab + ad + dc - dc = 0.$$ The computations for other pairs of basis vectors are similar.

**Definition of $\phi$:** For $q \in S^3$ and $a \in \mathbb{H}$ we define

$$\phi(q)(a) = qa\bar{q}.$$ 

**Note:** we do a left translation by $q$ and a right translation by $\bar{q}$. By Proposition 1 this is an orthogonal map of $\mathbb{H}$ to $\mathbb{H}$; i.e.,

$$\phi(q)(a) = a(q).$$

Since real quaternions commute with all other quaternions, if $x$ is a real quaternion

$$\phi(q)x = qx\bar{q} = x\bar{q} = x.$$ 

It is also that $\phi(\bar{q})$ is the inverse of $\phi(q)$ in the group $\phi(4)$ where $\phi(q)\phi(q^{-1})(a) = q\phi(a)\bar{q} = a$ and similarly for $\phi(\bar{q})\phi(q)$.

Together these two observations imply that $\phi(q)$ maps the 3-space spanned by $1, i, j, k$ to itself (Exercise #3). Thus $\phi(q)$ can be considered as an element of $SO(3)$ (Exercise #4).

**Fact (to be proved after Chapter VI):** $\phi(q)$ is in $SO(3)$. 

**Proposition 2:** $\phi : S^3 \to SO(3)$ is a surjective homomorphism and

$$\ker(\phi) = \{1, -1\} \subset S^3.$$ 

**Proof:** If $q_1, q_2 \in S^3$ and $a \in \text{Span}(1, i, j, k)$, then
\( \rho(q_1q_2)(a) = q_1q_2\overline{q_1q_2} = q_1(q_2\overline{q_2})\overline{q} = \overline{\rho(q_1)\rho(q_2)(a)}. \)

Thus \( \rho \) is a homomorphism.

Clearly \( \rho(1) \) and \( \rho(-1) \) are the identity in \( \text{SO}(3) \) so that 1 and \(-1\) are in \( \text{Ker} \, \rho \). Conversely, suppose \( \rho(q) \) is the identity with \( q = a + ib + jc + kd \). Then \( \rho(q)(1) = 1 \) gives
\[
(a + ib + jc + kd)(1)(a - ib - jc - kd) = 1. \quad \text{And from this we get}
\]
\[
a^2 + b^2 - c^2 - d^2 = 1. \quad \text{But} \quad a^2 + b^2 + c^2 + d^2 = 1 \quad \text{and we conclude that} \quad c = 0 - d. \quad \text{From} \quad \rho(q)j = j \quad \text{we get} \quad b = 0. \quad \text{Then} \quad a^2 = 1 \quad \text{so} \quad a \in \{1, -1\}.

Finally we need to show that \( \rho \) is surjective. This will be quite easy once we know some topology (Chapter VII) -- otherwise it is an almost hopelessly complicated computation. Here we will just show that we can find a \( q \in S^3 \) such that \( \rho(q) \) is the element of \( \text{SO}(3) \) which leaves \( k \) fixed, sends \( i \) to \( j \) and sends \( j \) to \( -i \).

Let \( q = a + ib + jc + kd \). We want
\[
(a + ib + jc + kd)(k)(a - ib - jc - kd) = k, \quad \text{or}
\]
\[
(a + ib + jc + kd)(ka - jb + ic + d) = k, \quad \text{so}
\]
\[
ad - bc + bc - ad = 0 \quad (\text{automatically}), \]
\[
ac + bd + ac + bd = 0 \quad \text{or} \quad 2(ac + bd) = 0, \]
\[
-ab + cd + dc - ab \quad \text{or} \quad 2(cd - ab) = 0. \]
\[
a^2 + d^2 - b^2 - c^2 = 1. \]
\[
a^2 + d^2 + b^2 + c^2 = 1.
\]

Now
\[
2(b^2 + c^2) = 0 \quad \text{or} \quad b = 0 = c.
\]

It's the only condition on \( q \) such that \( \rho(q)k = k \) is that
\[
q = a + dk \quad (\text{with} \quad a^2 + c^2 = 1).
\]

Next we want \( \rho(q)j = j \).
\[
(a + kd)(j)(a - kd) = j
\]
\[
(a + kd)(ja + jd) = j
\]
\[
a^2 - d^2 = 0, \quad a = \pm d. \quad \text{ad + ad = 1, if a = d, 2a^2 = 1,}
\]

and we can't have \( a = -d \). Finally we insist that \( \rho(q)j = -i \). So
\[
(a + ka)(j)(a - ka) = -i
\]
\[
(a + ka)(ja - ia) = -1
\]
\[
-2a^2 = -1 \quad \text{or} \quad a^2 = \frac{1}{2}. \quad \text{Both will give the desired element of} \quad \text{SO}(3). \quad (\text{This should be enough to convince us that we should not try the general proof of surjectivity at this stage.)}
\]

Note that this does not prove that \( S^3 \) and \( \text{SO}(3) \) are not isomorphic. \( \rho \) is not an isomorphism, but one might exist. In the next section we give a fairly easy proof that \( S^3 \neq \text{SO}(3) \).
In Exercise #4 of Chapter I the center \( C \) of a group \( G \) is defined as

\[
C = \{ x \in G \mid xy = yx \text{ for all } y \in G \},
\]

and was shown to be an abelian and normal subgroup of \( G \). We leave it as an exercise here to show that any isomorphisms of groups induce an isomorphism of their centers. We will show that \( S^3 \neq SO(3) \) by showing that their centers are not isomorphic.

**Proposition 3:** The center of \( S^3 = Sp(1) \) is \{1, -1\}, whereas the center of \( SO(3) \) is \{1\}.

**Proof:** Since real quaternions commute with all quaternions, it is clear that \{1, -1\} \subseteq Center \( S^3 \). Conversely, suppose \( q = a + ib + jc + kd \in S^3 \) is in the center. Then \( q^2 = i q \) gives

\[
a^2 - b^2 + c^2 + d^2 = a^2 - b^2 + c^2 - d^2
\]

so that \( c = 0 = d \). Then \( qj = jq \) gives

\[
(a + ib)j = j(a + ib)
\]

and this implies \( b = 0 \). So \( q = a \) and \( a^2 = 1 \). Thus Center \( S^3 = \{1, -1\} \).

Suppose \( A \in SO(3) \) is in the center. Since \( A \) commutes with all elements of \( SO(3) \) it surely commutes with all elements of

\[
T = \begin{pmatrix}
\cos \theta & \sin \theta & 0 \\
\sin \theta & \cos \theta & 0 \\
0 & 0 & 1
\end{pmatrix}
\]

since \( T \in SO(3) \). Consider the standard basis \( e_1 = (1,0,0), e_2 = (0,1,0), e_3 = (0,0,1) \) for \( \mathbb{R}^3 \).

**Claim:** \( A \) leaves \( e_3 \) fixed (or sends it to \( -e_3 \)).

**Cause:** \( B \in T \) which sends \( e_1 \) to \( e_2 \) and \( e_2 \) to \( e_1 \) and (automatically) leaves \( e_3 \) fixed. Then set \( Ae_3 = ae_1 + be_2 + ce_3 \). Then

\[
Ae_3 = ae_2 - be_1 + ce_3,
\]

whereas \( ABe_3 = Ae_3 \); this implies \( a = 0 \). and since \( A \) preserves length, we must have \( c = 1 \), or \( c = -1 \).

Thus \( A \) induces an orthogonal map of the \( e_1, e_2 \) plane. Actually, it is a rotation because:

**Sublemma:** Any element of \( \mathfrak{so}(2) \) which commutes with all rotations, is itself a rotation.

Let \( \varphi : \mathbb{R}^2 \to \mathbb{R}^2 \) denote such an element of \( \mathfrak{so}(2) \). For any rotation

\[
\varphi = \begin{pmatrix}
\cos \theta & \sin \theta \\
\sin \theta & -\cos \theta
\end{pmatrix}
\]

we must have \( \varphi t = t \varphi \). Let \( \varphi = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \). We get

\[
\alpha \cos \theta - \beta \cos \delta = \alpha \cos \theta + \gamma \sin \delta
\]

\[
\alpha \sin \theta + \beta \cos \delta = \beta \cos \theta + \delta \sin \theta
\]

holding for all \( \theta \). So \( \gamma = -\beta \) and \( \alpha = \delta \). Thus

\[
\varphi = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}
\]

and \( \det \varphi = \alpha^2 + \beta^2 \).
Since this cannot equal \(-1\) (and must be in \(\{1,-1\}\)), this proves the sublemma.

This also proves that \(c = 1\) (not \(-1\)) (since \(A \in SO(3)\)) and we conclude that

\[
A \in T. 
\]

We can now finish the proof.

\[
A = \begin{pmatrix}
\cos \theta & \sin \theta & 0 \\
-\sin \theta & \cos \theta & 0 \\
0 & 0 & 1
\end{pmatrix}
\]

and we let

\[
R = \begin{pmatrix}
0 & 0 & 1 \\
0 & 1 & 0 \\
-1 & 0 & 0
\end{pmatrix} \in SO(3). 
\]

Since \(A\) must commute with \(R\) we get

\[
AR = \begin{pmatrix}
0 & \sin \theta & \cos \theta \\
0 & \cos \theta & -\sin \theta \\
-1 & 0 & 0
\end{pmatrix} = RA.
\]

Thus we must have \(\cos \theta = 1\) and \(\sin \theta = 0\). Thus \(A = I\) and Proposition 3 is proved.

We will calculate the centers of all of the groups \(SO(n)\), \(U(n)\), \(SU(n)\), \(Sp(n)\) in a later chapter, after we know about maximal tori. We conclude this chapter with a bit more abstract theory which we will need later.

---

1. **Quotient groups**

If \(H\) is a subgroup of \(G\) we define an equivalence relation on \(G\) by

\[
x \sim y \text{ if } xy^{-1} \in H.
\]

This relation is **reflexive**, \(x \sim x\), since \(xx^{-1} = e \in H\). It is **symmetric**, \(x \sim y \iff y \sim x\), since \(xy^{-1} e H = (xy^{-1})^{-1} = yx^{-1} \in H\). It is **transitive**, \(x \sim y\) and \(y \sim z \implies x \sim z\), since \(xy^{-1} e H\) and \(yz^{-1} e H\) imply that \((xy^{-1})(yz^{-1}) = xz^{-1} e H\). Thus \(\sim\) divides \(G\) into equivalence classes.

Let \(C(x)\) denote the class containing \(x\). Then

\[
C(x) = Hx = \{hx \mid h \in H\}. 
\]

Also \(Hx = Hy \iff xy^{-1} \in H \iff y \in C(x) = x \in C(y)\). These equivalence classes are called **right cosets** of \(H\).

**Example:** Let \(G = S^3 \quad (= Sp(1))\) and \(H = \{1,-1\}\). Then

\[
S = \{q, -q\} = H(-q) \quad \text{so each equivalence class contains exactly two}
\]

points of \(S^3\).

**Example:** In \(G = U(3)\) let

\[
H = \{\lambda I \mid \lambda \text{ a complex number of unit length}\}. 
\]

Then \(H\) is a circle subgroup of \(G\) and the right cosets are circles.

Thus \(U(3)\) can be divided into disjoint circles filling up \(U(3)\).

Similarly, let \(G = S^3 = Sp(1)\) and let \(H\) be the circle

\[
\{a^2 + b^2 = 1\}. 
\]

Thus \(S^3\) can be divided up into circles.

One defines left cosets in a similar manner.
Recall (Exercise #3, Chapter I) that a subgroup $H$ is normal if $xHx^{-1} = H$ for all $x \in G$.

**Observation:** A subgroup $H$ of $G$ is normal (in $G$) if $xHx^{-1} = H$ for every $x \in G$.

Let $G/H$ denote the set whose elements are the right cosets of $H$ in $G$.

**Proposition 4:** If $H$ is a normal subgroup of $G$, then the operation on $G/H$ defined by

$$(hx)(hy) = h(xy)$$

makes $G/H$ into a group.

**Proof:** We need $H$ normal to show the operation on $G/H$ is well-defined. Suppose $hx = hx'$ and $hy = hy'$. We must show that $hxy = hx'y$. Well, $xy(zw)^{-1} = xzw^{-1}y^{-1}$ and $yw^{-1} = h_1 \in H$. Also, $z^{-1} = x^{-1}h_2$. So

$$xy(zw)^{-1} = xh_1x^{-1}h_2$$

and, since $H$ is normal, $xh_1x^{-1} = h_3 \in H$ so that

$$xy(zw)^{-1} = h_3h_2 \in H,$$

and we have proved that the operation is well defined.

The rest is easy. $H = He \in G/H$ is the identity and $Hx^{-1}$ is the inverse of $Hx$. (Associativity is inherited from $G = (Hx)(HyHz) = (HxHy)Hz$ since $x(yz) = (xy)z$).

Example: $G = Sp(1)$ and $H = \{1,-1\}$. $H$ is the center of $G$ and thus is a normal subgroup. Thus $G/H$ is a group. We know it is $SO(3)$.

There is a natural map $\eta: G \to G/H$ given by $\eta(x) = Hx$. In the exercises it is shown that $\eta$ is a surjective homomorphism with kernel $H$.

Let $G$ be a group and $x, y \in G$. Then the element

$$xyx^{-1}y^{-1}$$

is called the commutator of $x$ and $y$ (because $(xyx^{-1}y^{-1})(yx) = xy$).

Now the product of two commutators is not necessarily a commutator, but we set $[G,G] = \{all$ finite products of commutators\}.

**Proposition 5:** $[G,G]$ is a normal subgroup of $G$ and $\frac{G}{[G,G]}$ is an abelian group.

**Proof:** Closure and identity are clear and

$$(xyx^{-1}y^{-1})(yxy^{-1}x^{-1}) = e,$$

showing $[G,G]$ is a subgroup. Let $z \in G$ and $xyx^{-1}y^{-1} \in [G,G]$. Then

$$z(xyx^{-1}y^{-1})z^{-1} = zxy(z^{-1}(xy)^{-1}(xy)z)x^{-1}((yz)^{-1}(yz))y^{-1}z^{-1}$$

$$= (z(xy)x^{-1}(yz)x^{-1}((yz)^{-1}(yz))y^{-1}z^{-1}) \in [G,G].$$

This easily extends to products of commutators, so that $[G,G]$ is a normal subgroup.

\( xy(yx)^{-1} = xyx^{-1}y^{-1} \in [G,G] \).

q.e.d.

In most instances we will encounter, if \( G \) is a matrix group and \( C \) is its center, then \( G/C \) will have trivial center. But this need not always be the case.

**Proposition 6:** For \( x \in G \) define

\[ i(x) : G \rightarrow G \]

by \( i(x)(y) = xyx^{-1}y^{-1} \). Then \( G/C \) has nontrivial center \( gx \in G - C \) such that

\[ i(x)(g) \in C. \]

**Proof:**

\( x \notin G = Cx \neq C \) so \( Cx \) is not the identity in \( G/C \). But for any \( y \in G \) we have \( xyx^{-1}y^{-1} \in C \) so that \( CxCy = Cxy = Cyx = CyCx \) and \( Cx \in \text{center } G/C \).

Conversely, \( Cx \neq C \) with \( Cx \) in the center implies

\( CxCy = Cxy = Cyx = CyCx \) so that \( xyx^{-1}y^{-1} \in C \) for all \( y \in C \).

Once we have done a little topology (Chapter VI) we easily have:

**Corollary:** If \( G \) is connected and \( G \) is discrete (in particular if \( G \\text{ is finite} \) ), then \( G/C \) has no center.

### Exercises

1. **Let** \( G \) be a group and \( x \in G \). Show that left translation \( L_x : G \rightarrow G \) by \( x \) \( (L_x(g) = xg) \) is a one-to-one map of \( G \) onto \( G \).

Let \( R_x \) be right translation so that

\[ R_x \circ L_x(g) = xgx^{-1}. \]

Show that \( R_x \circ L_x \) is an isomorphism of \( G \) onto \( G \).

2. Do one more step in the proof of Proposition 1 by showing

\[ \langle q_1, L_q(k) \rangle = 0. \]

3, 4. These are listed by number in the text.

5. Show that \( \phi(1), \phi(j), \phi(k) \) are all in \( \text{SO}(3) \).

6. Show that the set \( T \) defined in the proof of Proposition 3 is an abelian subgroup of \( \text{SO}(3) \).

7. Let \( \phi : G \rightarrow K \) be a surjective homomorphism of groups and \( H = \text{Ker } \phi \). Then we have

\[ G = H \oplus K \]

\[ \phi \circ \eta^{-1} \]

\[ G/H \]

Show that \( \phi \circ \eta^{-1} \) is well defined and gives an isomorphism of \( G/H \) onto \( K \).

8. Show that (see Exercise #6) the abelian subgroup \( T \) of \( \text{SO}(3) \) is not a normal subgroup.

9. Show that the subgroup \( H = \{ a + ib \mid a^2 + b^2 = 1 \} \) of \( \text{Sp}(1) \)
is not a normal subgroup.

10. Show an isomorphism of groups induces an isomorphism of their centers.

Chapter 6
Topology

6. Introduction

Our matrix groups are all subsets of euclidean spaces, because they are all subsets of

\[ M_n(\mathbb{R}) = \mathbb{R}^{n^2} \quad \text{or} \quad M_n(\mathbb{C}) = \mathbb{C}^{n^2} \quad \text{or} \quad M_n(\mathbb{H}) = \mathbb{H}^{4n^2} \,.

There are certain topological properties, notably connectedness and compactness, which some of our groups have and others do not. These properties are preserved by continuous maps and so are surely invariants under isomorphisms of groups. So a connected matrix group could not be isomorphic with a nonconnected matrix group, and a similar statement holds for compactness. We will define these properties and decide which of our groups have them. This will be done in sections A and C.

In section D we define and discuss the notion of a countable basis for open sets, a concept we will need in our study of maximal tori in matrix groups. Finally, in section E we define manifold and show that all of our matrix groups are manifolds. Then we prove a theorem about manifolds which gives an easy proof that the homomorphism \( \varphi : \text{Sp}(1) \to \text{SO}(3) \) (defined in Chapter V) is surjective.
B. Continuity of functions, open sets, closed sets

Definition: A metric \( d \) on a set \( S \) is a way of assigning to each \( x, y \in S \) a real number \( d(x, y) \) (the distance from \( x \) to \( y \)) in such a way that:

(i) \( d(x, y) \geq 0 \) and \( d(x, y) = 0 \iff x = y \),
(ii) \( d(x, y) = d(y, x) \),
(iii) \( d(x, y) + d(y, z) \geq d(x, z) \).

Condition (iii) is called the triangle inequality.

![Diagram](image)

We will define such a metric \( d \) on \( \mathbb{R}^n \) and then for any \( S \subseteq \mathbb{R}^n \), will also clearly be a metric on \( S \). Recall that for \( x, y \in \mathbb{R}^n \),

\[
x = (x_1, \ldots, x_n) \\
y = (y_1, \ldots, y_n)
\]

defined an inner product

\[
\langle x, y \rangle = x_1y_1 + \ldots + x_ny_n
\]
distance function

\[
d(x, y) = \sqrt{\langle x-y, x-y \rangle}.
\]

(Thus we define \( d(x, y) \) to be the length of the vector \( x - y \).)

Proposition 1: This is a metric on \( \mathbb{R}^n \).

Proof: Properties (i) and (ii) follow from

\[
\langle x, x \rangle \geq 0 \quad \text{and} \quad \langle x, x \rangle = 0 \iff x = 0
\]

and symmetry of the inner product. To prove the triangle inequality we will prove the corresponding property of \( \langle , \rangle \) called the Schwarz inequality.

For any \( x, y \in \mathbb{R}^n \) and \( t \in \mathbb{R} \) we have

\[
\langle tx + ty, tx + ty \rangle \geq 0.
\]

Using the bilinearity and symmetry of \( \langle , \rangle \) this gives

\[
\langle x, x \rangle + 2\langle x, y \rangle t + \langle y, y \rangle t^2 \geq 0.
\]

This quadratic polynomial in \( t \) with real coefficients is always \( \geq 0 \) and thus it cannot have two distinct real roots. (A quadratic polynomial can have only one minimum.) Thus the discriminant cannot be positive; i.e.,

\[
(2\langle x, y \rangle)^2 - 4\langle y, y \rangle \langle x, x \rangle \leq 0.
\]

So

\[
\langle x, y \rangle^2 \leq \langle x, x \rangle \langle y, y \rangle.
\]

The inequality (*) is the Schwarz inequality.

We apply (*) to the vectors \( x - y \) and \( y - z \) to get

\[
\langle x-y, y-z \rangle \leq \sqrt{\langle x-y, x-y \rangle} \sqrt{\langle y-z, y-z \rangle}.
\]

If we square both sides of (iii), write it in terms of \( \langle , \rangle \) and use basic properties of \( \langle , \rangle \), we see that (iii) is equivalent to (**).

We use this metric \( d \) on \( \mathbb{R}^n \) to define open balls. Let \( x \in \mathbb{R}^n \) and \( r > 0 \) be a real number. Set

\[
B(x, r) = \{ y \in \mathbb{R}^n \mid d(x, y) < r \}.
\]
and call this the open ball with center \( x \) and radius \( r \). Open balls in euclidean spaces allow us to give a fairly direct generalization of the notion of continuity of a function on \( \mathbb{R} \) to functions defined on sets of dimension greater than one.

Let \( A \) be a subset of \( \mathbb{R}^n \) and

\[ f : A \to \mathbb{R}^m \]

a function defined on \( A \) and taking values in some euclidean space.

**Definition:** To say \( f \) is continuous at a point \( a \in A \) means:

Given any open ball \( B(f(a), \varepsilon) \) in \( \mathbb{R}^m \), there exists an open ball \( B(a, \delta) \) in \( \mathbb{R}^n \) such that any point \( x \in A \cap B(a, \delta) \) satisfies

\[ f(x) \in B(f(a), \varepsilon) . \]

Another way of saying this is: Given \( \varepsilon > 0 \) there exists \( \delta > 0 \) such that if \( x \in A \) satisfies \( d(a, x) < \delta \), then \( f(x) \) satisfies \( d(f(a), f(x)) < \varepsilon . \)

Both ways are just precise ways of saying that \( f \) is continuous and sends "nearby points" of \( A \) to "nearby points" in \( \mathbb{R}^m \).

It is important to notice that the continuity of \( f \) depends on the domain \( A \) of definition. For example define

\[ f : \mathbb{R} \to \mathbb{R} \]

\[ f(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}. \]

\( f \) is not continuous at 0. But suppose we take \( A = \mathbb{R} \) to be all \( x \geq 0 \) and restrict \( f \) to \( A \)

\[ f : A \to \mathbb{R}. \]

Then this restricted \( f \) is continuous at 0.

A function \( f : A \to \mathbb{R}^m \) (\( A \subset \mathbb{R}^n \)) is said to be continuous if it is continuous at each \( a \in A \).

**Example:** If \( A \subset \mathbb{R}^n \) is a finite set then any \( f : A \to \mathbb{R}^m \) is continuous.

**Proof:** For any \( a \in A \) let \( b_1, \ldots, b_k \) be all of the other points of \( A \). Let

\[ \varepsilon_i = d(a, b_i) , \quad i = 1, \ldots, k \]

and let \( \delta \) be the smallest of these. Then for any \( \varepsilon > 0 \) any element of \( A \) in \( B(a, \delta) \) goes into \( B(f(a), \varepsilon) \) (because \( a \) is the only such element of \( A \)).

**Proposition 2:** If \( A \subset \mathbb{R}^n \) and

\[ A \subset \mathbb{R}^n , \quad f(A) \subset \mathbb{R}^p \]

are continuous, then \( g \circ f \) is continuous.

**Proof:** Let \( a \in A \) and \( \varepsilon > 0 \). Since \( g \) is continuous, there exists \( \tau > 0 \) such that every element of \( f(A) \) in \( B(f(a), \tau) \) is sent by \( g \) into \( B(g(f(a)), \varepsilon) \). Since \( f \) is continuous there exists \( \delta > 0 \) such that every element of \( A \) in \( B(a, \delta) \) is sent by \( f \) into \( B(f(a), \tau) \) and is then sent by \( g \) into \( B(g(f(a)), \varepsilon) \).

q.e.d.
Some exercises on continuity are given at the end of this chapter.

**Definition:** A set $U \subset \mathbb{R}^n$ is an open set if each $x \in U$ lies some $B(x,r) \subset U$ (where $r$ will depend on $x$).

Clearly $\mathbb{R}$ is an open set. It is not quite so clear that the empty set $\emptyset$ is open; but since there is no $x \in \emptyset$, there is no requirement that some $B(x,r)$ be contained in $\emptyset$.

**Example:** Any open ball $B(y,s)$ is an open set. Let $x \in B(y,s)$, $d(x,y) < s$. We must find $r > 0$ such that $B(x,r) \subset B(y,s)$.

Let $r = s - d(x,y)$. If $z \in B(x,r)$ then $d(z,x) < s - d(x,y)$

thus

$$d(z,x) + d(z,y) < s.$$ \[\text{using \ the \ triangle \ inequality \ we \ have} \]

$$d(z,y) \leq d(z,x) + d(x,y) < s$$

when $z \in B(y,s)$.

**Example:** $(0,1) = \{x \in \mathbb{R} | 0 < x < 1\}$ is an open set in $\mathbb{R}$ because it is the open ball $B(\frac{1}{2}, \frac{1}{2})$.

But

$(0,1) = \{x \in \mathbb{R} | 0 < x < 1\}$ is not open in $\mathbb{R}$ because $1 \in (0,1)$ but no $B(1,r)$ lies in $(0,1)$ since every such ball contains numbers greater than 1.

**Definition:** A subset $C \subset \mathbb{R}^n$ is defined to be closed if its complement $\mathbb{R}^n - C$ is open.

$(0,1) \subset \mathbb{R}$ is neither open nor closed. We have seen it is not open. Let $T = \mathbb{R} - (0,1)$. Then $0 \in T$ but no $B(0,r)$ can lie in $T$ since each will contain points of $(0,1)$. Thus $T$ is not open so that $(0,1)$ is not closed.

$$[0,1] = \{x \in \mathbb{R} | 0 \leq x \leq 1\}$$ is a closed set.

**Example:** Let $K \subset \mathbb{R}^n$ be a finite set. Then $K$ is closed. For if $x \in \mathbb{R}^n - K$ there will be a minimum distance $\delta$ from $x$ to points of $K$. Then

$$B(x,\delta) \subset \mathbb{R}^n - K,$$

proving $K$ is closed.

**C. Connected sets, compact sets**

**Definition:** A set $D$ in $\mathbb{R}^n$ is connected if: Given $x, y \in D$ there exists a continuous function

$$\gamma : [0,1] \to D$$

(i.e., $\gamma : [0,1] \to \mathbb{R}^n$ with $\gamma([0,1]) \subset D$)

with $\gamma(0) = x$ and $\gamma(1) = y$. 

Such a function may be called a **path** from $x$ to $y$ in $D$.

### Examples:

$n$ is connected because 

$$
\gamma(t) = (x + t(y-x))
$$

Path in $\mathbb{R}^n$ from $x$ to $y$.

$$
\{x \in \mathbb{R} | x \neq 0\}
$$

is not connected. For example, no path from $-1$ to $1$ in $\mathbb{R}$ can lie in $D$.

$$
(x_1, x_2) \in \mathbb{R}^2 \setminus \{(0,0)\}
$$

is connected.

### Important Example:

$\mathfrak{g}(n) \subset \mathbb{R}^{n^2}$ is not connected. The matrices

$$
\begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix}
$$

and

$$
\begin{pmatrix}
1 & 0 \\
0 & -1
\end{pmatrix}
$$

are in $\mathfrak{g}(n)$. If $\gamma : [0,1]$ is a path from $A$ to $I$, then the composite function $[0,1] \ni t \mapsto \det \gamma(t)$ is continuous (Proposition 2). But it would be a path in $\mathbb{R}$ from $-1$ to $1$, contradicting the existence of $\gamma$.

Recall that $\mathfrak{so}(n) \subset M_n(\mathbb{R})$ consists of skew-symmetric matrices that if $A \in \mathfrak{so}(n)$, then $\exp A \in \mathfrak{so}(n)$.

### Proposition 3:

$\exp$ maps $\mathfrak{so}(n)$ into $\text{SO}(n)$.

**Proof:** For $B \in \mathfrak{so}(n)$ the path

$$
\gamma(t) = e^{tB}
$$

path from $e^0 = I$ to $e^B$. As seen above, this implies that $e^B = A$ so $e^B \in \text{SO}(n)$.

### Proposition 4:

Let $D \subset \mathbb{R}^n$ be connected and

$$
f : D \to \mathbb{R}^n
$$

be continuous, then $f(D)$ is connected.

**Proof:** Given $a, b \in f(D)$, choose $x, y \in D$ such that $f(x) = a$ and $f(y) = b$. Choose a path $\gamma$ from $x$ to $y$ in $D$. Then $f \circ \gamma$ is a path from $a$ to $b$ in $f(D)$.

**Definition:** A subset $W$ of $\mathbb{R}^n$ is **bounded** if $W$ lies in some open ball. This is clearly equivalent to: $W$ lies in some $B(0,r)$.

Now boundedness, unlike connectedness, is not preserved by continuous functions. For example, if $W = (0,1) \subset \mathbb{R}$ and $f : W \to \mathbb{R}$ is defined by $f(x) = \frac{1}{x}$, then $W$ is bounded but $f(W)$ is not bounded.

Neither is the property of being closed preserved by continuous functions. For example, $\mathbb{R}$ is closed in $\mathbb{R}$ and $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = e^x$ is continuous, but $f(\mathbb{R}) = \{y \in \mathbb{R} | y > 0\}$ is not closed.

However, when we put closed and bounded together, they are then both preserved.

**Definition:** $C \subset \mathbb{R}^n$ is **compact** if it is closed and bounded.

### Proposition 5:

If $C \subset \mathbb{R}^n$ is compact and

$$
f : C \to \mathbb{R}^n
$$

is continuous, then $f(C)$ is compact.

The proof is relegated to an appendix.
subspace topology, countable bases

Sometimes we will have a subset \( W \) of \( \mathbb{R}^n \) and will want to
which subsets of \( W \) we should call open sets in \( W \).

**Definition:** If \( U \subset W \subset \mathbb{R}^n \), we say \( U \) is an **open set in** \( W \)
exists an open set \( V \) in \( \mathbb{R}^n \) such that
\[ U = V \cap W. \]

For example, if \( W = [0,1] \subset \mathbb{R} \), then
\[ (1,1) = \{ x \in \mathbb{R} \mid \frac{1}{2} < x \leq 1 \} \]
is an open set in \( W \), but not an open set in \( \mathbb{R} \). \( U' = [\frac{1}{2},1] \) is not open in \( W \).

Note that if \( W \) is an open set in \( \mathbb{R}^n \), then \( U \subset W \) is open in \( \mathbb{R}^n \)
and only if it is open in \( \mathbb{R}^n \).

For \( W \subset \mathbb{R}^n \), the collection of all open sets of \( W \) is the
**open topology** of \( W \).

Recall that \( V \subset \mathbb{R}^n \) is defined to be open if any \( x \in V \) has
\( B(x,r) \subset V \). This is equivalent to saying that \( V \subset \mathbb{R}^n \) is open
is either the empty set or is a union of open balls. (Exercise.)

Of course, not every open set is an open ball, but open balls suffice to
all open sets by taking unions (the "empty" union being empty).

**Definition:** A collection \( \mathcal{V} = \{ V_\alpha \} \) of open sets in \( \mathbb{R}^n \) is an
**basis** for open sets if every open set in \( \mathbb{R}^n \) is a union of some of
\( V_\alpha \).

**Examples:** The set of all open squares in \( \mathbb{R}^2 \) is a basis for the
open sets in \( \mathbb{R}^2 \).

The set of all open intervals \( (a,b) \subset \mathbb{R} \) with \( a \) and \( b \)
rational is a basis for the open sets in \( \mathbb{R} \).

The set of all open balls in \( \mathbb{R}^n \) is, of course, a basis for open
sets. But so is
\[ \{ B(x,r) \mid x = (x_1, \ldots, x_n) \text{ with each } x_i \text{ rational, and } r \text{ is rational} \}. \]

(See Proposition 7.)

For a subset \( W \) of \( \mathbb{R}^n \),
and we can give the same definition as above for the notion of a
basis for the open sets in \( W \). Indeed, it is clear that if \( \mathcal{V} = \{ V_\alpha \} \)
is a basis for the open sets in \( \mathbb{R}^n \), then \( \{ V_\alpha \cap W \} \) is a basis for the
open sets of \( W \).

We want to get bases for open sets which are "minimal" in the
sense that they have no more sets than needed to do the job. The
notion that comes up is **countability.**

**Definition:** A set \( S \) is **countable** if its elements can all be
arranged in a finite or infinite sequence \( s_1, s_2, s_3, \ldots \); that is,
every element of \( S \) will be somewhere in the sequence.

**Examples:** The set \( \mathbb{Q} \) of all positive rational numbers is
countable; for example
\[ 1, 2, \frac{1}{2}, 3, \frac{3}{2}, 4, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \ldots \]
is a sequence containing all positive rationals. Similarly,
The set \( I = \{0,1\} = \{x \in \mathbb{R} \mid 0 \leq x \leq 1\} \) is not countable. We prove this contrapositively. Suppose

\[ x_1, x_2, x_3, \ldots \]

is a list of all elements of \( I \). It suffices to give an element of \( \mathbb{R} \) which cannot be in the list. Express the \( x_i \)'s as decimals

\[ x_i = 0.x_{i1}x_{i2}x_{i3} \ldots \]

Then, let

\[ r = 0.y_1y_2y_3\ldots \]

where \( y_j = 5 \) if \( x_{jj} \neq 5 \) and \( y_j = 1 \) if \( x_{jj} = 5 \). Then

\[ r \neq x_1 \text{ because } y_1 \neq x_{11} \]

\[ r \neq x_2 \text{ because } y_1 \neq x_{22}, \text{ etc.} \]

\[ r \in I. \]

**Proposition 6:** If \( A \) and \( B \) are countable sets, then so is their cartesian product \( A \times B \).

**Proof:** Let \( A = \{a_1, a_2, a_3, \ldots\} \) and \( B = \{b_1, b_2, b_3, \ldots\} \). Then we can write

\[ A \times B = \{(a_1, b_1), (a_1, b_2), (a_2, b_1), (a_3, b_1), (a_2, b_2), \ldots\} \]

Choose following the path shown below will include all of \( A \times B \).

\[ b_1 \quad b_2 \quad b_3 \quad b_4 \quad b_5 \ldots \]

\[ a_1 \quad a_2 \quad a_3 \quad a_4 \quad a_5 \quad \vdots \]

**Proposition 7:** \( \mathbb{R}^n \) (and hence any \( W \subset \mathbb{R}^n \)) has a countable basis for its open sets.

**Proof:** The set

\[ C = \{B(x, r) \mid x = (x_1, \ldots, x_n) \text{ each } x_i \in \mathbb{Q} \text{ and } r \in \mathbb{Q}\} \]

can be put in 1-1 correspondence with \((n+1)\)-tuples \((x_1, \ldots, x_n, r)\) of rational numbers (with \( r > 0 \)). By Proposition 6 this is a countable set of balls.

Let \( V \) be any open set in \( \mathbb{R}^n \). To show that \( V \) is a union of elements of \( C \) it suffices to show that for \( y \in V \) some \( B(x, r) \in C \) contains \( y \) and lies in \( V \). (For then \( V \) is the union of such \( B(x, r) \) -- one for each \( y \in V \).) Since \( V \) is open some

\[ B(y, s) \subset V. \]

Choose \( x \) with all coordinates rational such that

\[ d(x, y) < \frac{s}{3} \]

and let \( r \) be a rational number satisfying \( \frac{s}{3} < r < \frac{s}{2} \).
A space \( X \) is an \( n \)-manifold if each \( x \in X \) lies in some open set homeomorphic to some \( B(o,r) \subseteq \mathbb{R}^n \). An \( n \)-manifold is said to have dimension \( n \).

Proposition 8: A matrix group of dimension \( n \) is an \( n \)-manifold.

Proof: The exponential map from the \( n \)-dimensional tangent space \( T \) to \( G \) is continuous. It is one-to-one on some neighborhood \( V \) of \( 0 \) in \( T \) because it has an inverse \((\log)\). Also, this inverse is continuous. Take \( B(o,r) \subseteq V \) and we have that the identity matrix \( I \) has the right kind of neighborhood. For any \( x \in G \) we have

\[ L_x \circ \exp : B(o,r) \to G \]

being a homeomorphism (composition of homeomorphisms is a homeomorphism) onto a neighborhood of \( x \). Thus \( G \) is an \( n \)-manifold.

Definition: A manifold is called closed if it is compact (= closed and bounded).

Proposition 9: \( \text{GL}(n,k) \) is not closed, but \( \mathfrak{gl}(n,k) \) is closed.

Proof: Clearly \( \text{GL}(n,k) \) is not bounded, because for every non-zero real number \( r \), \( rI \in \text{GL}(n,k) \). (This also shows that \( \text{GL}(n,k) \) is not closed. Because \( 0 \in M_n(k) - \text{GL}(n,k) \) but every ball with center \( 0 \) will contain some \( rI \in \text{GL}(n,k) \).)

If \( A \in \mathfrak{gl}(n,k) \) then the rows are unit vectors so that as a vector in \( M_n(k) \) the length of \( A \) is \( \leq n \). Thus \( \mathfrak{gl}(n,k) \) is a bounded set.

To see that \( M_n(k) - \mathfrak{gl}(n,k) \) is open, suppose \( B \in M_n(k) - \mathfrak{gl}(n,k) \).

Then there exists \( x,y \in k^n \) such that...
\[ \langle xB, yB \rangle \neq \langle x, y \rangle. \]

\(\langle , \rangle\) is continuous, there is some open ball \(B(B, s)\) in \(m\) such that for \(B' \subset B(B, s)\) we have \(\langle xB', yB' \rangle \neq \langle x, y \rangle\). Thus \(c(n, k)\).

q.e.d.

We finish this chapter with a result which will be of substantial help to us later on.

**Proposition 10:** Let \(N\) and \(M\) be closed \(n\)-manifolds with \(M\). If \(M\) is connected, then \(N = M\).

**Proof:** We want to show that \(M - N\) is empty. If it isn’t, choose \(x \in M - N\) and \(y \in N\). Since \(M\) is connected, there exists a path \(\varphi : [0, 1] \to M\)

\(\varphi(0) = x\) and \(\varphi(1) = y\). Then \(\varphi^{-1}(M - N)\) is an open set in \([0, 1]\) (see Exercise 3) and it contains 1 but not 0. Let \(t_0\)

the largest element of the closed set \(I - \varphi^{-1}(M - N)\). Then

(i) every \(B(t_0, \varepsilon)\) contains points of \(\varphi^{-1}(M - N)\), but

(ii) since \(N\) is a manifold there is some open neighborhood \(U\) of \(\varphi(t_0)\) in \(N\). By continuity of \(\varphi\) some \(B(t_0, \varepsilon)\) maps by \(\varphi\)

\(\to N\). This contradiction shows \(N = M\).

**Corollary:** The map \(\varphi : \text{Sp}(1) \to SO(3)\) (see Chapter V) is surjective.

**Proof:** Since \(\varphi\) is a homeomorphism on some neighborhood of each point, we see that the image \(\varphi(\text{Sp}(1))\) is a 3-manifold. Since \(\varphi\) is continuous, this image is a closed 3-manifold (Proposition 5). It remains to prove that \(SO(3)\) is connected (so that we may apply Proposition 10). It suffices to show that any \(A \in SO(3)\) may be joined to the identity matrix \(I\) by a path in \(SO(3)\).

We have \(\det A = 1\) and that \(\{A_e_1, A_e_2, A_e_3\}\) is an orthonormal basis for \(\mathbb{R}^3\). Let \(B\) be a rotation sending \(e_1\) to \(A_e_1\) and leaving the direction perpendicular to the \((e_1, A_e_1)\) plane fixed. (If \(A_e_1 = e_1\) proceed directly to the next step. If \(e_1\) and \(A_e_1\) are antipodal on \(S^2\), then we have two choices for \(B\).) Clearly, there is a path \(\varphi\) from \(I\) to \(B\) in \(SO(3)\). Now \(B e_2 \in A_e_2\) so \(B e_2\) and \(B e_3\) are an orthonormal basis for the plane perpendicular to \(A_e_1\). Let \(C\) be a rotation of this plane sending \(B e_2\) to \(A_e_2\) and \(B e_3\) to \(A_e_3\). (If we can’t do this, we would have \(\det A = -1\).) There is a path \(\sigma\) from \(I\) to \(C\) in \(SO(3)\). Since \(A = BC\), we can multiply the paths \(\varphi\) and \(\sigma\) to get a path from \(I\) to \(A\) in \(SO(3)\).

q.e.d.

**Note:** In Chapter VIII we will prove that \(SO(n)\) is connected for all \(n\).

**F. Exercises**

1. Show that the definition of continuity reduces to the usual \(\varepsilon - \delta\) definition for \(f : (a, b) \to \mathbb{R}\).

2. Suppose we have \(A \in \mathbb{R}^n\) and have functions

\[ A \xrightarrow{f} \mathbb{R}^m, f(A) \in \mathbb{R}^r.\]
9. Let $D \subseteq \mathbb{R}^n$ be open and closed. Show that if $D$ is not empty, then $D = \mathbb{R}^n$. (See the proof of Proposition 10.)
Cartesian products of groups

If $G$ and $H$ are groups, we make $G \times H$ into a group by defining

$$(g,h)(g',h') = (gg', hh').$$

This works and if $G, H$ are abelian so is $G \times H$. (Exercises.)

**Example:** If $G$ is a group of $n \times n$ matrices and $H$ is a group of $m \times m$ matrices, we can represent elements of $G \times H$ as $\begin{pmatrix} n & m \\ 0 & m \end{pmatrix}$ x $(n + m)$ matrices by

$$(g,h) = \begin{pmatrix} g & 0 \\ 0 & h \end{pmatrix}.\]$$

Then matrix multiplication gives the operation described above on $G \times H$. (Exercise.) Let us look at an important special case of this.

Let $G = \{\begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}\}$ and $H = \{\begin{pmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{pmatrix}\}$. Then

$G \times H = \begin{pmatrix} \cos \theta & \sin \theta & 0 & 0 \\ -\sin \theta & \cos \theta & 0 & 0 \\ 0 & 0 & \cos \phi & \sin \phi \\ 0 & 0 & -\sin \phi & \cos \phi \end{pmatrix}$.

Both $G$ and $H$ are circle groups, and the (abelian) group $G \times H = S^1 \times S^1$ is called a 2-torus.

**Definition:** A $k$-torus is the Cartesian product of $k$ circle groups.

We have seen that a $k$-torus can be represented by a "block diagonal" $2k \times 2k$ real matrix. But it is easy to see that

$$T = \begin{pmatrix} I_{2k} \\ e \\ e \\ \cdots \end{pmatrix}$$

is a $k$-torus, so we can represent a $k$-torus as diagonal complex $k \times k$ matrices.

**Proposition 1:** If $G$ is an abelian matrix group and $\gamma, \sigma$ are one-parameter subgroups, then $\gamma \sigma$ is a one-parameter subgroup.

**Proof:**

$$(\gamma \sigma)(s+t) = \gamma(s+t)\sigma(s+t)$$

$= \gamma(s)\gamma(t)(\gamma(s)\sigma(t))$$

$= \gamma(s)\gamma(t)\gamma(s)\sigma(t)\gamma(t)$$

$= (\gamma \sigma)(s)(\gamma \sigma)(t).$$

**Corollary:** If $G$ is an abelian matrix group then $\exp : (TG)_e \to G$ is a homomorphism from the vector group of the tangent space $(TG)_e$ to $G$ at $e$.

**Proof:** Let $\xi = \gamma'(0)$, $\eta = \sigma'(0)$ with $\gamma, \sigma$ being one-parameter subgroups. Then from Chapter IV we know that $\exp(\xi) = \gamma(1)$ and
\( \gamma(0) = \gamma'(0) + \sigma'(0) = \xi + \eta \) (Chapter III), so \( \exp(\xi + \eta) \)
\( (1) = \gamma(1) \sigma(1) = \exp \xi \exp \eta \).

Now we know that \( \exp \) is 1-1 on some neighborhood \( V \) of \( 0 \) \( (G) \). So for \( G \) abelian we know that \( \ker(\exp) = L \) is a

Proposition 2: Let \( G \) be a connected matrix group and let \( H \) be

\( \exp : (G) \rightarrow G \) is surjective.

Proof: Since \( U \subset H \) and \( H \) is a subgroup, we must have

\[ U^2 = \{ xy \mid x, y \in U \}, U^3, U^4, \ldots \text{ all in } H. \]

\[ W = U \cup U^2 \cup U^3 \cup \ldots \subset H. \]

Each \( U^k \) is open, so \( W \) is an open set. (Exercise, Chapter VI.)

\( W \) is also closed. For, let \( x \) be a limit point of \( W \). Then

there is an open set containing \( x \) (\( U \subset W \)) and hence must contain some

point of \( W \). Thus

\[ xu = u_1 \ldots u_m \text{ for some } u, u_1, u_2, \ldots, u_m \in U. \]

Then \( x = u_1 \ldots u_m u^{-1} \in W \). In a connected space \( G \) only \( \emptyset \) and

\[ \emptyset \text{ are both open and closed (Exercise, Chapter VI). } W \not= \emptyset \text{ so } W = G \]

\( H = G. \)

Corollary: If \( G \) is a connected abelian matrix group, then

\[ \exp : (G) \rightarrow G \] is a surjective homomorphism with a discrete kernel.

Proof: \( \exp \) is a homomorphism so \( \exp((G) \) is a subgroup of

\( G \) and it contains some neighborhood \( U \) of \( e \). Thus \( \exp((G) \subset G \).

In the exercises it is proved that this implies that \( \exp((G) \)

\( = \emptyset \times \emptyset \) for some \( k \). So we have

Theorem 1: Any compact connected abelian matrix group \( G \) is a
torus.

B. Maximal tori in groups

Definition: A subgroup \( H \) of a matrix group \( G \) is a torus if it

is isomorphic with a \( k \)-torus for some \( k \). It is a maximal torus if it

is not contained in any larger torus subgroup of \( G \).

Proposition 3:

\[ T = \begin{pmatrix}
\cos \theta & \sin \theta & 0 \\
\sin \theta & \cos \theta & 0 \\
0 & 0 & 1
\end{pmatrix} \]

is a maximal torus in \( \text{SO}(3) \).

Proof: Clearly \( T \) is isomorphic with \( \begin{pmatrix}
\cos \theta & \sin \theta \\
\sin \theta & \cos \theta
\end{pmatrix} \) which

is a circle group and thus is a 1-torus.

Suppose there is a larger torus subgroup \( T' \) of \( \text{SO}(3) \), i.e.,

\[ T \not\subset T' \subset \text{SO}(3). \]

Since \( T' \) would be abelian, we would have: \( \theta \in \text{SO}(3) \) such that
but $\sigma$ commutes with every element of $T$. So it suffices to
prove that:

$\sigma \in \text{SO}(3)$ commutes with each $t \in T$, then $\sigma \in T$. Refer
to our proof that $\text{Center} \text{SO}(3) = \{I\}$ in Chapter V, and you will
see that we have already proved this fact.

**Proposition 4:**

$$
T = \begin{pmatrix}
\cos \theta_1 & \sin \theta_1 & 0 & 0 \\
-\sin \theta_1 & \cos \theta_1 & 0 & 0 \\
0 & 0 & \cos \theta_2 & \sin \theta_2 \\
0 & 0 & -\sin \theta_2 & \cos \theta_2 \\
\end{pmatrix}
$$

**Maximal torus in SO(4).**

**Proof:** This clearly is a 2-torus and is a subgroup of SO(4).

Moreover, it suffices to prove that if $\sigma \in \text{SO}(4)$ commutes with all
elements of $T$ then $\sigma \in T$.

Let $V$ be the 2-plane in $\mathbb{R}^4$ spanned by $e_1, e_2$ and $W$ be the
plane in $\mathbb{R}^4$ spanned by $e_3, e_4$. We see that $T$ consists of all

(rotation of $V$, rotation of $W$).

**Claim:**

$\sigma(e_1) \in V$.

Take $\sigma \in T$ such that $\sigma$ is the identity on $V$ but is not the iden-
tity on $W$. Then

$\sigma(e_1) = ae_1 + be_2 + ce_3 + de_4$

$\sigma^2(e_1) = ae_1 + be_2 + c'e_3 + d'e_4$

so $\sigma(e_1) = \sigma(e_1') = ae_1 + be_2 + ce_3 + de_4$.

This shows $c = 0 = d$, so $\sigma(e_1) \in V$. The same kind of proof shows
$\sigma(e_2) \in V$. Dually, $\sigma(e_3)$ and $\sigma(e_4)$ are in $W$.

So we know that $\sigma$ is orthogonal on $V$ and is orthogonal on $W$.
A priori, it could be a reflection in each and we would still have
$\sigma \in \text{SO}(4)$. But $\sigma$ commutes with all rotations on $V$ and is thus a
rotation on $V$. (See the proof of Proposition 3, Chapter V.) Similarly, $\sigma$ is a rotation on $W$, so $\sigma \in T$.

q.e.d.

From Propositions 1 and 2 the general result about maximal tori
in $\text{SO}(n)$ should be clear. We have $n/2$ of the $2 \times 2$ rotation
matrices for $n$ even and have a $1$ in the $n,n$ position for $n$
ond. The proof of the general case is an obvious extension of the
above proofs.

**Proposition 5:**

$$
T = \begin{pmatrix}
i^2 & 0 & 0 \\
e & i^2 & 0 \\
0 & 0 & e \\
\end{pmatrix}
$$

is a maximal torus in $U(n)$.

**Proof:** Let $\sigma \in U(n)$ commute with each $a \in T$. Consider any
$a$ of the form

$$
a = \begin{pmatrix}
1 & 0 & 0 \\
i & i & 0 \\
0 & 0 & e \\
\end{pmatrix} \in T.
$$

Then $a\sigma(a_1) = \sigma(a\sigma(a_1)) = \sigma(a_1)$. So $\sigma$ leaves $\sigma(a_1)$ fixed, but
such $\sigma$'s can move any vector which is not a multiple of $e_1$. Hence

$\sigma(e_1) = \lambda e_1$, $\lambda \in \mathbb{C}$.
For arguments give
\[ \sigma(e_j) = \lambda_j e_j \text{ for } j = 1, \ldots, n. \]

\[ \sigma = \begin{pmatrix} \lambda_1 & \circ & \cdots & \circ \\ \circ & \lambda_2 & \cdots & \circ \\ \vdots & \vdots & \ddots & \vdots \\ \circ & \circ & \cdots & \lambda_n \end{pmatrix} \]

Since \( \sigma \in U(n) \), each \( \lambda_j \) is of unit length. Thus \( \sigma \in T \) is maximal.

**Proposition 6:**
\[ T = \left\{ \begin{pmatrix} \theta_1 & \circ & \cdots & \circ \\ \cdots & \theta_n & \cdots & \circ \\ \circ & \cdots & \cdots & \theta_n \\ \circ & \cdots & \cdots & \circ \end{pmatrix} \mid \theta_1 + \cdots + \theta_n = 0 \right\} \]

is a maximal torus in \( SU(n) \).

**Proof:** A matrix \( \sigma \) of the form
\[ \begin{pmatrix} \theta_1 & \circ & \cdots & \circ \\ \cdots & \theta_n & \cdots & \circ \\ \circ & \cdots & \cdots & \theta_n \\ \circ & \cdots & \cdots & \circ \end{pmatrix} \]

has \( \sigma = e_1^{\theta_1} \cdots e_n^{\theta_n} \), so that \( \sigma \in SU(n) \iff \det \sigma = e^{\theta_1 + \cdots + \theta_n} = 1 \). Thus the \( T \) described here is just the intersection of \( SU(n) \) with the maximal torus given for \( U(n) \).

First we must check that this is an \( (n-1) \)-torus. To do this,

\[ \theta_1 = 0 \]

It is an exercise to show that this works.

Now for \( n > 2 \) the same proof as used for \( U(n) \) will work, but for \( n = 2 \)

\[ T = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \]

and we do not have matrices \( \sigma = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \) to use. But for \( n = 2 \)

we give a direct simple proof. If
\[
\sigma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SU(2) \text{ and } \sigma = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \in T
\]

then
\[ \varphi \sigma = \begin{pmatrix} a & -b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ -c & d \end{pmatrix} = a \sigma. \]

Thus \( b = 0 = c \) and \( \sigma \in T \).

**Proposition 7:** The maximal torus given for \( U(n) \) is also a maximal torus for \( Sp(n) \).

**Proof:** Just as for \( U(n) \) we can show that any element of \( Sp(n) \)
with commuting with all
\[
\begin{pmatrix}
1 & i_a \\
0 & e
\end{pmatrix}
\]
must be diagonal with the
eigenvalues having length 1. But now these elements are
prestics. However (Exercise), any quaternion which commutes with
must be a complex number.
q.e.d.

Centers again

Now that we know maximal tori in our matrix groups, we are able
to calculate the centers.

Proposition B: Center (Sp(n)) = \{I, -I\}.

Proof: We have seen that if any element commutes with all ele-
ments of the maximal torus we have described, then it must lie in
a maximal torus. Hence, in every case

\[
\text{Center} = T.
\]

If \(A \in \text{Center} \text{Sp}(n)\), then
\[
A = \begin{pmatrix}
1 & i_a \\
0 & e
\end{pmatrix}
\]

Since \(A\) must
commute with the matrix \(jjI\), it follows that the diagonal elements
must be real (and since they are of unit length) they are \(+1\). It
is an exercise to show that they all have the same sign. Now \(I\) and
\(-I\) are in the center.
q.e.d.

Proposition 9: Center \(U(n) = \{e^{i\theta}I\} \cong S^1\)

Center \(SU(n) = \{wI \mid w^n = 1\}\).

Proof: If \(B \in \text{Center} \text{U}(n)\) we get that \(B\) is diagonal with
diagonal elements complex numbers of unit length. Let
\[
B = \begin{pmatrix}
a_1 & \ldots & 0 \\
0 & \ddots & 0 \\
0 & \cdots & a_n
\end{pmatrix}
\]
with each \(|a_i| = 1\). Let
\[
A = \begin{pmatrix}
0 & 1 & 0 \ldots & 0 \\
1 & 0 & 0 \ldots & 0 \\
0 & 0 & 1 \ldots & 0 \\
\vdots & \ddots & \ddots & \ddots
\end{pmatrix}
\]

Then \(AB = BA\) shows \(a_1 = a_2\), etc., so all \(a\) are equal. Clearly
any \(e^{i\theta}I\) is in the center, so the center of \(U(n)\) is as asserted.

For \(SU(n)\) we note that the same argument will show that an
element must be of the form \(e^{i\theta}I\) to be in the center. But
\[
\det(e^{i\theta}I) = e^{i\theta},
\]
and since this must be 1, \(e^{i\theta}\) must be an \(n\)th root of unity.
q.e.d.

So Center \(SU(n) = SU(n) \cap \text{Center} \text{U}(n)\).

Finally, we want to calculate the center of \(SO(n)\). It turns
out that it depends on whether \(n\) is even or odd. The groups \(SO(2n)\)
and the groups \(SO(2n+1)\) are different in some important ways.

Now \(SO(2) = S^1\) is abelian so its center is the group itself. We
have already proved that the center of \(SO(3)\) is just \([I]\).
For \( k \geq 3 \) any element in the center of \( SO(k) \) must be a matrix.

Moreover, if \( A \in SO(k) \) is in the center, it must be in our maximal torus. So suppose \( A \in \text{Center } SO(k) \) is of the form

\[
A = \begin{pmatrix}
\cos \theta_1 & \sin \theta_1 & 0 & \cdots \\
-\sin \theta_1 & \cos \theta_1 & 0 & \cdots \\
0 & 0 & 1 & \cdots \\
& & & & \\
\end{pmatrix}
\]

\[
P = \begin{pmatrix}
0 & 0 & 1 & \cdots \\
0 & 1 & 0 & \cdots \\
-1 & 0 & 0 & \cdots \\
& & 1 & 0 & \\
& & & 1 & \\
\end{pmatrix} \in SO(k)
\]

has zero in the 1,2 position, whereas \( AP \) has \( \sin \theta_1 \) in the 1,2 position. Thus \( \sin \theta_1 = 0 \). Similar arguments show all diagonal terms are zero.

This allows also (since each \( \sin \theta_i = 0 \)) that each diagonal term is 1 or -1. So each 2 \times 2 block is \( \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \) or \( \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \). Arguments like we used for \( U(n) \) show that all diagonal terms must be equal. So we finally conclude that

Center \( SO(2n+1) = \{I\} \)

Center \( SO(2n) = \{I, -I\} \).

Since \( \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \) is in the center of \( SO(4) \), but \( \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \)

isn't in \( SO(3) \).

We now tabulate the information we have generated about our groups.

<table>
<thead>
<tr>
<th>Group</th>
<th>Dimension</th>
<th>Center</th>
<th>Standard Maximal torus</th>
</tr>
</thead>
<tbody>
<tr>
<td>( U(n) )</td>
<td>( n^2 )</td>
<td>( {e^{i\theta_1}} \equiv S^1 )</td>
<td>( \begin{pmatrix} e^{i\theta_1} \ \cdots \ e^{i\theta_n} \end{pmatrix} )</td>
</tr>
<tr>
<td>( SU(n) )</td>
<td>( n^2 - 1 )</td>
<td>( {e^{i\theta_1}} \equiv S^1 )</td>
<td>( \begin{pmatrix} e^{i\theta_1} \ \cdots \ e^{i\theta_n} \end{pmatrix} )</td>
</tr>
<tr>
<td>( SO(2n+1) )</td>
<td>( \frac{(2n+1)(2n)}{2} = 2n^2 + n )</td>
<td>( I )</td>
<td>( \begin{pmatrix} \text{rot} \theta_1 \ \cdots \ \text{rot} \theta_n \end{pmatrix} )</td>
</tr>
<tr>
<td>( SO(2n) )</td>
<td>( \frac{2n(2n-1)}{2} = 2n^2 - n )</td>
<td>( I, -I )</td>
<td>( \begin{pmatrix} \text{rot} \theta_1 \ \cdots \ \text{rot} \theta_n \end{pmatrix} )</td>
</tr>
<tr>
<td>( Sp(n) )</td>
<td>( 2n^2 + n )</td>
<td>( I, -I )</td>
<td>( \begin{pmatrix} \text{rot} \theta_1 \ \cdots \ \text{rot} \theta_n \end{pmatrix} )</td>
</tr>
</tbody>
</table>

Note that we have nothing to distinguish \( SO(2n+1) \) and \( Sp(n) \).

A good part of the remainder of this book is devoted to deciding for which \( n \) these are isomorphic.
Exercises

1. Show that the operation defined on \( G \times H \) does make it into a group. Prove that if \( G \) and \( H \) are abelian, so is \( G \times H \).

2. Do the exercise in the first example of §A.

3. Show directly that the product of the matrices

\[
\begin{pmatrix}
\cos \alpha & \sin \alpha \\
-\sin \alpha & \cos \alpha
\end{pmatrix}
\begin{pmatrix}
\cos \beta & \sin \beta \\
-\sin \beta & \cos \beta
\end{pmatrix}
\]

is the matrix for a rotation through angle \( \alpha + \beta \).

4. Let \( T \) be a maximal torus in a matrix group \( G \) and let \( G \) be a maximal torus in \( G \).

5. Prove that if \( q \) is a quaternion such that \( qi = 1 \), then \( q \) is a complex number.

6. Show that \( \frac{U(n)}{SU(n)} \) is isomorphic to \( \frac{SU(n)}{SU(n)} \).

7. A lattice subgroup \( K \) of \( \mathbb{R}^n \) consists of all integer linear combinations of some set of linearly independent vectors. More explicitly, let

\[ v_1, v_2, \ldots, v_k \]

linearly independent vectors in \( \mathbb{R}^n \). For any integers \( a_1, \ldots, a_k \), the form

\[ a_1 v_1 + \ldots + a_k v_k \]

is in \( K = \{ a_1 v_1 + \ldots + a_k v_k \mid a_i \in \mathbb{Z} \} \). It is routine to verify that \( K \) is a subgroup of \( \mathbb{R}^n \).

8. Show that if \( L \) is a lattice group in \( \mathbb{R}^n \) generated by \( v_1, \ldots, v_k \), then \( \mathbb{R}^n/L \) is isomorphic with the product of \( k \)-torus and \( \mathbb{R}^{n-k} \).

A subgroup \( H \) of \( \mathbb{R}^n \) is discrete if some neighborhood of \( 0 \) in \( \mathbb{R}^n \) contains no point of \( H \) other than \( 0 \). Prove that: A discrete subgroup of \( \mathbb{R}^n \) is a lattice subgroup. (Choose a nonzero vector \( v_1 \in H \) such that no element of \( H \) lies in the interval \( [0,v_1] \) in \( \mathbb{R}^n \). Show that \( \mathbb{Z} v_1 \) contains all integral multiples of \( v_1 \) but no other elements of \( H \). Choose \( v_2 \) in \( H \) with \( v_2 \) not an integral multiple of \( v_1 \). Show that the span of \( v_1 \) and \( v_2 \) in \( \mathbb{R}^n \) contains all integral linear combinations of \( v_1 \) and \( v_2 \) but no other elements of \( H \). etc.)