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

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

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 $N = \{1,2,3,\dots\}$ 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 .

<u>Definition</u>: A <u>group</u> G is a set G along with a binary

ration

$$\phi : G \times G \to G$$

isfying certain properties. To state these properties it is con-

ient to adopt a simple notation--for $\phi(a,b)$ we just write ab.

Required properties of the operation:

(i) The operation is <u>associative</u>. This means that for any

$,c \in G$ we have

$$(ab)c = a(bc) .$$

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

$(a,b),c) = \phi(a,\phi(b,c))$.

(ii) There exists an <u>identity</u> element e of G. This means

t for any $a \in G$ we have $ea = ae = a$.

(iii) <u>Inverses</u> 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

t there may be more than one identity element and that an element

have more than one inverse. But neither of these can happen.

<u>Proposition</u> 1: <u>A group G has exactly one identity element and</u>

<u>h</u> $a \in G$ <u>has exactly one inverse</u>.

<u>Proof</u>: Suppose e and f are identity elements of G. Then

$fe = e$ since f is an identity element, and

$fe = f$ since e is an identity element .

Suppose both b and c are inverses of a . Then

$$b = eb = (ca)b = c(ab) = ce = c .$$

<u>Examples</u>

(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) $\mathbb{R}^+ = \{x \in \mathbb{R} \mid x > 0\}$ 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

$$\mathbf{0} = (0,0,\ldots,0)$$

l 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
nctions) of S onto S . For example $f : S \to S$ given by
) = b , $f(b) = c$, $f(c) = a$ is one element of G . We define an
ration on G as follows: if $f,g \in G$ we let

$$f \circ g : S \to S$$

defined by $(f\circ g)(a) = f(g(a))$, $(f\circ g(b) = f(g(b))$, $(f\circ g)(c)$
(g(c)) , i.e., $f\circ g$ means first apply g to S and then apply f.
$i : S \to S$ be the identity element $(i(a) = a,\ i(b) = b,\ i(c) = c)$.
n 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
ration. Thus G is a group. It is called the symmetric group on
b,c} (or just the symmetric group on three elements).

Definition: A group G is abelian if for every $a,b \in G$ we
e $ab = ba$.

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

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

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 σ , or we can map a and b into H by σ 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 $\sigma(e) = \sigma(ee) = \sigma(e)\sigma(e)$ and $\sigma(e)$ has an inverse, call it h , 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 $\sigma : \mathbb{R} \to \mathbb{R}^2$ ($\mathbb{R}$ = additive group of reals, $\mathbb{R}^2$ as in example (6)) by $\sigma(x) = (x,x)$, then σ is a homomorphism but is not surjective because $\sigma(\mathbb{R})$ is just the diagonal line in $\mathbb{R}^2$. But $\rho : \mathbb{R}^2 \to \mathbb{R}$ defined by $\rho(x,y) = x$ 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 $\sigma : \mathbb{R} \to \mathbb{R}$ $(\sigma(x) = (x,x))$ is injective, and the map

$^2 \to \mathbb{R}$ $(\rho(x,y) = x)$ is not injective.

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

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

$$\sigma : \mathbb{R} \to \mathbb{R}^+$$

$$\sigma(x) = a^x .$$

...n σ is a homomorphism

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

...o, σ is injective. For, suppose $\sigma(x) = \sigma(y)$. This means ...$= a^y$ and so $a^{-y}a^x = a^{-y}a^y = 1$ and $a^{x-y} = 1$ which implies ...$y = 0$ or $x = y$. Also, σ is surjective. For, if y is any ...itive real number $x = \log_a y$ has the property that $a^x = y$. ...s these two groups are isomorphic--not only that, but there are lots ...isomorphisms.

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

$$\sigma \text{ is injective} \Leftrightarrow \sigma^{-1}(e') = e .$$

$$\sigma(a) = \sigma(b) \Leftrightarrow \sigma(a)\sigma(b)^{-1} = e' \Leftrightarrow \sigma(ab^{-1}) = e'$$

and

$$ab^{-1} = e \Leftrightarrow a = b .$$

B. Fields, Quaternions

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.

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

$$(x_1,x_2)(y_1,y_2) = (x_1y_1, x_2y_2)$$

(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
field because:

<u>Proposition</u> 3: <u>In a field</u> k <u>if</u> $a \neq 0$ <u>and</u> $b \neq 0$, <u>then</u>

$$ab \neq 0 .$$

<u>Proof</u>: If $a \neq 0$ then $a \in k - (0)$ which by (iii) is required
be a group under multiplication. Thus there is an a^{-1} in $k - (0)$
h that $a^{-1}a = 1$ (the multiplicative identity). Thus if $ab = 0$
have

$$a^{-1}(ab) = (a^{-1})(0) = 0$$

$a^{-1}(ab) = (a^{-1}a)b = 1b = 0$ so $b = 0$.

The statement of Proposition 3 is equivalent to the statement
t a field has "<u>no divisors of zero</u>."

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.

$$(a,b)((c,d) + (e,f)) = (a,b)((c + e, d + f))$$

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

s should equal $(a,b)(c,d) + (a,b)(e,f)$. This latter equals
$- bd, ad + bf) + (ad - bf, af + be)$ and we easily check that these
equal. Next we need to see that if $(a,b) \neq (0,0)$ then it has a

multiplicative inverse. Well, $(a,b) \neq (0,0) \Rightarrow a^2 + b^2 \neq 0$; in which case, we need to find a multiplicative inverse for (a,b) . The multiplicative identity clearly is $(1,0)$ and

$$(a,b)\left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2}\right) = (1,0) ,$$

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

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

$$(a + ib)(c + id) = ac + aid + ibc + ibid$$

$$= ac + iad + ibc + i^2bd$$

$$= (ac - bd) + i(ad + bc) .$$

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

$$x \in \mathbb{R} \text{ be } x + i0 .$$

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

$$x + y = x + i0 + y + i0 = (x + y) + i0$$

$$xy = (x + i0)(Y + i0) = (xy) + i0 .$$

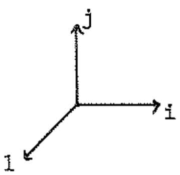
So we have taken the field $\mathbb{R}$ as all $(x,0)$ in $\mathbb{R}^2$ and <u>extended</u>

e operations in $\mathbb{R}$ to $\mathbb{R}^2$ to get a field.

This strongly suggests that we try to extend the field on $\mathbb{R}^2$

a field on $\mathbb{R}^3$. Now for the bad news.

Proposition 4: The operations on $\mathbb{C}$ cannot be extended to make into a field.



Proof: Take basis vectors 1,i,j so

at any element of $\mathbb{R}^3$ can be written

iquely as $a+ib+jc$ with $a,b,c \in \mathbb{R}$.

if we are to have a multiplication extending that of $\mathbb{C}$ we must

ve $ij = a+ib+jc$ for some three real numbers a,b,c. But then

$j) = ia+i^2b+ijc$; so

$$-j = ia - b + ijc$$

$$-j = ia - b + (a + ib + jc)c$$

$$-j = (a - b) + i(a + b) + jc^2 .$$

s implies $c^2 = -1$, contradicting $c \in \mathbb{R}$.

The main thrust of this proof is that if we insist that the

oduct ij be in $\mathbb{R}^3$ we get into trouble. Maybe if we had one more

nension it would work. This is almost true; we can define a multi-

ation on $\mathbb{R}^4$ which satisfies conditions (i) and (ii) for a

eld but (iii) must be replaced by (iii)′ $k - (0)$ is a group under

tiplication--it is not an abelian group. We will just describe how

s can be done. You may be interested in reading "Hamilton's dis-

very of the quaternions" by B. L. van der Waerden in the Mathematics

gazine (vol. 49, #5, (1976)). We take a basis 1,i,j,k for $\mathbb{R}^4$

and define

	1	i	j	k
1	1	i	j	k
i	i	-1	k	-j
j	j	-k	-1	i
k	k	j	-i	-1

Thus 1 acts as identity, $ij = k$, $ji = -k$, etc.

This tells us how to multiply quadruples of real numbers:

$$(a + ib + jc + kd)(x + iy + jz + kw) = (ax - by - cz - dw)$$

$$+ i(ay + bx + cw - dz) + j(az + cx + dy - bw)$$

$$+ k(aw + bx + bz - cy) .$$

$\mathbb{R}^4$ with this multiplication is called the quaternions. It is easy to verify that this does extend the multiplication in $\mathbb{C}$ 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

$$q = a + ib + jc + bd$$

is not the zero $(0 + i0 + j0 + k0)$ 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 ments of k. Define addition on k^n by

$$x = (x_1, \ldots, x_n) \qquad y = (y_1, \ldots, y_n)$$

$$x + y = (x_1 + y_1, \ldots, x_n + y_n) .$$

s makes k^n into an abelian group with identity $\mathbb{0} = (0, \ldots, 0)$. $c \in k$ we define

$$cx = (cx_1, \ldots, cx_n)$$

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

Definition: A map $k^n \xrightarrow{\phi} k^n$ is linear if it respects linear binations; i.e., if $c, d \in k$ and $x, y \in k^n$ then

$$(*) \qquad \phi(cx + dy) = c\phi(x) + d\phi(y) .$$

particular, $\phi(x + y) = \phi(x) + \phi(y)$, so that a linear map is a omorphism of the additive group of k^n. Also

$$\phi(cx) = c\phi(x) ,$$

these two conditions together are equivalent to $(*)$.

Proposition 5: If $k^n \xrightarrow{\phi} k^n \xrightarrow{\psi} k^n$ are both linear, then

so is $\psi \circ \phi$.

Proof: $(\psi \circ \phi)(cx + dy) = \psi(c\phi(x) + d\phi(y))$

$$= c(\psi \circ \phi)(x) + d(\psi \circ \phi)(y) .$$

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 $\phi(M)$ by

$$\phi(M)(x_1, \ldots, x_n) = (x_1, \ldots, x_n)(m_{ij})$$

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

$$\phi(M)(cx + dy) = (cx + dy)\ (m_{ij})$$

$$= c(x_1, \ldots, x_n)(m_{ij}) + d(y_1, \ldots, y_n)(m_{ij}) .$$

We use row vectors instead of column vectors because we 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 $(x_1 c, \ldots, x_n c)$ in general. If we use column vectors and multiply by matrices on the left we do not always get linear maps. For $q, c, d \in \mathbb{H}$ and $x, y \in \mathbb{H}^n$ consider

$$\begin{pmatrix} q & & 0 \\ & \ddots & \\ 0 & & q \end{pmatrix} \begin{pmatrix} cx_1 + dy_1 \\ \vdots \\ cx_n + dy_n \end{pmatrix} = \begin{pmatrix} qcx_1 + qdy_1 \\ \vdots \\ qcx_n + qdy_n \end{pmatrix}$$

d we certainly can't expect this to equal

$$c \begin{pmatrix} qx_1 \\ \vdots \\ qx_n \end{pmatrix} + d \begin{pmatrix} qy_1 \\ \vdots \\ qy_n \end{pmatrix} .$$

ake $n = 1$, $x = 1$, $y = 1$, $d = 0$, $c = i$ and $q = j$.)

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

Note that if the matrix A gives the linear map ϕ and the rix B gives the linear map ψ then AB gives $\psi \circ \phi$. A linear ϕ is an <u>isomorphism</u> if it is injective and surjective (same initions as for group homomorphisms). Then ϕ^{-1} is also a linear morphism and $\phi \circ \phi^{-1}$ = identity map = $\phi^{-1} \circ \phi$. For the corresponding rices this means that $M(\phi^{-1})M(\phi) = I = M(\phi)M(\phi^{-1})$ so that $M(\phi^{-1})$ a 2-sided inverse for $M(\phi)$. So <u>if</u> A^{-1} <u>is a left inverse for then it is also a right inverse for</u> A.

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

(i) If $A = (a_{ij})$ and $B = (b_{ij})$, then

$$A + B = (a_{ij} + b_{ij}) ;$$

(ii) If $A = (a_{ij})$ and $c \in k$, then

$$cA = (ca_{ij}) .$$

This is really no different from the way we made k^n into a vector space, but we are now working with n^2-tuples. However, there is nothing to be gained by writing n^2 elements in a line instead of an $n \times n$ array.

But $M_n(k)$ is not just a vector space. It also has a multiplication which distributes over addition (on either side).

$$A(B + C) = AB + AC$$

$$(B + C)A = BA + CA .$$

Such a system is called an <u>algebra</u>. When we use the word algebra we will always mean one with a two-sided multiplicative identity. For $M_n(k)$,

$$I = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

is the multiplicative identity.

D. <u>General Linear Groups</u>

<u>Definition</u>: If G is an algebra, $x \in G$ is a <u>unit</u> if there exists some $y \in G$ such that $xy = 1 = yx$, i.e., if it has a multiplicative inverse.

<u>Proposition 6</u>: <u>If G is an algebra with an associative</u>

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 $M_n(\mathbb{R})$ is denoted by $GL(n,\mathbb{R})$, in $M_n(\mathbb{C})$ by $GL(n,\mathbb{C})$ and in $M_n(\mathbb{H})$ by $GL(n,\mathbb{H})$. These are the general linear groups.

Note that: $A \in M_n(k)$ is a unit $\Leftrightarrow$ 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 ($H \subset G$ 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 1×1 matrix over k is just an element of k and matrix multiplication of two is just multiplication in k. So we see that

$$GL(1,\mathbb{R}) = \mathbb{R} - (0)$$
$$GL(1,\mathbb{C}) = \mathbb{C} - (0)$$
$$GL(1,\mathbb{H}) = \mathbb{H} - (0)$$

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

$$GL(2,\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \;\middle|\; \begin{matrix} a,b,c,d \in \mathbb{R} \\ ad - bc \neq 0 \end{matrix} \right\},$$

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

For $\mathbb{R}$ and $\mathbb{C}$ we have determinants defined on $M_n(\mathbb{R})$ and $M_n(\mathbb{C})$ and from linear algebra we know that

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

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

Suppose we define a "determinant" on $M_2(\mathbb{H})$ by

$$\det\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \alpha\delta - \beta\gamma.$$

Then $\det\begin{pmatrix} i & j \\ i & j \end{pmatrix} = k - (-k) = 2k \neq 0$, but this matrix cannot be a unit or the corresponding linear map would be an isomorphism, whereas

$$(j,-j)\begin{pmatrix} i & j \\ i & j \end{pmatrix} = (0,0)$$

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

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

Proof: $\phi(\text{id}) = \text{id}$ 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

$$xy = \phi(a)\phi(b) = \phi(ab) \in \phi(G) .$$

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

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

$$\Psi : GL(n,\mathbb{H}) \to GL(2n,\mathbb{C}) ,$$

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

We begin with

$$\psi : \mathbb{H} \to M_2(\mathbb{C})$$

defined by

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

Lemma 9: (i) $\psi(\alpha+\beta) = \psi(\alpha) + \psi(\beta)$

(ii) $\psi(\alpha\beta) = \psi(\alpha)\psi(\beta)$

(iii) ψ is injective .

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

Next, for $A \in Mn(\mathbb{H})$ we set

$$\Psi(A) = (\psi(a_{ij}))$$

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

Lemma 10: $\Psi(AB) = \Psi(A)\Psi(B)$.

Proof: Let $A = (\alpha_{uv})$, $B = (\beta_{uv})$. Then

$$(AB)_{ij} = \alpha_{i1}\beta_{1j}+\cdots+\alpha_{in}\beta_{nj} . \quad \text{By Lemma 9}$$

$$(\Psi(AB)_{ij}) = \psi(\alpha_{i1})\psi(\beta_{1j})+\cdots+\psi(\alpha_{in})\psi(\beta_{nj})$$

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

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

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

$$\Psi(AA^{-1}) = I$$

and Ψ is injective so $AA^{-1} = I$. Thus A is nonsingular.

8. Exercises

1. Let $\phi : G \to H$ be a homomorphism of groups. The kernel of ϕ is defined to be

$$\ker \phi = \{x \in G \mid \phi(x) = \text{identity of } H\} .$$

Show that $\ker \phi$ is a subgroup of G .

2. A subgroup W of a group G is <u>normal</u> if for each $x \in G$ we have

$$xWx^{-1} = W .$$

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

3. The <u>center</u> 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 <u>centralizer</u> $C(S)$ of S by

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

Show that $C(S)$ 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 <u>normalizer</u> of S . Show that $C(S) \subset N(S)$ and that $N(S)$ is a subgroup of G . Show that if S is a subgroup of G , then $S \subset N(S)$ 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 <u>subgroup generated by</u> 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 $G = GL(n,2)$

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} .$$

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 $A^r B^s$.)

9. We say that a subgroup K of G <u>normalizes</u> 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

$$\phi : C \to M_2(R)$$

as follows: represent $\alpha \in C$ as $\alpha = \rho e^{i\theta}$ with $\rho \geq 0$ and set

$$\phi(\alpha) = \sqrt{\rho} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} .$$

Show that $\phi(\alpha\beta) = \phi(\alpha)\phi(\beta)$ but that $\phi(\alpha+\beta)$ need not equal $\phi(\alpha)+\phi(\beta)$.

11. Let G be the multiplicative group of complex numbers of unit length. We say that $\alpha \in G$ is a <u>primitive n^{th} root of unity</u> if $\alpha^n = 1$, but none of $\alpha, \alpha^2, \ldots, \alpha^{n-1}$ 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 R^3. Take the two "purely imaginary" quaternions

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

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

how that if α' and β' are multiplied as quaternions, then

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

s just the usual cross product of vectors in R^3.

Chapter 2
Orthogonal Groups

A. <u>Inner products</u>

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

$$\text{for } x \in \mathbb{R}, \quad \bar{x} = x \ .$$

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

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

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

$$(\overline{\alpha + \beta}) = \bar{\alpha} + \bar{\beta} \ .$$

It is an exercise to prove that

$$\overline{\alpha\beta} = \overline{\beta\alpha} \ .$$

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

$$\overline{\alpha\beta} = \bar{\alpha}\bar{\beta} \ .$$

Let $k \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$ and define an <u>inner product</u> $\langle \ , \ \rangle$ on k^n by

$$\langle x,y\rangle = x_1\bar{y}_1 + \ldots + x_n\bar{y}_n \quad .$$

Proposition 1: $\langle\ ,\ \rangle$ has the following properties:

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

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

(iii) $a\langle x,y\rangle = \langle ax,y\rangle$, $\langle x,ay\rangle = \langle x,y\rangle\bar{a}$

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

(v) $\langle x,x\rangle$ is always a real number ≥ 0 and $\langle x,x\rangle = 0 \Leftrightarrow x = (0,\ldots,0)$.

(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 $\overline{a_{ij}}$; its transpose ${}^tA$ is obtained by replacing each a_{ij} by a_{ji} . These two operations commute so that the symbol

$${}^t\bar{A}$$

(the conjugate transpose of A) is unambiguous .

Recall that for $\mathbb{H}^n$ we must operate on the right (since we defined (scalar)(vector) on the left). So we do the same for $\mathbb{R}^n$ 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 $A = (a_{ij})$.

$$xA = (x_1a_{11} + \ldots + x_na_{n1},\ldots,x_1a_{1n} + \ldots + x_na_{nn})$$

$$y\,{}^t\bar{A} = (y_1\bar{a}_{11} + \ldots + y_n\bar{a}_{1n},\ldots,y_1\bar{a}_{n1} + \ldots + y_n\bar{a}_{nn})$$

Thus the left hand side $\langle xA,y\rangle$ equals

$$(x_1a_{11} + \ldots + x_na_{n1})\bar{y}_1 + \ldots + (x_1a_{1n} + \ldots + x_na_{nn})\bar{y}_n ,$$

and the right hand side $\langle x,y\,{}^t\bar{A}\rangle$ equals

$$x_1(a_{11}\bar{y}_1 + \ldots + a_{1n}\bar{y}_n) + \ldots + x_n(a_{n1}\bar{y}_1 + \ldots + a_{nn}\bar{y}_n) .$$

It is easy to see that these contain exactly the same terms.

Orthogonal groups

Again let $k \in \{\mathbb{R},\mathbb{C},\mathbb{H}\}$.

Definition:

$\mathcal{O}(n,k) = \{A \in M_n(k) \mid \langle xA, yA\rangle = \langle x,y\rangle \text{ for all } x,y \in k^n\}$.

Proposition 3: $\mathcal{O}(n,k)$ is a group .

Proof: If $A,B \in \mathcal{O}(n,k)$, then

$$\langle xAB, yAB\rangle = \langle xA, yA\rangle = \langle x,y\rangle$$

so that

$$AB \in \mathcal{O}(n,k) .$$

Clearly the identity matrix I is in $\mathcal{O}(n,k)$.

If $A \in \mathcal{O}(n,k)$ we have

$$\langle e_iA, e_jA\rangle = \langle e_i, e_j\rangle = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} .$$

Now e_iA is just the i^{th} row of A and we see that $\langle e_iA, e_jA\rangle$ is just the ij entry in the product

$$A\,{}^t\bar{A} .$$

Thus $A\,{}^t\bar{A} = I$. But then ${}^t\bar{A}A$ is also the identity since ${}^t(\overline{{}^t\bar{A}A}) = {}^t(\bar{A}\,{}^tA) = A\,{}^t\bar{A} = I$. Thus ${}^t\bar{A} = A^{-1}$, 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 $A^{-1} \in \mathcal{O}(n,k)$. q.e.d.

Definition: For $k = \mathbb{R}$ we write $\mathcal{O}(n,k)$ as $\mathcal{O}(n)$ 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 $A \in M_n(k)$. Then the following conditons are equivalent:

(i) $A \in \mathcal{O}(n,k)$

(ii) $\langle e_iA, e_jA\rangle = \delta_{ij}$

(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) ${}^t\bar{A} = A^{-1}$.

Proof: Exercise.

Proposition 5: Let $A \in M_n(\mathbb{R})$. Then $A \in \mathcal{O}(n) \Leftrightarrow A$ preserves lengths.

Proof: A preserves lengths $\Leftrightarrow \langle xA, xA\rangle = \langle x,x\rangle$ for all $x \in \mathbb{R}^n$. So $\Rightarrow$ is trivial. Conversely, we have

$$\langle (x+y)A, (x+y)A\rangle = \langle x+y, x+y\rangle = \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle$$

$$= \langle xA,xA\rangle + \langle xA,yA\rangle + \langle yA,xA\rangle + \langle yA,yA\rangle .$$

This gives $\langle x,y\rangle + \langle y,x\rangle = \langle xA,yA\rangle + \langle yA,xA\rangle$ and since $\langle\ ,\ \rangle$ over $\mathbb{R}$ is symmetric, this proves

$$\langle xA, yA\rangle = \langle x, y\rangle\ , \text{ i.e., } \ A \in \mathbb{O}(n)\ .$$

Proposition 5 bis: Proposition 5 also holds for $\mathbb{C}$ and $\mathbb{H}$.

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

$$\langle e_iA, e_jA\rangle + \langle e_jA, e_iA\rangle = 0\ .$$

Then consider $x = x_ie_i + x_je_j$ and calculate $\langle xA, xA\rangle$. We get $x_i\overline{x_j}\langle e_iA, e_jA\rangle + x_j\overline{x_i}\langle e_jA, e_iA\rangle = 0$ and thus

$$\langle e_iA, e_jA\rangle(x_i\bar{x}_j - \bar{x}_jx_i) = 0$$

and this forces $\langle e_iA, e_jA\rangle = 0$. q.e.d.

Let us look at $\mathbb{O}(n)$, $U(n)$ and $Sp(n)$ for small n . $\mathbb{O}(1)$ is the set of all real numbers of length one, so $\mathbb{O}(1) = \{1,-1\}$. $U(1)$ is just the set of all complex numbers of length one. This is the circle group S^1 . $Sp(1)$ 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

$$\mathbb{O}(1) = S^0\ ,\quad U(1) = S^1\ ,\ Sp(1) = S^3\ .$$

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 $A \in \mathbb{O}(n,k)$, then

$$(\det A)\overline{(\det A)} = 1\ .$$

Proof: $A\,{}^t\bar{A} = I \Rightarrow (\det A)(\det {}^t\bar{A}) = 1$, and clearly

$$\det {}^t\bar{A} = \det \bar{A} = \overline{\det A}\ . \qquad \text{q.e.d.}$$

Thus if $A \in \mathbb{O}(n)$ $(= \mathbb{O}(n,\mathbb{R}))$, then $\det A \in \{1,-1\}$. We define

$$SO(n) = \{A \in \mathbb{O}(n) \mid \det A = 1\}$$

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

$$SU(n) = \{A \in U(n) \mid \det A = 1\}$$

and call this the special unitary group.

An example of an element of $\mathbb{O}(2) - SO(2)$ 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 just 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 defined quite differently were isomorphic. We have now defined several series of groups ($GL(n,k)$ for $n = 1,2,\ldots$ and $k \in \{\mathbb{R},\mathbb{C},\mathbb{H}\}$ and $\mathbb{O}(n,k)$ for $n = 1,2,\ldots$ and $k \in \{\mathbb{R},\mathbb{C},\mathbb{H}\}$) and our major goal is to find out which of these are isomorphic. The basic idea will be to develop invariants of matrix groups (dimension, rank, etc.), i.e., two groups which are isomorphic must have the same invariants. This will make it possible to say that certain groups are not isomorphic. But if 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 $Sp(1)$ and $SU(2)$ are isomorphic. How would you try to find an isomorphism? $Sp(1)$ is the set of all quaternions of unit length and $SU(2)$ is the set of all complex 2×2 matrices A such that $A{}^t\bar{A} = I$ and $\det A = 1$. The operation in $Sp(1)$ is multiplication of quaternions, in $SU(2)$ it is matrix multiplication.

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

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

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

(i) $A \in Sp(1) \Rightarrow \phi(A) \in SU(2)$ and

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

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

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

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

Let $B = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in SU(2)$. Using $\det B = 1$ and the fact that the rows are orthogonal unit vectors, we find that

$$\delta = \bar{\alpha} \quad \text{and} \quad \gamma = \bar{\beta} .$$

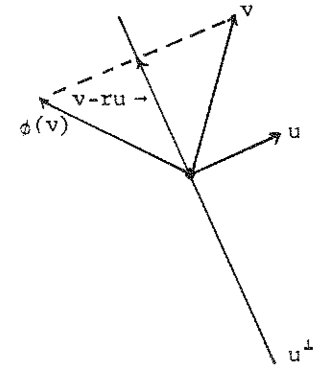
$-\bar{\beta}$

So, if $\alpha = a + ib$ and $\beta = c - id$, we may take $A = a + ib + jc + kd$ and have $\phi(A) = B$ (and $a^2 + b^2 + c^2 + d^2 = 1$).

D. <u>Reflections</u> in $\mathbb{R}^n$

Let u be a unit vector in $\mathbb{R}^n$ and let

$$u^\perp = \{x \in \mathbb{R}^n \mid \langle x,u\rangle = 0\}$$



be its orthogonal complement.

The <u>projection</u> of a vector v into $u^\perp$ is to be $v - ru$, where $r \in \mathbb{R}$ is to be chosen so that $v - ru$ is in $u^\perp$. So $0 = \langle v-ru,u\rangle = \langle v,u\rangle - r\langle u,u\rangle$ and so

$$r = \langle v,u\rangle .$$

Clearly then the <u>reflection</u> of v in $u^\perp$ is to be

$$\phi(v) = v - 2ru = v - 2\langle v,u\rangle u .$$

Choose an orthonormal basis $u_1,\ldots,u_n$ with $u_1 = u$. Then, in this basis, the reflection ϕ is given by the matrix

. Let A be the linear map of $\mathbb{R}^n$ given by sending

... to $u_1,\ldots,u_n$. By Proposition 4 A is orthogonal. So relative to our standard basis $e_1, \ldots, e_n$ the reflection ϕ is given by

$$A \begin{pmatrix} -1 & & 0 \\ & 1 & \\ & & \ddots \\ 0 & & & 1 \end{pmatrix} A^{-1} = A \begin{pmatrix} -1 & & 0 \\ & 1 & \\ & & \ddots \\ 0 & & & 1 \end{pmatrix} {}^tA \ .$$

Conversely we see that such a matrix represents a reflection in the orthogonal complement of the vector $e_1 A$.

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

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

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

so

$$A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \ .$$

Thus the matrix giving reflection in $u^{\perp}$ is

$$\Phi = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

$$\Phi = \begin{pmatrix} -\cos 2\alpha & \sin 2\alpha \\ \sin 2\alpha & \cos 2\alpha \end{pmatrix} \ .$$

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

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

§. Exercises

1. Prove Proposition 1.

2. Prove Proposition 4.

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

$$\mathcal{O}(n) - SO(n) = \{BA \mid B \in SO(n)\} \ .$$

4. Show that any element of $SO(2)$ can be written as

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \ .$$

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

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 rotation through twice the angle between L_1 and L_2.

7. A matrix $A \in M_n(\mathbb{R})$ is said to be <u>idempotent</u> 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 <u>projection</u> of $\mathbb{R}^n$ onto its image.

What can you say about the determinant of a idempotent matrix ? What is the image of $\mathbb{R}^2$ under $A = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$.

8. A matrix A is <u>nilpotent</u> if some power of it is the zero matrix. For example $A = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix}$ has its third power zero. Prove that a nilpotent matrix is singular. Prove that any $A = (a_{ij})$ with

$$a_{ij} = 0 \quad \text{whenever} \quad i > j$$

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

$$a_{ij} = 0 \quad \text{whenever} \quad i > j \ .$$

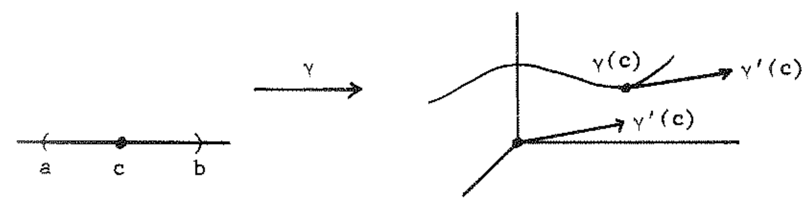
Prove that U is a group under matrix multiplication (the group of <u>unipotent</u> matrices in $M_n(\mathbb{R})$) .

Chapter 3
Homomorphisms

A. <u>Curves in a vector space</u>

We are going to define our first invariant of a matrix group, its <u>dimension</u>. 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 <u>curve</u> in V we mean a continuous function $\gamma:(a,b) \to V$ where (a,b) is an open interval in $\mathbb{R}$.



If $c \in (a,b)$ we say γ is <u>differentiable at</u> c if

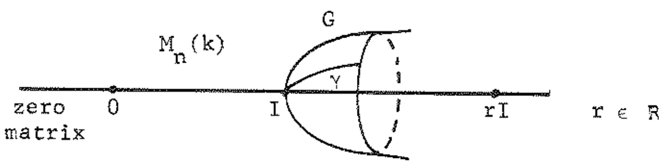
$$\lim_{h\to 0} \frac{\gamma(c+h) - \gamma(c)}{h}$$

exists. When this limit exists, it is a vector in V . We denote it $\gamma'(c)$ and call it the <u>tangent vector</u> to γ at $\gamma(c)$.

It is a standard result from calculus that if we choose a basis for V and thus represent γ as $(\gamma_1,\ldots,\gamma_n)$ (γ_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 $M_n(\mathbb{R})$, $M_n(\mathbb{C})$, $M_n(\mathbb{H})$ 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 <u>curve in</u> G is a curve in $M_n(k)$ with all values $\gamma(u)$ for $u \in (a,b)$ lying in G.



Suppose we have curves $\gamma,\sigma : (a,b) \to G$. Then we can define a new curve, the <u>product curve</u>, by

$$(\gamma\sigma)(u) = \gamma(u)\sigma(u) .$$

<u>Proposition</u> 1: <u>Let</u> $\gamma,\sigma : (a,b) \to G$ <u>be curves, both of which are differentiable at</u> $c \in (a,b)$. <u>Then the product curve</u> $\gamma\sigma$ <u>is differentiable at</u> c <u>and</u>

$$(\gamma\sigma)'(c) = \gamma(c)\sigma'(c) + \gamma'(c)\sigma(c) .$$

<u>Proof</u>: Let $\gamma(u) = (\gamma_{ij}(u))$, $\sigma(u) = (\sigma_{ij}(u))$. Then

$$(\gamma\sigma)(u) = (\sum_k \gamma_{ik}(u)\sigma_{kj}(u)) ,$$

so that

$$(\gamma\sigma)'(u) = (\sum_k \{\gamma_{ik}'(u)\sigma_{kj}(u) + \gamma_{ik}(u)\sigma_{kj}'(n)\})$$
$$= \gamma'(u)\sigma(u) + \gamma(u)\sigma'(u) .$$

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

<u>Proof</u>: 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

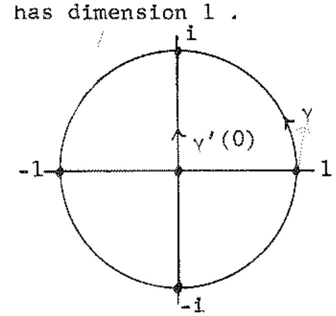
$$\sigma(u) = \gamma(ru) .$$

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

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

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

<u>Example</u> 1: $U(1)$ has dimension 1.



U(1)

Example 2: $\dim Sp(1) = 3$.

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

$$\gamma(t) = x(t) + i\, y(t) + j\, z(t) + k\, w(t)$$

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 $\gamma'(0) = 0 + i\, y'(0) + j\, z'(0) + k\, w'(0)$, as asserted.

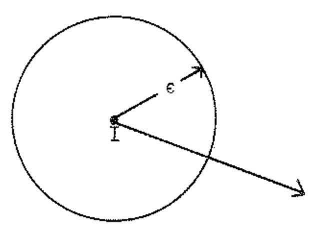
Conversely, let $q = i\mu + j\nu + k\lambda$ be any quaternion with zero real part. We claim that there exists a smooth curve γ in $Sp(1)$ such that $\gamma'(0) = q$. Indeed,

$$\gamma(t) = \sqrt{1 - \sin^2 \mu t - \sin^2 \nu t - \sin^2 \lambda t} + i \sin \mu t + j \sin \nu t + k \sin kt$$

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

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

The determinant function $\det : M_n(\mathbb{R}) \to \mathbb{R}$ is continuous and $\det(I) = 1$. So there is some ϵ-ball about I in $M_n(\mathbb{R})$ such that for each A in this ball $\det A \neq 0$; i.e., $A \in GL(n,\mathbb{R})$. If v is any vector in $M_n(\mathbb{R})$ define a curve σ in $M_n(\mathbb{R})$ by



$$\sigma(t) = tv + I .$$

Then $\sigma(0) = I$ and $\sigma'(0) = v$ and for small t, $\sigma(t)$ is in $GL(n,\mathbb{R})$. Hence the tangent space T is all of $M_n(\mathbb{R})$ which has dimension n^2 .

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

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

Definition: $A \in M_n(\mathbb{R})$ is said to be skew-symmetric if

$$A + {}^tA = 0 ;$$

i.e. if $a_{ij} = -a_{ji}$ for each i,j . In particular, the diagonal terms must all be zero.

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

Proof: The zero matrix is in $so(n)$, and if A,B belong to $so(n)$, then

$$(A + B) + {}^t(A + B) = A + {}^tA + B + {}^tB = 0 ,$$

so that $so(n)$ is closed under vector addition. It is also closed under scalar multiplication, for is $A \in so(n)$ and $r \in \mathbb{R}$, then ${}^t(rA) = r\,{}^tA$ so that $rA + {}^t(rA) = r(A + {}^tA) = 0$.

To check the dimension of $so(n)$ 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_{ij} 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

$$(n-1) + (n-2) + \ldots + 1 = \frac{n(n-1)}{2} \quad \text{of them} .$$

Definition: A matrix $B \in M_n(\mathbb{C})$ is skew-Hermitian if

$$B + {}^t\bar{B} = 0 .$$

Thus if $b_{jk} = c + id$, then $\bar{b}_{kj} = -b_{jk} = -c - id$ and $b_{kj} = -c + id$. In particular if $j = k$ we have $c + id = -c + id$, so that the diagonal terms of a skew-Hermitian matrix are purely imaginar[y].

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

Proposition 4: $su(n) \subset M_n(\mathbb{C})$ is a real vector space of dimensio[n]

$$n + 2\,\frac{n(n-1)}{2} = n^2 .$$

Proof: Exercise.

We make a similar definition for matrices in $M_n(\mathbb{H})$, and call $C \in M_n(\mathbb{H})$ skew-symplectic if

$$C + {}^t\bar{C} = 0 .$$

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 β is a curve through the identity $(\beta(0) = I)$

in $O(n)$ then $\beta'(0)$ is skew-symmetric

in $U(n)$ then $\beta'(0)$ is skew-Hermitian

in $Sp(n)$ then $\beta'(0)$ is skew-symplectic .

Proof: In each case we have that the product curve is constant

$$\beta(u)\,{}^t\bar{\beta}(u) = I .$$

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

Corollary:

$$\operatorname{Dim} O(n) \le \frac{n(n-1)}{2}$$

$$\operatorname{Dim} U(n) \le n^2$$

$$\operatorname{Dim} Sp(n) \le n(2n+1) .$$

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 ϕ to be continuous. From now on homomorphism always means continuous homomorphism. This being so, a curve

$$\rho : (a,b) \to G$$

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

Definition: A homomorphism $\phi : G \to H$ of matrix groups is smooth if for every differentiable curve ρ in G, $\phi\circ\rho$ is differentiable.

<u>Definition</u>: Let $\phi : G \to H$ be a smooth homomorphism of matrix groups. If $\gamma'(0)$ is a tangent vector to G at I we define a tangent vector $d\phi(\gamma'(0))$ to H at I by

$$d\phi(\gamma'(0)) = (\phi\circ\rho)'(0) .$$

The resulting map $d\phi : T_G \to T_H$ is called the <u>differential</u> of ϕ.

<u>Proposition</u> 6: $d\phi : T_G \to T_H$ <u>is a linear map</u>.

<u>Proof</u>: If $\rho'(0)$ and $\sigma'(0)$ are in T_G, consider

$$d\phi(a\rho'(0) + b\sigma'(0))$$

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

$$\{\phi\circ(a\rho + b\sigma)\}'(0) = \{a(\phi\circ\rho) + b(\phi\circ\sigma)\}'(0)$$

$$= a(\phi\circ\rho)'(0) + b(\phi\circ\sigma)'(0) = a\, d\phi(\rho'(0)) + b\, d\phi(\sigma'(0)) ,$$

proving that $d\phi$ is linear.

<u>Proposition</u> 7: <u>If</u> $G \overset{\phi}{\to} H \overset{\psi}{\to} K$ <u>are smooth homomorphisms, then so is</u> $\psi\circ\phi$ <u>and</u>

$$d(\psi\circ\phi) = d\psi\circ d\phi .$$

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

$$d(\psi\circ\phi)(\gamma'(0)) = (\psi\circ\phi\circ\gamma)'(0) = d\psi(\phi\circ\gamma)'(0) = d\psi\circ d\phi(\gamma'(0)) .$$

<u>Corollary</u>: <u>If</u> $\phi : G \to H$ <u>is a smooth isomorphism, then</u>

$d\phi : T_G \to T_H$ <u>is a linear isomorphism and</u> $\dim G = \dim H$.

<u>Proof</u>: $\phi^{-1}\circ\phi$ is the identity, so $d\phi^{-1}\, d\phi : T_G \to T_G$ is the identity. Thus $d\phi$ is injective and $d\phi^{-1}$ is surjective. $\phi\,\phi^{-1}$ is the identity, so $d\phi\, d\phi^{-1} : T_H \to T_H$ is the identity. Thus $d\phi^{-1}$ is injective and $d\phi$ is surjective. q.e.d.

<u>Exercises</u>

1. Let $\gamma : (-1,1) \to M_3(\mathbb{R})$ be given by

$$\gamma(t) = \begin{pmatrix} \cos t & \sin t & 0 \\ -\sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

Show that γ is a curve in $SO(3)$ and find $\gamma'(0)$. Show that $(\gamma^2)'(0) = 2\gamma'(0)$.

2. Let $\sigma : (-1,1) \to M_3(\mathbb{R})$ be given by

$$\sigma(t) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \cos t & \sin t \\ 0 & -\sin t & \cos t \end{pmatrix} .$$

Calculate $\sigma'(0)$. Write the matrix for $\gamma(t)\sigma(t)$ and verify that $(\gamma\sigma)'(0) = \gamma'(0) + \sigma'(0)$.

3. Let $\rho : (-1,1) \to M_3(\mathbb{C})$ be given by

$$\rho(t) = \begin{pmatrix} e^{i\pi t} & 0 & 0 \\ 0 & e^{i\frac{\pi t}{2}} & 0 \\ 0 & 0 & e^{i\frac{\pi t}{2}} \end{pmatrix} .$$

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

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

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

Show that α 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 $\dim H \le \dim G$.

6. Show that the set $sp(n)$ of $n \times n$ skew-symplectic 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

A. 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 T to G and G to T and study their properties. We will work with real matrices -- developments for $\mathbb{C}$ and $\mathbb{H}$ are quite analogous. (We need these maps to determine dimensions of some of our matrix groups.)

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

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

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

$$(I)_{ij} + (A)_{ij} + (\frac{A^2}{2!})_{ij} + (\frac{A^3}{3!})_{ij} + \dots$$

converges.

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

$$I + A + \frac{A^2}{2!} + \dots$$

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 $\leq \frac{nm^2}{2!}$.

The biggest element in the fourth term is $\leq \frac{n^2m^3}{3!}$, etc.

Any ij sequence is dominated by $1, m, \frac{nm^2}{2!}, \frac{n^2m^3}{3!}, \ldots,$

$\frac{n^{k-2}m^{k-1}}{(k-1)!}, \ldots$.

Applying the ratio test to this maximal sequence gives

$$\frac{n^{k-1}m^k}{k!} \frac{(k-1)!}{n^{k-2}m^{k-1}} = \frac{nm}{k} .$$

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

$$e^{A+B} = e^Ae^B .$$

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

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

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

$$= I + A + B + \frac{A^2}{2} + AB + \frac{B^2}{2} + \frac{A^3}{6} + \frac{A^2B}{2} + \frac{AB^2}{2} + \frac{B^3}{6} + \ldots .$$

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

Proof: A and $-A$ commute, so $I = e^0 = e^{A-A} = e^Ae^{-A}$ and $1 = (\det e^A)(\det e^{-A})$ and $\det e^A \neq 0$.

From this corollary we see that the map $\exp : M_n(\mathbb{R}) \to M_n(\mathbb{R})$, $(A) = e^A$, actually maps $M_n(\mathbb{R})$ into $GL(n,\mathbb{R})$.

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

Proof: We have $I = e^0 = e^{A+{}^tA} = e^Ae^{{}^tA} = (e^A)^t(e^A)$, proving e^A is orthogonal.

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

$$\exp : so(n) \to \mathcal{O}(n) .$$

It is important to note two things which Proposition 3 does not (i) it does not say that every orthogonal matrix is some e^A A skew-symmetric (i.e. it does not say $\exp : so(n) \to \mathcal{O}(n)$ is ...tive), and (ii) it does not say that e^A orthogonal implies is skew-symmetric. It is instructive to examine the case $n = 2$ detail.

The general 2×2 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^{α}, we calculate the powers of α. $\alpha^2 = \begin{pmatrix} -x^2 & 0 \\ 0 & -x^2 \end{pmatrix}$, $\alpha^3 = \begin{pmatrix} 0 & -x^3 \\ x^3 & 0 \end{pmatrix}$, $\alpha^4 = \begin{pmatrix} x^4 & 0 \\ 0 & x^4 \end{pmatrix}$, $\alpha^5 = \begin{pmatrix} 0 & x^5 \\ -x^5 & 0 \end{pmatrix}$, etc. Then

$$e^{\alpha} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix} + \frac{1}{2!}\begin{pmatrix} -x^2 & 0 \\ 0 & -x^2 \end{pmatrix} + \frac{1}{3!}\begin{pmatrix} 0 & -x^3 \\ x^3 & 0 \end{pmatrix} + \frac{1}{4!}\begin{pmatrix} x^4 & 0 \\ 0 & x^4 \end{pmatrix} + \frac{1}{5!}\begin{pmatrix} 0 & x^5 \\ -x^5 & 0 \end{pmatrix} + \ldots .$$

From the 1,1 position we get

$$1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots = \cos x , \text{ etc.}$$

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 2×2 skew-symmetric matrix α we have

$$\det e^{\alpha} = 1 , \text{ i.e., } e^{\alpha} \in SO(2) .$$

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

Note also that $e^{\alpha} = I$ does not imply that α is the zero matr[illegible] ($\alpha = \begin{pmatrix} 0 & 2\pi \\ -2\pi & 0 \end{pmatrix}$ 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.

<u>Proposition</u> 4: <u>If</u> A,B <u>are</u> $n \times n$ <u>matrices over</u> $k \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$ <u>and</u> B <u>is nonsingular</u>, <u>then</u>

$$e^{BAB^{-1}} = Be^{A}B^{-1} .$$

<u>Proof</u>: $(BAB^{-1})^n = (BAB^{-1})(BAB^{-1})\ldots(BAB^{-1}) = BA^nB^{-1}$, and $B(C+D)B^{-1} = BCB^{-1} + BDB^{-1}$; these and the definition of the exponential of a matrix yield the result.

[illegible] <u>Logarithm</u>

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

$$\log X = (X-I) - \frac{(X-I)^2}{2} + \frac{(X-I)^3}{3} - \frac{(X-I)^4}{4} + \ldots .$$

<u>Proposition</u> 5: <u>For</u> X <u>near</u> I <u>this series converges</u>.

<u>Proof</u>: Let $Y = X - I$ and $Y = (y_{ij})$, and suppose each [illegible] . Then

$$|(Y)_{ij}| \le \epsilon , \quad |(\frac{Y^2}{2})_{ij}| \le \frac{n\epsilon^2}{2} ,$$

$$|(\frac{Y^3}{3})_{ij}| \le \frac{n^2\epsilon^3}{3} , \quad |(\frac{Y^k}{k})_{ij}| \le \frac{n^{k-1}\epsilon^k}{k} .$$

[illegible] ratio test gives

$$\frac{n^k\epsilon^{k+1}}{k+1} \frac{k}{n^{k-1}\epsilon^k} = \frac{k}{k+1} n\epsilon \to n\epsilon .$$

[illegible] series converges for any X such that each entry of $X - I$ [illegible] in magnitude.

<u>Proposition</u> 6: <u>In</u> $M_n(\mathbb{R})$ <u>let</u> U <u>be a neighborhood of</u> I <u>on</u>

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. $A \in V \Rightarrow e^A \in U \Rightarrow \log e^A$ is defined (i.e., the series converges). $e^A - I = A + \frac{A^2}{2!} + \frac{A^3}{3!} + \ldots$. So

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

$$= A + [\frac{A^2}{2!} - \frac{A^2}{2}] + [\frac{A^3}{6} - \frac{A^3}{2} + \frac{A^3}{3}] + \ldots = A .$$

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

$$e^{\log X} = \{I + (X-I) - \frac{(X-I)^2}{2} + \ldots\} + \frac{1}{2!}\{(X-I) - \frac{(X-I)^2}{2} + \ldots\}^2 + \frac{1}{3!}\{(X-I) - \frac{(X-I)^2}{2} + \ldots\}^3 + \ldots$$

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

$$= X .$$

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.

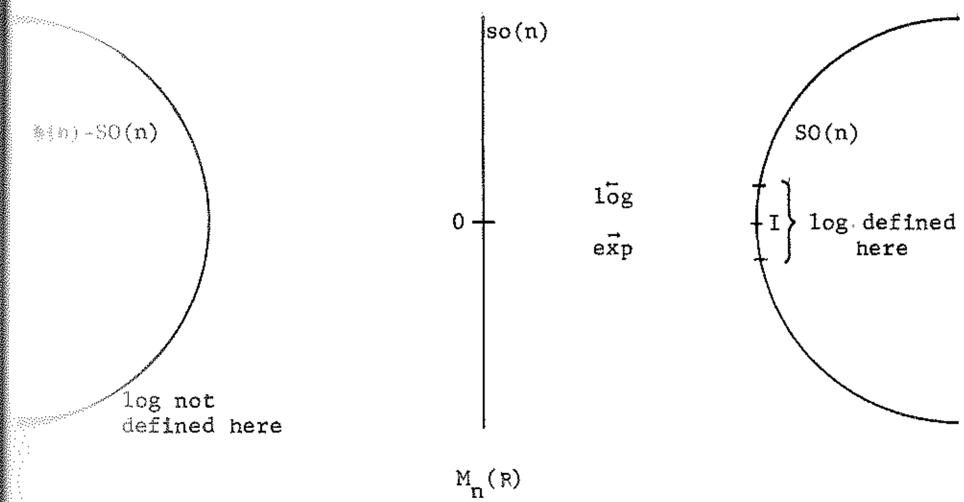
Proof: $e^{\log XY} = XY = e^{\log X} e^{\log Y} = e^{\log X + \log Y}$, and e is one-to-one near 0 .

Next X and ${}^tX$ commute so that $\log X$ and $\log {}^tX$ commute. If X is orthogonal

$$I = X^tX$$

$0 = \log X^tX = \log X + \log {}^tX = \log X + {}^t(\log X)$ showing $\log X$ is skew-symmetric.

We will be able to say more after we have done a little topology. The picture will be



One-parameter subgroups

Definition: A one-parameter subgroup γ in a matrix group G smooth homomorphism

$$\gamma : \mathbb{R} \to G .$$

that it suffices to know γ on some open neighborhood U of 0 . For $x \in \mathbb{R}$, some $\frac{1}{n}x \in U$ and $\gamma(x) = (\gamma(\frac{1}{n}x))^n$.

<u>Example</u>: Let $k \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$ and $A \in M_n(k)$. Then

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

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

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

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

<u>Proof</u>: Let $\sigma(u) = \log \gamma(u)$. Then σ 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.

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

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

Now $u + v = v + u$ and γ 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 $\mathfrak{G}(n,k)$.

<u>Proposition</u> 9: <u>Let A be a tangent vector to $\mathfrak{G}(n,k)$. Then there exists a unique one-parameter subgroup γ in $\mathfrak{G}(n,k)$ such that</u>

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

<u>Proof</u>: By definition $A = \rho'(0)$ where ρ is a curve in $\mathfrak{G}(n,k)$. Thus

$$\rho(u)\,{}^t\overline{\rho(u)} = I$$

so that

$$\rho'(0) + {}^t\overline{\rho'(0)} = 0 \text{ , i.e.,}$$

$$A + {}^t\bar{A} = 0 .$$

Now $\gamma(u) = e^{uA}$ is a one-parameter subgroup of $GL(n,k)$, but it lies in $\mathfrak{G}(n,k)$ because

$$\gamma(u)\ {}^t\overline{\gamma(u)} = e^{uA} e^{u\,{}^t\bar{A}} = e^{u(A + {}^t\bar{A})} = I .$$

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

Taking $k = \mathbb{R}$ we have that the tangent space to $\mathfrak{G}(n) = \mathfrak{G}(n,\mathbb{R})$ is $so(n)$, the vector space of all skew-symmetric $n \times n$ matrices. Thus $\dim \mathfrak{G}(n) = \dim so(n) = \frac{n(n-1)}{2}$.

Taking $k = \mathbb{C}$ we have that the tangent space to $U(n) = \mathfrak{G}(n,\mathbb{C})$

is su(n), the vector space of all skew-Hermitian $n \times n$ complex matrices. Thus

$$\dim U(n) = \dim su(n) = n^2 .$$

Taking $k = \mathbb{H}$ 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.

<u>Definition</u>: The <u>trace</u> of a matrix $A = (a_{ij})$ is the sum of the diagonal terms;

$$Tr(A) = a_{11} + a_{22} + \ldots + a_{nn} .$$

We clearly have

(i) $Tr(A+B) = Tr(A) + Tr(B)$, and $Tr(aA) = a\,Tr(A)$, (so Tr is linear).

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

(ii) $Tr(AB) = Tr(BA)$.

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 $\mathbb{R}$ and $\mathbb{C}$ are commutative one easily checks that these are equal.

Clearly

(iii) $Tr(I) = n$.

Also

(iv) If B is nonsingular, then

$$Tr(BAB^{-1}) = Tr(A) .$$

<u>Proof</u>: By (ii), $Tr(B(AB^{-1})) = Tr((AB^{-1})B) = Tr(AI) = TrA$.

Now we come to the crucial relation.

<u>Theorem</u>: <u>If</u> A <u>is a real or complex matrix, then</u>

$$(+) \qquad e^{Tr(A)} = \det (e^A) .$$

We will prove this later, but a few comments are in order here. First off, (+) looks wrong because the left hand side depends only on the diagonal elements of A and it is not immediately clear that this is true for the right-hand side. The point is that det and exp are also invariant under conjugation just as (iv) for Tr; so if B is nonsingular

$$\det (e^{BAB^{-1}}) = \det (Be^AB^{-1}) = \det (e^A) .$$

We will prove (+) once we have found how to put matrices in simpler forms by conjugation (in Chapter VIII).

Suppose we know (+). The linear map

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

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 $\mathrm{Tr}(u(n))$ is all of $i\mathbb{R}$. From the rank theorem in linear algebra we know that (all vector spaces being over $\mathbb{R}$ now)

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

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

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

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

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

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

$$[A,B] = AB - BA .$$

Then $so(n)$, $su(n)$ and $sp(n)$ are closed under $[\ ,\ ]$.

Proof: We need to show that

$$(AB - BA) + {}^t\overline{(AB - BA)} = 0 .$$

The left-hand side is

$$AB-BA+{}^t\bar{B}\,{}^t\bar{A}-{}^t\bar{A}\,{}^t\bar{B} = AB+(A\,{}^t\bar{B}-A\,{}^t\bar{B})-BA+(-B\,{}^t\bar{A}+B\,{}^t\bar{A})+{}^t\bar{B}\,{}^t\bar{A}-{}^t\bar{A}\,{}^t\bar{B}$$

$$= A(B+{}^t\bar{B})-(A+{}^t\bar{A})\,{}^t\bar{B}-B(A+{}^t\bar{A})+(B+{}^t\bar{B})\,{}^t\bar{A}$$

$$= 0 .$$

Thus $so(n)$, $su(n)$ and $sp(n)$ become algebras (over $\mathbb{R}$) with this bracket multiplication. This product has some obvious properties.

(i) $[A,B] = -[B,A]$

(ii) $[A,B+C] = [A,B] + [A,C]$

$[A+B,C] = [A,C] + [B,C]$

(iii) For $r \in \mathbb{R}$, $r[A,B] = [rA,B] = [A,rB]$.

Finally, $[\ ,\ ]$ has one nonobvious property.

(iv) $[A,[B,C]] + [B,[A,C]] + [C,[A,B]] = 0$.

Property (iv) is called the Jacobi identity and its proof is a routine verification.

Definition: A real vector space with a product satisfying (i)...(iv) is called a Lie algebra. (One could clearly consider complex Lie algebras, but we will have no occasion to do so.)

Let us consider low dimensional Lie algebras. For dim 1 the vector space is just $\mathbb{R}$ and if $x,y \in \mathbb{R}$ we have

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

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

Consider $\mathbb{R}^2$ with basis e_1, e_2

We must have

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

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

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

so

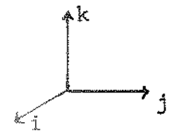
$$b(ae_1 + be_2) + [e_1,(-ae_1 - be_2)] = 0$$

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.

$$so(3) = \left\{ \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix} \middle| \ a,b,c \in \mathbb{R} \right\}$$

clearly has dimension three. Also using basis



and defining
$$[i,j] = k$$
$$[j,k] = i$$
$$[k,i] = j$$

gives a 3-dimensional Lie algebra.

Exercises

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

2. Let $B \in O(3) - SO(3)$. Show that the series for $\log B$ does not converge.

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

4. Prove that any two nontrivial Lie algebras on $\mathbb{R}^2$ are isomorphic as Lie algebras.

5. Show that the two 3-dimensional Lie algebras defined above are isomorphic.

Chapter 5
SO(3) and Sp(1)

A. The homomorphism $\rho : S^3 \to SO(3)$

We have seen that $Sp(1)$, which is all quaternions of unit length, is just the unit 3-sphere in $\mathbb{R}^4$ $(= \mathbb{H})$. 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 $\mathbb{R}^4$ to $\mathbb{R}^4$.

Proof: As vector spaces over $\mathbb{R}$, $\mathbb{H}$ and $\mathbb{R}^4$ are the same. So L_q is surely a linear map of $\mathbb{R}^4$, for if $a,b \in \mathbb{R}$ and $\alpha,\beta \in \mathbb{H}$ we have

$$L_q(a\alpha + b\beta) = q(a\alpha + b\beta) = aq\alpha + bq\beta = aL_q(\alpha) + bL_q(\beta) .$$

To see that L_q is orthogonal, it suffices to show that L_q preserves the perpendicularity (using $\langle\ ,\ \rangle$ for $\mathbb{R}^4$) of the four unit vectors $1,i,j,k$. For example, let $q = a + ib + jc + kd$ and

calculate $\langle L_q(i), L_q(j)\rangle$ (using $\langle\ ,\ \rangle$ for $\mathbb{R}^4$). We get $ad + bc - bc - da = 0$. For $\langle L_q(1), L_q(i)\rangle$ we get

$$\langle a + ib + jc + kd,\ ai - b - kc + jd\rangle = -ab + ab + dc - dc = 0 .$$

The computations for other pairs of basis vectors are similar.

Definition of ρ: For $q \in S^3$ and $\alpha \in \mathbb{H}$ we define

$$\rho(q)(\alpha)\ q\alpha\bar{q} .$$

That is, we do a left translation by q and a right translation by $\bar{q}$. By Proposition 1 this is an orthogonal map of $\mathbb{R}^4$ to $\mathbb{R}^4$; i.e., $\rho(q) \in \mathcal{O}(4)$.

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

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

Note also that $\rho(\bar{q})$ is the inverse of $\rho(q)$ in the group $\mathcal{O}(4)$ since $\rho(q)\rho(q^{-1})(\alpha) = q(\bar{q}\alpha q)\bar{q} = \alpha$ and similarly for $\rho(\bar{q})\rho(q)$. Together these two observations imply that $\rho(q)$ maps the 3-space spanned by i,j,k to itself (Exercise #3). Thus $\rho(q)$ can be considered as an element of $\mathcal{O}(3)$ (Exercise #4).

Fact (to be proved after Chapter VI): $\rho(q)$ is in $S\mathcal{O}(3)$.

Proposition 2: $\rho : S^3 \to S\mathcal{O}(3)$ is a surjective homomorphism and

$$\mathrm{Ker}(\rho) = \{1,-1\} \subset S^3 .$$

Proof: If $q_1, q_2 \in S^3$ and $\alpha \in \mathrm{Span}\langle i,j,k\rangle$, then

$$\rho(q_1q_2)(\alpha) = q_1q_2\alpha\overline{q_1q_2} = q_1(q_2\alpha\overline{q_2})\bar{q} = \rho(q_1)\rho(q_2)(\alpha) .$$

Thus ρ is a homomorphism.

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

Finally we need to show that ρ is surjective. This will be quite easy once we know some topology (Chapter VI)--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 $SO(3)$ which leaves k fixed, sends i to j and sends j to $-i$.

Let $q = a + ib + jc + kd$. We want $(a + ib + jc + kd)(k)(a - ib - jc - kd) = k$, or $(a + ib + jc + kd)(ka - jb + ic + d) = k$, so

$$ad - bc + bc - ad = 0 \quad \text{(automatically)} ,$$

$$ac + bd + ac + bd = 0 \quad \text{or} \quad 2(ac + bd) = 0 ,$$

$$-ab + cd + dc - ab \quad \text{or} \quad 2(cd - ab) = 0 .$$

$$a^2 + d^2 - b^2 - c^2 = 1 .$$

Now

$$a^2 + d^2 + b^2 + c^2 = 1$$

so

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

So the only condition on q such that $\rho(q)k = k$ is that $q = a + dk$ (with $a^2 + d^2 = 1$).

Next we want $\rho(q)i = j$.

$$(a + kd)(i)(a - kd) = j$$

$$(a + kd)(ia + jd) = j$$

$$a^2 - d^2 = 0 , \ a = \pm d$$

$$ad + ad = 1 , \quad \text{if} \quad a = d , \ 2a^2 = 1 ,$$

and we can't have $a = -d$. Finally we insist that $\rho(q)j = -i$. So

$$(a + ka)(j)(a - ka) = -i$$

$$(a + ka)(ja - ia) = -1$$

$$-2a^2 = -1$$

$$a^2 - a^2 = 0 .$$

So $q = \frac{1}{\sqrt{2}} + k\frac{1}{\sqrt{2}}$ or $q = -\frac{1}{\sqrt{2}} - k\frac{1}{\sqrt{2}}$. Both will give the desired element of $SO(3)$. (This should be enough to convice us that we should not try the general proof of surjectivity at this stage.)

Note that this <u>does not prove</u> that S^3 and $SO(3)$ are not isomorphic. ρ is not an isomorphism, but one might exist. In the next section we give a fairly easy proof that $S^3 \not\approx SO(3)$.

B. Centers

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 induces an isomorphism of their centers. We will show that $S^3 \not\approx SO(3)$ 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 $\{I\}$.

Proof: Since real quaternions commute with all quaternions, it is clear that $\{1,-1\} \subset \text{Center } S^3$. Conversely, suppose $q = a + ib + jc + kd \in S^3$ is in the center. Then $qi = iq$ gives

$$ai - b - ck + dj = ai - b + ck - dj$$

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 $T \subset SO(3)$. Consider the standard basis $e_1 = (1,0,0)$, $e_2 = (0,1,0)$, $e_3 = (0,0,1)$ for R^3.

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

Choose $B \in T$ 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 $BAe_3 = ae_2 - be_1 + ce_3$, whereas $ABe_3 = Ae_3$; this implies $a = 0 = b$ and since A preserves length, we must have $c = 1$, or $c = -1$.

Thus A induces an orthogonal map of the e_1e_2 plane. Actually, it is a rotation because:

Sublemma: Any element of $\mathcal{O}(2)$ which commutes with all rotations, is itself a rotation.

Let $\phi : R^2 \to R^2$ denote such an element of $\mathcal{O}(2)$. For any rotation

$$t = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

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

$$\alpha\cos\theta - \beta\cos\theta = \alpha\cos\theta + \gamma\sin\theta$$

$$\alpha\sin\theta + \beta\cos\theta = \beta\cos\theta + \delta\sin\theta,$$

holding for all θ. So $\gamma = -\beta$ and $\alpha = \delta$. Thus

$$\phi = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \text{ and } \det\phi = \alpha^2 + \beta^2.$$

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

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

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.

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 <u>reflexive</u>, $x \sim x$, since $xx^{-1} = e \in H$. It is <u>symmetric</u>, $x \sim y \Rightarrow y \sim x$, since $xy^{-1} \in H \Rightarrow (xy^{-1})^{-1} = yx^{-1} \in H$. It is <u>transitive</u>, $x \sim y$ and $y \sim z \Rightarrow x \sim z$, since $xy^{-1} \in H$ and $yz^{-1} \in H$ imply that $(xy^{-1})(yz^{-1}) = xz^{-1} \in H$. 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 $Hx = Hy \Leftrightarrow xy^{-1} \in H \Leftrightarrow y \in C(x) \Leftrightarrow x \in C(y)$. These equivalence classes are called <u>right cosets</u> of H.

<u>Example</u>: Let $G = S^3$ ($= Sp(1)$) and $H = \{1,-1\}$. Then $Hq = \{q,-q\} = H(-q)$ so each equivalence class contains exactly two points of S^3.

<u>Example</u>: 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 $\{a + ib \mid a^2 + b^2 = 1\}$. Thus S^3 can be divided up into circles.

One defines left cosets in a similar manner

$$xH = \{xh \mid h \in H\} .$$

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) $\Leftrightarrow xH = Hx$ for every $x \in G$.

Let G/H denote the set whose elements are the rightcosets 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 $Hx = Hz$ and $Hy = Hw$. We must show that $Hxy = Hzw$. Well, $xy(zw)^{-1} = xyw^{-1}z^{-1}$ 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 η 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

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

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

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

Finally, $[G,G]x[G,G]y = [G,G]xy = [G,G]yx = [G,G]y[G,G]x$ since

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

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.

<u>Proposition</u> 6: <u>For</u> $x \in G$ <u>define</u>

$$\psi(x) : G \to G$$

<u>by</u> $\psi(x)(y) = xyx^{-1}y^{-1}$. <u>Then</u> G/C <u>has</u> <u>nontrivial</u> <u>center</u> $\Leftrightarrow \exists x \in G-C$ <u>such</u> <u>that</u>

$$\psi(x)(G) \subset C .$$

<u>Proof</u>: $\Leftarrow$

$x \notin C \Rightarrow Cx \neq C$ 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 $Cx \in$ center G/C .

$\Rightarrow$

Conversely, $Cx \neq C$ with Cx in the center implies $CxCy = Cxy = Cyx = CyCx$ so that $xyx^{-1}y^{-1} \in C$ for all $y \in G$.

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

<u>Corollary</u>: <u>If</u> G <u>is</u> <u>connected</u> <u>and</u> C <u>is</u> <u>discrete</u> (<u>in</u> <u>particular</u> <u>if</u> C <u>is</u> <u>finite</u>), <u>then</u> G/C <u>has</u> <u>no</u> <u>center</u>.

D. <u>Exercises</u>

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

$$R_{x^{-1}} \circ L_x(g) = xgx^{-1} .$$

Show that $R_{x^{-1}} \circ L_x$ is an isomorphism of G onto G .

2. Do one more step in the proof of Proposition 1 by showing $\langle L_q(i), L_q(k) \rangle = 0$.

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

5. Show that $\rho(i)$, $\rho(j)$, $\rho(k)$ are all in $SO(3)$.

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

7. Let $\phi : G \to K$ be a surjective homomorphism of groups and $H = \operatorname{Ker} \phi$. Then we have

$$\begin{array}{l} G \xrightarrow{\phi} K \\ \downarrow \eta \\ G/H \end{array} .$$

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 $SO(3)$ is not a normal subgroup.

9. Show that the subgroup $H = \{a + ib \mid a^2 + b^2 = 1\}$ of $Sp(1)$

is not a normal subgroup.

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

Chapter 6
Topology

A. Introduction

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

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

There are certain <u>topological</u> 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 B 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 <u>manifold</u> 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 $\rho : Sp(1) \to SO(3)$ (defined in Chapter V) is surjective.

B. Continuity of functions, open sets, closed sets

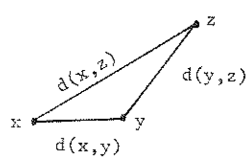
Definition: A metric d on a set S is a way of assigning to each $x,y \in S$ a real number $d(x,y)$ (the distance from x to y) in such a way that:

(i) $d(x,y) \geq 0$ and $d(x,y) = 0 \Leftrightarrow x = y$,

(ii) $d(x,y) = d(y,x)$,

(iii) $d(x,y) + d(y,z) \geq d(x,z)$.

ondition (iii) is called the triangle inequality.



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

$$x = (x_1,\dots,x_n) \quad y = (y_1,\dots,y_n)$$

defined an inner product

$$\langle x,y\rangle = x_1y_1 + \dots + x_ny_n .$$

$d(x,y) = \sqrt{\langle x-y,x-y\rangle}$. (Thus we define $d(x,y)$ to be the length the vector $x - y$.)

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

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

$$\langle x,x\rangle \geq 0 \text{ and } \langle x,x\rangle = 0 \Leftrightarrow x = 0$$

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

$$\langle x+ty,x+ty\rangle \geq 0 .$$

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 ≥ 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

$$(*) \qquad \langle x,y\rangle^2 \leq \langle x,x\rangle\langle y,y\rangle .$$

The inequality (*) is the Schwarz inequality.

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

$$(**) \qquad \langle x-y,y-z\rangle \leq \sqrt{\langle x-y,x-y\rangle}\,\sqrt{\langle y-z,y-z\rangle} .$$

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

$$B(x,r) = \{y \in \mathbb{R}^n \mid d(x,y) < r\}$$

d call this the open ball with center x and radius r. Open balls euclidean spaces allow us to give a fairly direct generalization of e notion of continuity of a function on $\mathbb{R}$ to functions defined on aces 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 $B(f(a),\epsilon)$ in $\mathbb{R}^m$, there exists an open l $B(a,\delta)$ in $\mathbb{R}^n$ such that any point $x \in A \cap B(a,\delta)$ satisfies

$$f(x) \in B(f(a),\epsilon) .$$

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

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

It is important to notice that the continuity of f depends on 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 $A \subset \mathbb{R}$ 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

$$\delta_i = d(a,b_i) \ , \quad i = i,\ldots,k$$

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

__Proposition__ 2: __If__ $A \subset \mathbb{R}^n$ __and__

$$A \xrightarrow{f} \mathbb{R}^m \ , \ f(A) \xrightarrow{g} \mathbb{R}^p$$

__are continuous__, __then__ $g \circ f$ __is continuous__.

__Proof__: Let $a \in A$ and $\epsilon > 0$. Since g is continuous, there exists $\eta > 0$ such that every element of $f(A)$ in $B(f(a),\eta)$ is sent by g into $B(g(f(a)),\epsilon)$. Since f is continuous there exists $\delta > 0$ such that every element of A in $B(a,\delta)$ is sent by f into $B(f(a),\eta)$ and is then sent by g into $B(g(f(a)),\epsilon)$.

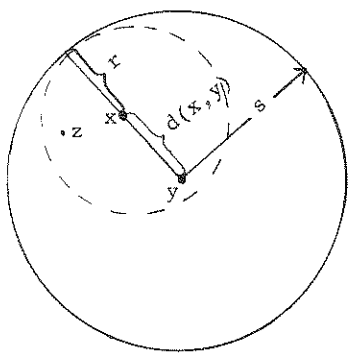
q.e.d.

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 [in s]ome $B(x,r) \subset U$ (where r will depend on x) .

Clearly $\mathbb{R}^n$ is an open set. It is not quite so clear that the [empt]y set ϕ is open; but since there is no $x \in \phi$, there is no [requ]irement that some $B(x,r)$ be contained in ϕ .

Example: Any open ball $B(y,s)$ is an open set. Let $x \in B(y,s)$ [(i.e]., $d(x,y) < s$) . We must find $r > 0$ such that $B(x,r) \subset B(y,s)$. [Indeed], let $r = s - d(x,y)$. If $z \in B(x,r)$ then $d(z,x) < s - d(x,y)$



thus

$$d(z,x) + d(x,y) < s .$$

[Usin]g the triangle inequality we have

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

[prov]ing $z \in B(y,s)$.

Example: $(0,1) = \{x \in \mathbb{R} \mid 0 < x < 1\}$ is an open set in $\mathbb{R}$ [(bec]ause it is the open ball $B(\frac{1}{2},\frac{1}{2})$) . But

$(0,1] = \{x \in \mathbb{R} \mid 0 < x \leq 1\}$ 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.

$[0,1] = \{x \in \mathbb{R} \mid 0 \leq x \leq 1\}$ 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 δ 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

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

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

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

uch a function may be called a <u>path</u> from x to y in D .

xamples: $\mathbb{R}^n$ is connected because

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

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

$= \{x \in \mathbb{R} \mid x \neq 0\}$ is not connected. For example, no path

-1 to 1 in $\mathbb{R}$ can lie in D .

$= \{(x_1,x_2) \in \mathbb{R}^2 \mid (x_1,x_2) \neq (0,0)\}$ is connected.

mportant example: $\mathbb{O}(n) \subset \mathbb{R}^{n^2}$ is not connected. The matrices

$$\begin{pmatrix} 1 & & \bigcirc \\ & 1 & \\ & & \ddots \\ \bigcirc & & & 1 \end{pmatrix} \quad \text{and} \quad I = \begin{pmatrix} 1 & & \bigcirc \\ & 1 & \\ & & \ddots \\ \bigcirc & & & 1 \end{pmatrix}$$ are in $\mathbb{O}(n)$. If $\gamma : [0,1]$

is a path from A to I , then the composite function

$$[0,1] \xrightarrow{\gamma} \mathbb{O}(n) \xrightarrow{\det} \mathbb{R}$$

be continuous (Proposition 2) . But it would be a path in

$\subset \mathbb{R}$ from -1 to 1 , contradicting the existence of γ .

Recall that $so(n) \subset M_n(\mathbb{R})$ consists of skew-symmetric matrices

nat if $A \in so(n)$, then $\exp A \subset \mathbb{O}(n)$.

<u>Proposition</u> 3: exp <u>maps</u> so(n) <u>into</u> SO(n) .

<u>Proof</u>: For $B \in so(n)$ the path

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

path from $e^0 = I$ to e^B . As seen above, this implies that

$^B = +1$ so $e^B \in SO(n)$.

<u>Proposition</u> 4: <u>Let</u> $D \subset \mathbb{R}^n$ <u>be connected and</u>

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

<u>be continuous, then</u> $f(D)$ <u>is connected.</u>

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

<u>Definition</u>: A subset W of $\mathbb{R}^n$ is <u>bounded</u> 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 $f(\mathbb{R}) = \{y \in \mathbb{R} \mid y > 0\}$ is not closed.

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

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

<u>Proposition</u> 5: <u>If</u> $C \subset \mathbb{R}^n$ <u>is compact and</u>

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

<u>is continuous, then</u> $f(C)$ <u>is compact.</u>

The proof is relegated to an appendix.

ubspace 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 ere 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 $\frac{1}{2},1] = \{x \in \mathbb{R} \mid \frac{1}{2} < x \leq 1\}$ is an open set in W , but not an 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 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 ace 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.) urse, not every open set is an open ball, but open balls suffice ve all open sets by taking unions (the "empty" union being ed).

Definition: A collection $\mathcal{V} = \{V_\alpha\}$ of open sets in $\mathbb{R}^n$ is a for open sets if every open set in $\mathbb{R}^n$ is a union of some of V_α's .

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

$$\{B(x,r) \mid x = (x_1,\dots,x_n) \text{ with each } x_i \text{ rational, and } r \text{ is rational}\}$$

(See Proposition 7.)

For a subset W of $\mathbb{R}^n$ we know which are the open sets in W 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 "minimial" 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 $s_1,s_2,s_3,\dots$; 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

$$1, 2, \frac{1}{2}, \frac{3}{2}, 3, \frac{1}{3}, \frac{2}{3}, \frac{4}{3}, \frac{5}{3}, \frac{7}{3}, \frac{8}{3}, 4, \frac{1}{4}, \dots$$

is a sequence containing all positive rationals. Similarly,

..., 2, -2, $\frac{1}{2}$, $-\frac{1}{2}$, $\frac{3}{2}$, $-\frac{3}{2}$, 3, -3,... contains all of $\mathbb{Q}$.

The set $I = [0,1] = \{x \in \mathbb{R} \mid 0 \le x \le 1\}$ is not countable. We ... this contrapositively. Suppose

$$r_1, r_2, r_3, \ldots$$

... list of all elements of I . It suffices to give an element of ...ich cannot be in the list. Express the r_i's as decimals

$$r_i = .x_{i_1} x_{i_2} x_{i_3} \ldots \quad .$$

... $r = .y_1 y_2 y_3 \ldots$ where $y_j = 5$ if $x_{jj} \ne 5$ and $y_j = 1$ if ... $= 5$. Then

$$r \ne r_1 \quad \text{because} \quad y_1 \ne x_{11}$$

$$r \ne r_2 \quad \text{because} \quad y_2 \ne x_{22} \ , \quad \text{etc.}$$

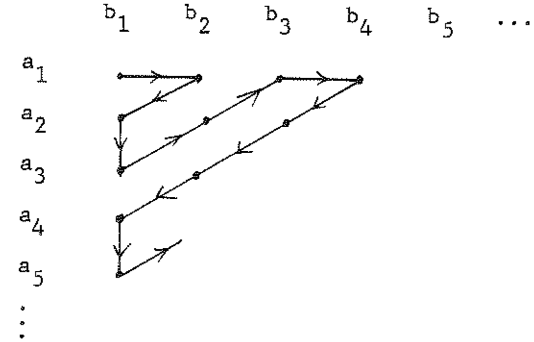
... $r \in I$.

<u>Proposition</u> 6: <u>If</u> A <u>and</u> B <u>are countable sets</u>, <u>then so is</u> ...r <u>cartesian product</u> $A \times B$.

<u>Proof</u>: Let $A = \{a_1, a_2, a_3, \ldots\}$ and $B = \{b_1 b_2, b_3, \ldots\}$. Then ...an write

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

...use following the path shown below will include all of $A \times B$.



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

<u>Proof</u>: 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.

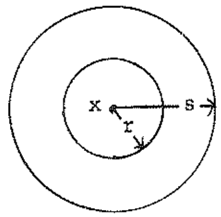
Let V be any open set in $\mathbb{R}^n$. To show that V is a union of elements of C it suffices to show that for $y \in V$ some $B(x,r) \in C$ contains y and lies in V . (For then V is the union of such $B(x,r)$ -- one for each $y \in V$.) Since V is open some

$$B(y,s) \subset V \ .$$

Choose x with all coordinates rational such that

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

and let r be a rational number satisfying $\frac{s}{3} < r < \frac{s}{2}$.



$y \in B(x,r)$ and $B(x,r) \subset B(y,s) \subset V$. q.e.d.

This proposition will be used in an essential way in our study of
al tori in matrix groups.

anifolds

Definition: By a space we mean some subset of some R^n with the
ace topology. A map

$$f : X \to Y$$

aces is a homeomorphism if it is one-to-one, f is continuous,
f^{-1} is continuous.

Example: $f(x) = e^{ix}$ is a one-to-one continuous map of
) $\subset R$ to the unit circle $S^1 \subset R^2$. But f is not a homeo-
ism because f^{-1} is not continuous. (It "tears" the circle
)

Example: Let G be a matrix group and $x \in G$. Then left
slation L_x by x $(L_x(g) = xg)$ is a homeomorphism $L_x : G \to G$.
rcise.)

A manifold is a space which "locally" looks like some R^n .

Definition: A space X is an n-manifold if each $x \in X$ lies in some open set homeomorphic to some $B(o,r) \subset R^n$. 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 $B(o,r) \subset V$ 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 homeomorphisim) 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: $GL(n,k)$ is not closed, but $\mathcal{O}(n,k)$ is closed.

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

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

To see that $M_n(k) - \mathcal{O}(n,k)$ is open, suppose $B \in M_n(k) - \mathcal{O}(n,k)$. Then there exists $x,y \in k^n$ such that

$$\langle xB, yB\rangle \neq \langle x,y\rangle .$$

e $\langle\ ,\ \rangle$ is continuous, there is some open ball $B(B,s)$ in
) such that for $B' \in B(B,s)$ we have $\langle xB', yB'\rangle \neq \langle x,y\rangle$. Thus
$\mathbb{R}(n,k)$.

q.e.d.

We finish this chapter with a result which will be of substantial
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
$M - N$ and $x \in N$. Since M is connected, there exists a path

$$\rho : [0,1] \to M$$

a $\rho(0) = x$ and $\rho(1) = y$. Then $\rho^{-1}(M - N)$ is an open set in
1] (see Exercise 3) and it contains 1 but not 0. Let t_0
the largest element of the closed set $I - \rho^{-1}(M - N)$. Then

(i) every $B(t_0,\epsilon)$ contains points of $\rho^{-1}(M - N)$, but

(ii) since N is a manifold there is some open neighborhood U
$\rho(t_0)$ in N. By continuity of ρ some $B(t_0,\epsilon)$ maps by ρ
o N. This contradiction shows $N = M$.

Corollary: The map $\rho : Sp(1) \to SO(3)$ (see Chapter V) is sur-
tive.

Proof: Since ρ is a homeomorphism on some neighborhood of each
nt, we see that the image $\rho(Sp(1))$ is a 3-manifold. Since ρ is
tinuous, 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 $\{Ae_1, Ae_2, Ae_3\}$ is an orthonormal basis for $\mathbb{R}^3$. Let B be a rotation sending e_1 to Ae_1 and leaving the direction perpendicular to the (e_1, Ae_1) plane fixed. (If $Ae_1 = e_1$ proceed directly to the next step. If e_1 and Ae_1 are antipodal on S^2, then we have two choices for B.) Clearly, there is a path ω from I to B in $SO(3)$. Now $Be_1 = Ae_1$ so Be_2 and Be_3 are an orthonormal basis for the plane perpendicular to Ae_1. Let C be a rotation of this plane sending Be_2 to Ae_2 and Be_3 to Ae_3. (If we can't do this, we would have $\det A = -1$.) There is a path σ from I to C in $SO(3)$. Since $A = BC$, we can multiply the paths ω and σ 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 $\epsilon - \delta$ definition for $f : (a,b) \to \mathbb{R}$.

2. Suppose we have $A \subset \mathbb{R}^n$ and have functions

$$A \xrightarrow{f} \mathbb{R}^m \qquad f(A) \xrightarrow{g} \mathbb{R}^p .$$

ave seen that f and g continuous implies that $g \circ f$ is con-
ous. Give examples to show that:

f continuous and $g \circ f$ continuous $\not\Rightarrow g$ continuous .

g continuous and $g \circ f$ continuous $\not\Rightarrow f$ continuous .

3. Show that for $A \subset \mathbb{R}^n$ and $f : A \to \mathbb{R}^m$ then f is continuous
for each open set U in $\mathbb{R}^m$, $f^{-1}(U)$ is an open set in A .

4. Show that if A,B are connected sets in $\mathbb{R}^n$ and $A \cap B \neq \phi$,
n $A \cup B$ is connected.

5. Let H be any connected subgroup of a matrix group G . Show
t

$$S = \bigcup_{x \in G} xHx^{-1}$$

connected.

6. Show that matrix multiplication is continuous (with one matrix
ed; i.e., $A \in M_n(k)$, $L_A : M_n(k) \to M_n(k)$ given by $L_A(B) = AB$
continuous) .

7. Show that an arbitrary union of open sets is an open set.

8. Let $A \subset \mathbb{R}^n$ and $x \in \mathbb{R}^n$. We say that x is a limit point
A (x lp A) if every

$$B(x,r) \cap A$$

an infinite set. Show that $C \subset \mathbb{R}^n$ is a closed set $\Leftrightarrow$
x lp $C \Rightarrow x \in C$) .

9. Let $D \subset \mathbb{R}^n$ be open and closed. Show that if D is not empty, then $D = \mathbb{R}^n$. (See the proof of Proposition 10.)

apter 7
aximal Tori

Cartesian products of groups

If G and H are groups, we make $G \times H$ into a group by defining

$$(g,h)(g',h') = (gg',hh') .$$

is 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 oup of $m \times m$ matrices, we can represent elements of $G \times H$ as $+ m) \times (n + m)$ matrices by

$$(g,h) = \begin{array}{c} n\{ \\ \\ \end{array}\left(\begin{array}{c|c} g & 0 \\ \hline \underbrace{0}_{n} & \underbrace{h}_{m} \end{array}\right)\begin{array}{c} \\ \}m \end{array} .$$

en matrix multiplication gives the operation described above on $\times H$. (Exercise.) Let us look at an important special case of is.

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

$$G \times H = \left\{\begin{pmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & \cos\phi & \sin\phi \\ 0 & 0 & -\sin\phi & \cos\phi \end{pmatrix}\right\} .$$

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

$$T = \begin{pmatrix} e^{i\theta_1} & & & \\ & e^{i\theta_2} & & \bigcirc \\ & \bigcirc & \ddots & \\ & & & e^{i\theta_k} \end{pmatrix}$$

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 γ,σ are one-parameter subgroups, then $\gamma\sigma$ is a one - parameter subgroup.

Proof:
$$\begin{aligned} (\gamma\sigma)(s+t) &= \gamma(s+t)\sigma(s+t) \\ &= \gamma(s)\gamma(t)\sigma(s)\sigma(t) \\ &= \gamma(s)\sigma(s)\gamma(t)\sigma(t) \\ &= (\gamma\sigma)(s)(\gamma\sigma)(t) . \end{aligned}$$

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 γ,σ being one-parameter subgroups. Then from Chapter IV we know that $\exp(\xi) = \gamma(1)$ and

$= \sigma(1)$. We have that $\gamma\sigma$ is a one-parameter subgroup and

$(0) = \gamma'(0) + \sigma'(0) = \xi + \eta$ (Chapter III), so $\exp(\xi+\eta)$

$(1) = \gamma(1)\sigma(1) = \exp \xi \exp \eta$.

ow we know that exp is 1-1 on some neighborhood V of 0

$TG)_e$. So for G abelian we know that $\ker(\exp) = L$ is a

ete subgroup of the vector group $(TG)_e$ (i.e. some neighborhood

contains no point of L except 0). Next we consider when

$(TG)_e \to G$ is surjective.

roposition 2: Let G be a connected matrix group and let H be

ubgroup of G containing an open neighborhood U of e . Then

.

roof: 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

is an open set containing x $(e \in U)$ and hence must contain some

t of W . Thus

$$xu = u_1 \ldots u_m \quad \text{for some } u,u_1,u_2,\ldots,u_m \in U .$$

then $x = u_1 \ldots u_m u^{-1} \in W$. In a connected space G only ϕ and

are both open and closed (Exercise, Chapter VI), $W \neq \phi$ so $W = G$

$H = G$.

Corollary: If G is a connected abelian matrix group, then

$\exp : (TG)_e \to G$ is a surjective homomorphism with a discrete kernel.

Proof: exp is a homomorphism so $\exp((TG)_e)$ is a subgroup of G and it contains some neighborhood U of e . Thus $\exp(TG)_e = G$.

In the exercises it is proved that this implies that $\exp(TG)_e = T^k \times R^{n-k}$ 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:

$$T = \left\{ \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\}$$

is a maximal torus in $SO(3)$.

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 $SO(3)$, i.e.,

$$T \subsetneq T' \subset SO(3) .$$

Since T' would be abelian, we would have: $\phi \in SO(3)$ such that

but ϕ commutes with every element of T . So it suffices to
that:

$\phi \in SO(3)$ commutes with each $t \in T$, then $\phi \in T$. Refer
o our proof that Center $SO(3) = \{I\}$ in Chapter V, and you will
at we have already proved this fact.

oposition 4:

$$T = \left\{ \begin{pmatrix} \cos\theta_1 & \sin\theta_1 & 0 & 0 \\ -\sin\theta_1 & \cos\theta_1 & 0 & 0 \\ 0 & 0 & \cos\theta_2 & \sin\theta_2 \\ 0 & 0 & -\sin\theta_2 & \cos\theta_2 \end{pmatrix} \right\}$$

aximal torus in $SO(4)$.

oof: This clearly is a 2-torus and is a subgroup of $SO(4)$.
ore, it suffices to prove that if $\phi \in SO(4)$ commutes with all
ts of T then $\phi \in T$.
t V be the 2-plane in R^4 spanned by e_1, e_2 and W be the
e in R^4 spanned by e_3, e_4 . We see that T consists of all

(rotation of V , rotation of W) .

aim: $\phi(e_1) \in V$.

$\alpha \in T$ such that α is the identity on V but is not the iden-
on W . Then

$$\phi(e_1) = ae_1 + be_2 + ce_3 + de_4$$
$$\alpha\phi(e_1) = ae_1 + be_2 + c'e_3 + d'e_4$$
$$\phi\alpha(e_1) = \phi(e_1) = ae_1 + be_2 + ce_3 + de_4 .$$

This shows $c = 0 = d$, so $\phi(e_1) \in V$. The same kind of proof shows $\phi(e_2) \in V$. Dually, $\phi(e_3)$ and $\phi(e_4)$ are in W .

So we know that ϕ is <u>orthogonal on</u> V and is <u>orthogonal on</u> W . A priori, it could be a <u>reflection</u> in each and we would still have $\phi \in SO(4)$. But ϕ commutes with all rotations on V and is thus a rotation on V . (See the proof of Proposition 3, Chapter V.) Similarly, ϕ is a rotation on W , so $\phi \in T$.

q.e.d.

From Propositions 1 and 2 the general result about maximal tori in $SO(n)$ should be clear. We have $n/2$ of the 2×2 rotation matrices for n even and have a 1 in the n,n position for n odd. The proof of the general case is an obvious extension of the above proofs.

<u>Proposition</u> 5:

$$T = \left\{ \begin{pmatrix} e^{i\theta_1} & & \bigcirc \\ & e^{i\theta_2} & \\ \bigcirc & & e^{i\theta_n} \end{pmatrix} \right\}$$

<u>is a maximal torus in</u> $U(n)$.

<u>Proof</u>: Let $\phi \in U(n)$ commute with each $\alpha \in T$. Consider any α of the form

$$\alpha = \begin{pmatrix} 1 & & \bigcirc \\ & e^{i\theta_2} & \\ \bigcirc & & e^{i\theta_n} \end{pmatrix} \in T .$$

Then $\alpha\phi(e_1) = \phi(\alpha(e_1)) = \phi(e_1)$. So α leaves $\phi(e_1)$ fixed, but such α's can move any vector which is not a multiple of e_1 . Hence

$$\phi(e_1) = \lambda_1 e_1 , \lambda_1 \in C .$$

r arguments give

$$\phi(e_j) = \lambda_j e_j \quad \text{for} \quad j = 1,\ldots,n .$$

$$\phi = \begin{pmatrix} \lambda_1 & & \bigcirc \\ & \lambda_2 & \\ \bigcirc & & \ddots \; \lambda_n \end{pmatrix}$$

since $\phi \in U(n)$, each λ_j is of unit length. Thus $\phi \in T$

T is maximal.

roposition 6: $T = \left\{ \begin{pmatrix} e^{i\theta_1} & & \bigcirc \\ & \ddots & \\ \bigcirc & & e^{i\theta_n} \end{pmatrix} \middle| \; \theta_1 + \ldots + \theta_n = 0 \right\}$

maximal torus in $SU(n)$.

roof: A matrix α of the form $\begin{pmatrix} e^{i\theta_1} & & \bigcirc \\ & \ddots & \\ \bigcirc & & e^{i\theta_n} \end{pmatrix}$ has

$\alpha = e^{i\theta_1} \ldots e^{i\theta_n} = e^{i(\theta_1+\ldots+\theta_n)}$ so that $\alpha \in SU(n) \Leftrightarrow \det \alpha$

$= \theta_1 + \ldots + \theta_n = 0$. So the T described here is just the

rsection 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,

$$\begin{pmatrix} e^{i\theta_1} & & \bigcirc \\ & \ddots & \\ \bigcirc & & e^{i\theta_n} \end{pmatrix} \leftrightarrow \begin{pmatrix} e^{i(\theta_1-\theta_n)} & & & \bigcirc \\ & e^{i(\theta_2-\theta_n)} & & \\ & & \ddots & \\ \bigcirc & & & e^{i(\theta_{n-1}-\theta_n)} \\ & & & & 1 \end{pmatrix}$$

$$\Sigma\theta_i = 0$$

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 $\alpha = \begin{pmatrix} 1 & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$ to use. But for $n = 2$

we give a direct simple proof. If

$$\phi = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SU(2) \quad \text{and} \quad \alpha = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \in T$$

then

$$\phi\alpha = \begin{pmatrix} ai & -bi \\ ci & -di \end{pmatrix} = \begin{pmatrix} ai & bi \\ -ci & -di \end{pmatrix} = \alpha\phi \; .$$

Thus $b = 0 = c$ and $\phi \in T$.

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 $Sp(n)$

h commutes with all $\begin{pmatrix} e^{i\theta_1} & & \bigcirc \\ & \ddots & \\ \bigcirc & & e^{i\theta_n} \end{pmatrix}$ must be diagonal with the

onal elements having length 1 . But now these elements are ernions. However (Exercise), any quaternion which commutes with ust be a complex number.

q.e.d.

Centers again

Now that we know maximal tori in our matrix groups, we are able alculate the centers.

Proposition 8: Center $(Sp(n)) = \{I,-I\}$.

Proof: We have seen that if any element commutes with all ele- s of the maximal torus we have described, then it must lie in maximal torus. Hence, in every case

$$\text{Center} \subset T .$$

If $A \in$ Center $Sp(n)$, then $A = \begin{pmatrix} e^{i\theta_1} & & \bigcirc \\ & \ddots & \\ \bigcirc & & e^{i\theta_n} \end{pmatrix}$. Since A must ute with the matrix jI , it follows that the diagonal elements be real (and since they are of unit length) they are ± 1 . It an exercise to show that they all have the same sign. Now I and 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 $B \in$ Center $U(n)$ we get that B is diagonal with diagonal elements complex numbers of unit length. Let

$$B = \begin{pmatrix} \alpha_1 & & & \bigcirc \\ & \alpha_2 & & \\ & & \ddots & \\ \bigcirc & & & \alpha_n \end{pmatrix}$$ with each $|\alpha_\ell| = 1$. Let

$$A = \begin{pmatrix} 0 & 1 & 0\ldots0 \\ 1 & 0 & 0\ldots0 \\ 0 & 0 & 1 \ddots 0 \\ & \vdots & 0 \ddots 1 \end{pmatrix} .$$

Then $AB = BA$ shows $\alpha_1 = \alpha_2$, etc., so all α 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

$$\det(e^{i\theta}I) = e^{in\theta} ,$$

and since this must be 1 , $e^{i\theta}$ must be an n^{th} root of unity.

q.e.d.

So Center $SU(n) = SU(n) \cap$ Center $U(n)$.

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

matrix.

ore, if $A \in SO(k)$ is in the center, it must be in our

aximal torus. So suppose $A \in$ Center $SO(k)$ is of the form

$$A = \begin{pmatrix} \cos\theta_1 & \sin\theta_1 & 0 \ldots \\ -\sin\theta_1 & \cos\theta_1 & 0 \ldots \\ 0 & 0 & * \ldots \end{pmatrix}$$

$$P = \begin{pmatrix} 0 & 0 & 1 & & & \\ 0 & 1 & 0 & & \bigcirc & \\ -1 & 0 & 0 & & & \\ & & & 1 & & 0 \\ & \bigcirc & & & \ddots & \\ & & & 0 & & 1 \end{pmatrix} \in SO(k) \ .$$

has zero in the 1,2 position, whereas AP has $\sin\theta_1$

2 position. Thus $\sin\theta_1 = 0$. Similar arguments show all

al terms are zero.

lows also (since each $\sin\theta_i = 0$) that each diagonal term

is 1 or -1. So each 2×2 block is $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ or

. Arguments like we used for $U(n)$ show that all diagonal

be equal. So we finally conclude that

$$\text{Center } SO(2n+1) = \{I\}$$

$$\text{Center } SO(2n) = \{I,-I\} \ .$$

le, $\begin{pmatrix} -1 & & & \bigcirc \\ & -1 & & \\ & & -1 & \\ \bigcirc & & & -1 \end{pmatrix}$ is in the center of $SO(4)$, but $\begin{pmatrix} -1 & & \bigcirc \\ & -1 & \\ \bigcirc & & -1 \end{pmatrix}$

en in $SO(3)$.

We now tabulate the information we have generated about our groups.

<u>Dimensions</u>, <u>Centers</u>, <u>Maximal tori</u>

Group	Dimension	Center	Standard Maximal torus
$U(n)$	n^2	$\{e^{i\theta}I\} \cong S^1$	$\left\{\begin{pmatrix} e^{i\theta_1} & & \bigcirc \\ & \ddots & \\ \bigcirc & & e^{i\theta_n} \end{pmatrix}\right\}$
$SU(n)$	n^2-1	$\{wI \mid w^n=1\} \cong \frac{\mathbb{Z}}{n}$	$\left\{\begin{pmatrix} e^{i\theta_1} & & \bigcirc \\ & \ddots & \\ \bigcirc & & e^{i\theta_n} \end{pmatrix} \middle\vert\, \Sigma\theta_i = 0\right\}$
$SO(2n+1)$	$\frac{(2n+1)(2n)}{2} = 2n^2+n$	$\{I\}$	$\left\{\begin{pmatrix} \text{rot } \theta_1 & & \bigcirc \\ & \ddots & \\ \bigcirc & & \text{rot } \theta_{n_1} \end{pmatrix}\right\}$
$SO(2n)$	$\frac{2n(2n-1)}{2} = 2n^2-n$	$\{I,-I\}$	$\left\{\begin{pmatrix} \text{rot } \theta_1 & & \bigcirc \\ & \ddots & \\ \bigcirc & & \text{rot } \theta_n \end{pmatrix}\right\}$
$Sp(n)$	$2n^2+n$	$\{I,-I\}$	$\left\{\begin{pmatrix} e^{i\theta_1} & & \bigcirc \\ & \ddots & \\ \bigcirc & & e^{i\theta_n} \end{pmatrix}\right\}$

Note that we have nothing to distinguish $SO(2n+1)$ and $\frac{Sp(n)}{\text{Center}}$. 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
oup. 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

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \text{ and } \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}$$

he matrix for a rotation through angle $\theta + \phi$.

4. Let T be a maximal torus in a matrix group G and let
G. Prove that xTx^{-1} is also a maximal torus in G.

5. Prove that if q is a quaternion such that $qi = iq$, then
s a complex number.

6. Show that $\dfrac{U(n)}{\text{center}} \cong \dfrac{SU(n)}{\text{center}}$.

7. A lattice subgroup K of $\mathbb{R}^n$ consists of all integar linear
binations of some set of linearly independent vectors. More
licitely, let

$$v_1, v_2, \ldots, v_k$$

linearly independent vectors in $\mathbb{R}^n$. For any integers $a_1, \ldots, a_k$
form

$$a_1v_1 + \ldots + a_kv_k .$$

n $K = \{a_1v_1 + \ldots + a_kv_k\} \mid a_i \in \mathbb{Z}\}$. It is routine to verify that
is a subgroup of $\mathbb{R}^n$.

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 $\mathbb{R}v_1$. Show that $\mathbb{R}v_1$ 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.)

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 $\mathbb{R}^{n-k}$.