Chapter 8

Linear Groups

In these days the angel of topology and the devil of abstract algebra fight for the soul of every individual discipline of mathematics.

Hermann Weyl

1. THE CLASSICAL LINEAR GROUPS

Subgroups of the general linear group GL_n are called *linear groups*. In this chapter we will study the most important ones: the orthogonal, unitary, and symplectic groups. They are called the *classical groups*.

The classical groups arise as stabilizers for some natural operations of GL_n on the space of $n \times n$ matrices. The first of these operations is that which describes change of basis in a bilinear form. The rule

$$P, A \rightsquigarrow (P^t)^{-1} A P^{-1} \tag{1.1}$$

is an operation of GL_n on the set of all $n \times n$ matrices. This is true for any field of scalars, but we will be interested in the real and complex cases. As we have seen in Chapter 7 (1.15), the orbit of a matrix A under this operation is the set of matrices A' which represent the form $X^t A Y$, but with respect to different bases. It is customary to call matrices in the same orbit *congruent*. We can set $Q = (P^t)^{-1}$ to obtain the equivalent definition

(1.2) *A and A' are congruent if $A' = QAQ^t$ for some $Q \in GL_m(F)$.*

Sylvester's Law [Chapter 7 (2.11)] describes the different orbits or congruence classes of real symmetric matrices. Every congruence class of real symmetric matrices contains exactly one matrix of the form Chapter 7 (2.10). The *orthogonal group*, which we have defined before, is the stabilizer of the identity matrix for this operation. As before, we will denote the real orthogonal group by the symbol O_n:

$$O_n = \{P \in GL_n(\mathbb{R}) \mid P^t P = I\}. \tag{1.3}$$

The complex orthogonal group is defined analogously:

$$O_n(\mathbb{C}) = \{P \in GL_n(\mathbb{C}) \mid P^tP = I\}.$$

The stabilizer of the *Lorentz form* [Chapter 7 (2.16)], defined by the matrix

$$I_{3,1} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{bmatrix},$$

is called the *Lorentz group*. It is denoted by $O_{3,1}(\mathbb{R})$ or $O_{3,1}$:

$$O_{3,1} = \{P \in GL_n(\mathbb{R}) \mid P^tI_{3,1}P = I_{3,1}\}. \tag{1.4}$$

The linear operators represented by these matrices are often called *Lorentz transformations*. The subscript (3, 1) indicates the signature of the matrix, the number of $+1$'*s* and -1'*s*. In this way an analogous group $O_{p,q}$ can be defined for any signature (p, q).

The operation (1.1) also describes change of basis in forms which are not symmetric. Thus Theorem (8.6) of Chapter 7 tells us this:

(1.5) **Corollary.** There is exactly one congruence class of real nonsingular skew-symmetric $m \times m$ matrices, if m is even. □

The standard skew-symmetric form is defined by the $2n \times 2n$ matrix J (Chapter 7 (8.5)), and its stabilizer is called the *symplectic group*

$$SP_{2n}(\mathbb{R}) = \{P \in GL_{2n}(\mathbb{R}) \mid P^tJP = J\}. \tag{1.6}$$

Again, the complex symplectic group $SP_{2n}(\mathbb{C})$ is defined analogously.

Finally, the *unitary group* is defined in terms of the operation

$$P, A \rightsquigarrow (P^*)^{-1}AP^{-1}. \tag{1.7}$$

This definition makes sense only when the field of scalars is the complex field. Exactly as with bilinear forms, the orbit of a matrix A consists of the matrices which define the form $\langle X, Y\rangle = X^*AY$ with respect to different bases (see [Chapter 7 (4.12)]). The unitary group is the stabilizer of the identity matrix for this action:

$$U_n = \{P \mid P^*P = I\}. \tag{1.8}$$

Thus U_n is the group of matrices representing changes of basis which leaves the hermitian dot product [Chapter 7 (4.2)] X^*Y invariant.

The word *special* is added to indicate the subgroup of matrices with determinant 1. This gives us some more groups:

Special linear group	$SL_n(\mathbb{R})$: $n \times n$ matrices P with determinant 1;
Special orthogonal group	$SO_n(\mathbb{R})$: the intersection $SL_n(\mathbb{R}) \cap O_n(\mathbb{R})$;
Special unitary group	SU_n: the intersection $SL_n(\mathbb{C}) \cap U_n$.

Though this is not obvious from the definition, symplectic matrices have determinant 1, so the two uses of the letter S do not cause conflict.

2. THE SPECIAL UNITARY GROUP SU_2

The main object of this chapter is to describe the geometric properties of the classical linear groups, by considering them as subsets of the spaces $\mathbb{R}^{n\times n}$ or $\mathbb{C}^{n\times n}$ of all matrices. We know the geometry of a few groups already. For example, $GL_1(\mathbb{C}) = \mathbb{C}^\times$ is the "punctured plane" $\mathbb{C} - \{0\}$. Also, if p is a 1×1 matrix, then $p^* = \overline{p}$. Thus

$$U_1 = \{p \in \mathbb{C}^\times \mid \overline{p}p = 1\}. \tag{2.1}$$

This is the set of complex numbers of absolute value 1—the unit circle in the complex plane. We can identify it with the unit circle in $\mathbb{R}^2$,

$$x_1^2 + x_2^2 = 1,$$

by sending $x_1 + x_2 i \rightsquigarrow (x_1, x_2)$. The group SO_2 of rotations of the plane is isomorphic to U_1. It is also a circle, embedded into $\mathbb{R}^{2\times 2}$ by the map

$$(x_1, x_2) \rightsquigarrow \begin{bmatrix} x_1 & -x_2 \\ x_2 & x_1 \end{bmatrix}. \tag{2.2}$$

We will describe some more of the groups in the following sections.

The *dimension* of a linear group G is, roughly speaking, the number of degrees of freedom of a matrix in G. The group SO_2, for example, has dimension 1. A matrix in SO_2 represents rotation by an angle θ, and this angle is the single parameter needed to determine the rotation. We will discuss dimension more carefully in Section 7, but we want to describe some of the low-dimensional groups explicitly first. The smallest dimension in which really interesting groups appear is 3, and three of these—SU_2, SO_3, and $SL_2(\mathbb{R})$—are very important. We will study the special unitary group SU_2 in this section.

Let $P = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be an element of SU_2, with $a, b, c, d \in \mathbb{C}$. The equations defining SU_2 are $P^*P = I$ and $\det P = 1$. By Cramer's Rule,

$$P^{-1} = (\det P)^{-1}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Since $P^{-1} = P^*$ for a matrix in SU_2, we find $\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} \overline{a} & \overline{c} \\ \overline{b} & \overline{d} \end{bmatrix}$, or

$$\overline{a} = d, \quad \text{and} \quad \overline{b} = -c. \tag{2.3}$$

Thus

$$P = \begin{bmatrix} a & b \\ -\overline{b} & \overline{a} \end{bmatrix}. \tag{2.4}$$

The condition $\det P = 1$ has become lost in the computation and must be put back:

$$\bar{a}a + \bar{b}b = 1. \tag{2.5}$$

Equations (2.3) and (2.5) provide a complete list of conditions describing the entries of a matrix in SU_2. The matrix P is described by the vector $(a, b) \in \mathbb{C}^2$ of length 1, and any such vector gives us a matrix $P \in SU_2$ by the rule (2.4).

If we write out a, b in terms of their real and imaginary parts, equation (2.5) gives us a bijective correspondence between SU_2 and points of $\mathbb{R}^4$ lying on the locus

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1. \tag{2.6}$$

This equation is equivalent to (2.5) if we set $a = x_1 + x_2 i$ and $b = x_3 + x_4 i$.

The locus (2.6) is called the *unit 3-sphere* in $\mathbb{R}^4$, in analogy with the unit sphere in $\mathbb{R}^3$. The number 3 refers to its dimension, the number of degrees of freedom of a point on the sphere. Thus the unit sphere

$$x_1^2 + x_2^2 + x_3^2 = 1$$

in $\mathbb{R}^3$, being a surface, is called a *2-sphere*. The unit circle in $\mathbb{R}^2$, a curve, is called a *1-sphere*. We will sometimes denote a sphere of dimension d by S^d.

A bijective map $f\colon S \longrightarrow S'$ between subsets of Euclidean spaces is called a *homeomorphism* if f and f^{-1} are continuous maps (Appendix, Section 3). The correspondence between SU_2, considered as a subset of $\mathbb{C}^{2\times 2}$, and the sphere (2.6) is obviously continuous, as is its inverse. Therefore these two spaces are homeomorphic.

(2.7) *SU_2 is homeomorphic to the unit 3-sphere in $\mathbb{R}^4$.*

It is convenient to identify SU_2 with the 3-sphere. We can do this if we represent the matrix (2.4) by its top row, the vector $(a, b) \in \mathbb{C}^2$, or by the vector $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$. These representations can be thought of as different notations for the same element P of the group, and we will pass informally from one representation to the other. For geometric visualization, the representations $P = (a, b)$ and $P = (x_1, x_2, x_3, x_4)$, being in lower-dimensional spaces, are more convenient.

The fact that the 3-sphere has a group structure is remarkable, because there is no way to make the 2-sphere into a group with a continuous law of composition. In fact, a famous theorem of topology asserts that the only spheres with continuous group laws are the 1-sphere, which is realized as the rotation group SO_2, and the 3-sphere SU_2.

We will now describe the algebraic structures on SU_2 analogous to the curves of constant latitude and longitude on the 2-sphere. The matrices $I, -I$ will play the roles of the north and south poles. In our vector notation, they are the points $(\pm 1, 0, 0, 0)$ of the sphere.

If the poles of the 2-sphere $x_1^2 + x_2^2 + x_3^2 = 1$ are placed at the points $(\pm 1, 0, 0)$, then the latitudes are the circles $x_1 = c$, $-1 < c < 1$. The analogues on the 3-sphere SU_2 of these latitudes are the surfaces on which the x_1-coordinate is constant. They are two-dimensional spheres, embedded into $\mathbb{R}^4$ by

$$x_1 = c \quad \text{and} \quad x_2^2 + x_3^2 + x_4^2 = (1 - c^2), \quad -1 < c < 1. \tag{2.8}$$

These sets can be described algebraically as *conjugacy classes* in SU_2.

(2.9) **Proposition.** Except for two special classes, the conjugacy classes in SU_2 are the *latitudes*, the sets defined by the equations (2.8). For a given c in the interval $(-1, 1)$, this set consists of all matrices $P \in SU_2$ such that trace $P = 2c$. The remaining conjugacy classes are $\{I\}$ and $\{-I\}$, each consisting of one element. These two classes make up the center $Z = \{\pm I\}$ of the group SU_2.

Proof. The characteristic polynomial of the matrix P (2.4) is

$$t^2 - (a + \bar{a})t + 1 = t^2 - 2x_1 t + 1. \tag{2.10}$$

This polynomial has a pair $\lambda, \bar{\lambda}$ of complex conjugate roots on the unit circle, and the roots, the eigenvalues of P, depend only on trace $P = 2x_1$. Furthermore, two matrices with different traces have different eigenvalues. The proposition will follow if we show that the conjugacy class of P contains every matrix in SU_2 with the same eigenvalues. The cases $x_1 = 1, -1$ correspond to the two special conjugacy classes $\{I\}, \{-I\}$, so the proof is completed by the next lemma.

(2.11) **Lemma.** Let P be an element of SU_2, with eigenvalues $\lambda, \bar{\lambda}$. Then P is conjugate in SU_2 to the matrix

$$\begin{bmatrix} \lambda & \\ & \bar{\lambda} \end{bmatrix}.$$

Proof. By the Spectral Theorem for normal operators [Chapter 7 (7.3)], there is a unitary matrix Q so that QPQ^* is diagonal. We only have to show that Q can be chosen so as to have determinant 1. Say that $\det Q = \delta$. Since $Q^*Q = I$, $(\det Q^*)(\det Q) = \bar{\delta}\delta = 1$; hence δ has absolute value 1. Let ϵ be a square root of δ. Then $\bar{\epsilon}\epsilon = 1$ too. The matrix $Q_1 = \bar{\epsilon}Q$ is in SU_2, and $P_1 = Q_1 P Q_1^*$ is also diagonal. The diagonal entries of P_1 are the eigenvalues $\lambda, \bar{\lambda}$. The eigenvalues can be interchanged, if desired, by conjugating by the matrix

$$Q_2 = \begin{bmatrix} & 1 \\ -1 & \end{bmatrix}, \tag{2.12}$$

which is also an element of SU_2. □

Next we will introduce the longitudes of SU_2. The longitudes on the 2-sphere $x_1^2 + x_2^2 + x_3^2 = 1$ can be described as intersections of the sphere with planes containing the two poles $(\pm 1, 0, 0)$. When we add a fourth variable x_4 to get the equation of the 3-sphere, a natural way to extend this definition is to form the intersection with a two-dimensional subspace of $\mathbb{R}^4$ containing the two poles $\pm I$. This is a circle in SU_2, and we will think of these circles as the longitudes. Thus while the latitudes on SU_2 are 2-spheres, the *longitudes* are 1-spheres, the "great circles" through the poles.

Note that every point $P = (x_1, x_2, x_3, x_4)$ of SU_2 except for the poles is contained in exactly one longitude. This is because if P is not a pole, then P and I will

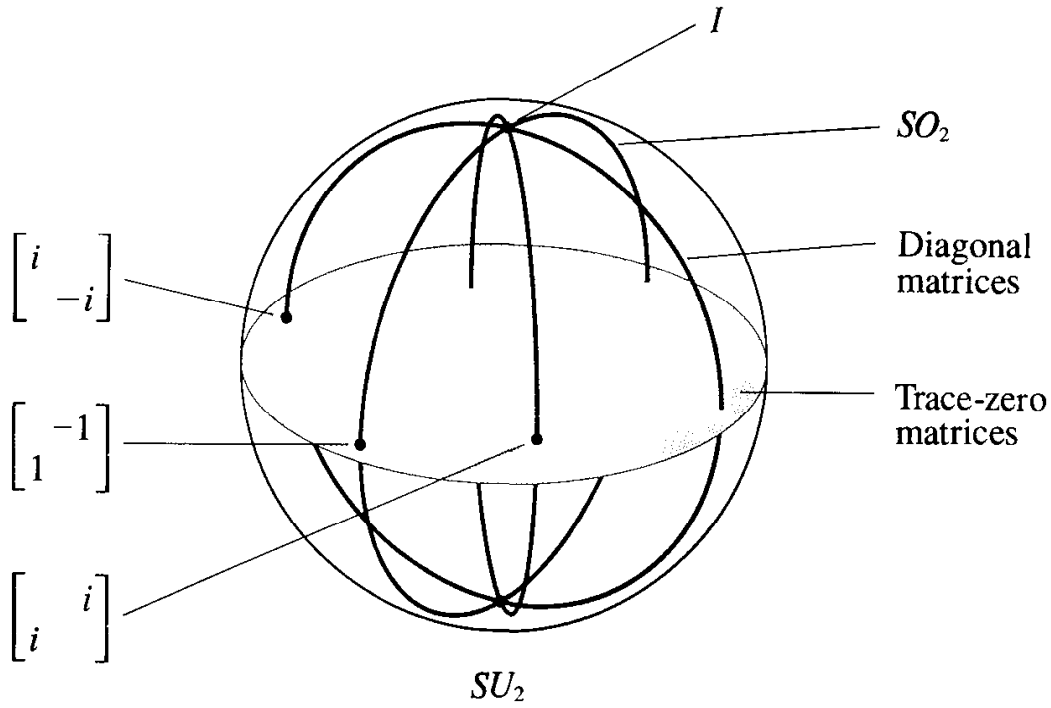
be linearly independent and thus will span a subspace V of $\mathbb{R}^4$ of dimension 2. The intersection $SU_2 \cap V$ is the unique longitude containing P.

The intersection of SU_2 with the plane W defined by $x_3 = x_4 = 0$ is a particularly nice longitude. In matrix notation, this great circle consists of the diagonal matrices in SU_2, which form a subgroup T:

$$T = \left\{ \begin{bmatrix} \lambda & \\ & \bar{\lambda} \end{bmatrix} \middle| \lambda\bar{\lambda} = 1 \right\}. \tag{2.13}$$

The other longitudes are described in the following proposition.

(2.14) **Proposition.** The longitudes of SU_2 are the conjugate subgroups QTQ^* of the subgroup T.



(2.15) **Figure.** Some latitudes and longitudes in SU_2.

In Figure (2.15) the 3-sphere SU_2 is projected from $\mathbb{R}^4$ onto the unit disc in the plane. The conjugacy class shown is the "equatorial" latitude in $\mathbb{R}^4$, which is defined by the equation $x_1 = 0$. Just as the orthogonal projection of a circle from $\mathbb{R}^3$ to $\mathbb{R}^2$ is an ellipse, the projection of this 2-sphere from $\mathbb{R}^4$ to $\mathbb{R}^3$ is an ellipsoid, and the further projection of this ellipsoid to the plane is the elliptical disc shown.

Proof of Proposition (2.14). The point here is to show that any conjugate subgroup QTQ^* is a longitude. Lemma (2.11) tells us that every element $P \in SU_2$ lies in one of these conjugate subgroups (though the roles of Q and Q^* have been reversed). Since every $P \neq \pm I$ is contained in exactly one longitude, it will follow that every longitude is one of the subgroups QTQ^*.

So let us show that a conjugate subgroup QTQ^* is a longitude. The reason this is true is that conjugation by a fixed element Q is a linear operator which sends the

subspace W to another subspace. We will compute the conjugate explicitly to make this clear. Say that Q is the matrix (2.4). Let $w = (w_1, w_2, 0, 0)$ denote a variable element of W, and set $z = w_1 + w_2 i$. Then

$$\begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix} \begin{bmatrix} z & \\ & \bar{z} \end{bmatrix} \begin{bmatrix} a & -b \\ \bar{b} & a \end{bmatrix} = \begin{bmatrix} a\bar{a}z + b\bar{b}\bar{z} & ab(\bar{z}-z) \\ * & * \end{bmatrix}.$$

Computing these entries, we find that w is sent to the vector $u = (u_1, u_2, u_3, u_4)$, where

$$u_1 = w_1, \quad u_2 = (x_1^2 + x_2^2 - x_3^2 - x_4^2)w_2,$$
$$u_3 = 2(x_1x_4 + x_2x_3)w_2, \quad u_4 = 2(x_2x_4 - x_1x_3)w_2.$$

The coordinates u_i are real linear combinations of (w_1, w_2). This shows that the map $w \rightsquigarrow u$ is a real linear transformation. So its image V is a subspace of $\mathbb{R}^4$. The conjugate group QTQ^* is $SU_2 \cap V$. Since QTQ^* contains the poles $\pm I$, so does V, and this shows that QTQ^* is a longitude. □

We will describe another geometric configuration briefly: As we have seen, the subgroup T of diagonal matrices is a great circle in the 3-sphere SU_2. The left cosets of this subgroup, the sets of the form QT for $Q \in SU_2$, are also great circles, and they partition the group SU_2. Thus the 3-sphere is partitioned into great circles. This very interesting configuration is called the *Hopf fibration*.

3. THE ORTHOGONAL REPRESENTATION OF SU_2

We saw in the last section that the conjugacy classes in the special unitary group SU_2 are two-dimensional spheres. Since conjugacy classes are orbits for the operation of conjugation, SU_2 operates on these spheres. In this section we will show that conjugation by an element $P \in SU_2$ acts on each of the spheres as a *rotation*, and that the map sending P to the matrix of this rotation defines a surjective homomorphism

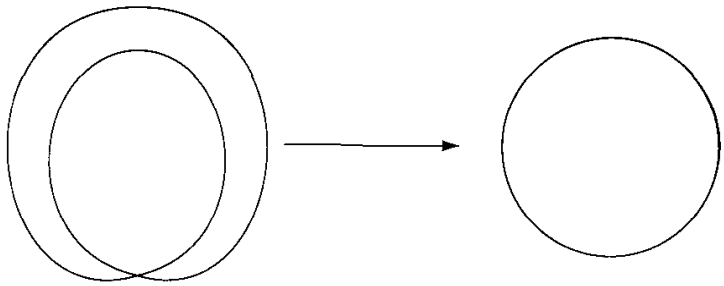
$$\varphi: SU_2 \longrightarrow SO_3, \tag{3.1}$$

whose kernel is the center $Z = \{\pm I\}$ of SU_2. This homomorphism is called the *orthogonal representation* of SU_2. It represents a complex 2×2 matrix P in SU_2 by a real 3×3 rotation matrix $\varphi(P)$.

The safest way to show that P operates by rotating a conjugacy class may be to write the matrix representing the rotation down explicitly. This is done in (3.12). However, the formula for $\varphi(P)$ is complicated and not particularly enlightening. It is better to describe φ indirectly, as we will do presently. Let us discuss the geometry of the map first.

Since the kernel of φ is $\{\pm I\}$, its cosets are the sets $\{\pm P\}$. They form the fibres of the homomorphism. Thus every element of SO_3 corresponds to a pair of unitary matrices which differ by sign. Because of this, the group SU_2 is called a *double covering* of the group SO_3.

The map μ: $SO_2 \longrightarrow SO_2$ of the 1-sphere to itself defined by $\rho_\theta \rightsquigarrow \rho_{2\theta}$ is another example of a double covering. Its kernel also consists of two elements, the identity and the rotation by π. Every fibre of μ contains two rotations ρ_θ and $\rho_{\pi+\theta}$.



(3.2) **Figure.** A double covering of the 1-sphere.

The orthogonal representation can be used to identify the topological structure of the rotation group. In vector notation, if $P = (x_1,\dots,x_4)$, then $-P = (-x_1,\dots,-x_4)$, and the point $-P$ is called the *antipode* of P. So since points of the rotation group correspond to cosets $\{\pm P\}$, the group SO_3 can be obtained by identifying antipodal points on the 3-sphere SU_2. The space obtained in this way is called the *real projective 3-space:*

(3.3) *SO_3 is homeomorphic to the real projective 3-space.*

The number 3 refers again to the dimension of the space. Points of the real projective 3-space are also in bijective correspondence with lines through the origin (or one-dimensional subspaces) of $\mathbb{R}^4$. Every line through the origin meets the unit sphere in a pair of antipodal points.

As we noted in Section 8 of Chapter 4, every element of SO_3 except the identity can be described in terms of a pair (v,θ), where v is a unit vector in the axis of rotation and where θ is the angle of rotation. However, the two pairs (v,θ) and $(-v,-\theta)$ represent the same rotation. The choice of one of these pairs is referred to by physicists as the choice of a *spin*. It is not possible to make a choice of spin which varies continuously over the whole group. Instead, the two possible choices define a double covering of $SO_3 - \{I\}$. We may realize the set of all pairs (v,θ) as the product space $S \times \Theta$, where S is the 2-sphere of unit vectors in $\mathbb{R}^3$, and where Θ is the set of nonzero angles $0 < \theta < 2\pi$. This product space maps to SO_3:

$$\psi\colon S \times \Theta \longrightarrow SO_3 - \{I\}, \tag{3.4}$$

by sending (v,θ) to the rotation about v through the angle θ. The map ψ is a double covering of $SO_3 - \{I\}$ because every nontrivial rotation is associated to two pairs $(v,\theta),(-v,-\theta)$.

We now have two double coverings of $SO_3 - \{I\}$, namely $S \times \Theta$ and also $SU_2 - \{\pm I\}$, and it is plausible that they are equivalent. This is true:

(3.5) **Proposition.** There is a homeomorphism h: $(SU_2 - \{\pm I\}) \longrightarrow S \times \Theta$ which is compatible with the maps SO_3, i.e., such that $\psi \circ h = \varphi$.

This map h is not a group homomorphism. In fact, neither its domain nor its range is a group.

Proposition (3.5) is not very difficult to prove, but the proof is slightly elusive because there are two such homeomorphisms. They differ by a switch of the spin. On the other hand, the fact that this homeomorphism exists follows from a general theorem of topology, because the space $SU_2 - \{\pm I\}$ is *simply connected.* (A simply connected space is one which is path connected and such that every loop in the space can be contracted continuously to a point.) It is better to leave this proof to the topologists. □

Therefore every element of SU_2 except $\pm I$ can be described as a rotation of $\mathbb{R}^3$ together with a choice of spin. Because of this, SU_2 is often called the *Spin group.*

We now proceed to compute the homomorphism φ, and to begin, we must select a conjugacy class. It is convenient to choose the one consisting of the trace-zero matrices in SU_2, which is the one defined by $x_1 = 0$ and which is illustrated in Figure (2.15). The group operates in the same way on the other classes. Let us call the conjugacy class of trace-zero matrices C. An element A of C will be a matrix of the form

$$A = \begin{bmatrix} y_2 i & y_3 + y_4 i \\ -y_3 + y_4 i & -y_2 i \end{bmatrix}, \tag{3.6}$$

where

$$y_2^2 + y_3^2 + y_4^2 = 1. \tag{3.7}$$

Notice that this matrix is *skew-hermitian,* that is, it has the property

$$A^* = -A. \tag{3.8}$$

(We haven't run across skew-hermitian matrices before, but they aren't very different from hermitian matrices. In fact, A is a skew-hermitian matrix if and only if $H = iA$ is hermitian.) The 2×2 skew-hermitian matrices with trace zero form a real vector space V of dimension 3, with basis

$$\mathbf{B} = \left[\begin{bmatrix} i & \\ & -i \end{bmatrix}, \begin{bmatrix} & 1 \\ -1 & \end{bmatrix}, \begin{bmatrix} & i \\ i & \end{bmatrix} \right]. \tag{3.9}$$

In the notation of (3.6), $A = \mathbf{B}Y$, where $Y = (y_2, y_3, y_4)^t$. So the basis $\mathbf{B}$ corresponds to the standard basis (e_2, e_3, e_4) in the space $\mathbb{R}^3$, and (3.7) tells us that our conjugacy class is represented as the unit sphere in this space.

Note that SU_2 operates by conjugation on the whole space V of trace-zero, skew-hermitian matrices, not only on its unit sphere: If $A \in V$, $P \in SU_2$, and if $B = PAP^* = PAP^{-1}$, then trace $B = 0$, and $B^* = (PAP^*)^* = PA^*P^* = (P(-A)P^* = -B$. Also, conjugation by a fixed matrix P gives a linear operator on V, because $P(A + A')P^* = PAP^* + PA'P^*$, and if r is a real number, then $P(rA)P^* = rPAP^*$. The matrix of this linear operator is defined to be $\varphi(P)$. To determine the matrix ex-

plicitly, we conjugate the basis (3.7) by P and rewrite the result in terms of the basis. For example,

$$\begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix}\begin{bmatrix} i & \\ & -i \end{bmatrix}\begin{bmatrix} \bar{a} & -b \\ \bar{b} & a \end{bmatrix} = i\begin{bmatrix} a\bar{a}-b\bar{b} & -2ab \\ -2\bar{a}\bar{b} & b\bar{b}-a\bar{a} \end{bmatrix}. \tag{3.10}$$

The coordinates of this matrix are $y_2 = a\bar{a} - b\bar{b}$, $y_3 = i(-ab + \bar{a}\bar{b})$, and $y_4 = -(ab + \bar{a}\bar{b})$. They form the first column of the matrix $\varphi(P)$. Similar computation for the other columns yields

$$\begin{bmatrix} (a\bar{a}-b\bar{b}) & i(\bar{a}b-a\bar{b}) & (\bar{a}b+a\bar{b}) \\ i(\bar{a}\bar{b}-ab) & \frac{1}{2}(a^2+\bar{a}^2+b^2+\bar{b}^2) & \frac{i}{2}(a^2-\bar{a}^2-b^2+\bar{b}^2) \\ -(\bar{a}\bar{b}+ab) & \frac{i}{2}(\bar{a}^2-a^2+\bar{b}^2-b^2) & \frac{1}{2}(a^2+\bar{a}^2-b^2-\bar{b}^2) \end{bmatrix}. \tag{3.11}$$

We will not make use of the above computation. Even without it, we know that $\varphi(P)$ is a real 3×3 matrix because it is the matrix of a linear operator on a real vector space V of dimension 3.

(3.12) **Lemma.** The map $P \rightsquigarrow \varphi(P)$ defines a homomorphism $SU_2 \longrightarrow GL_3(\mathbb{R})$.

Proof. It follows from the associative law [Chapter 5 (5.1)] for the operation of conjugation that φ is compatible with multiplication: The operation of a product PQ on a matrix A is $(PQ)A(PQ)^* = P(QAQ^*)P^*$. This is the composition of the operations of conjugation by P and by Q. Since the matrix of the composition of linear operators is the product matrix, $\varphi(PQ) = \varphi(P)\varphi(Q)$. Being compatible with multiplication, $\varphi(P^{-1})\varphi(P) = \varphi(I_2) = I_3$. Therefore $\varphi(P)$ is invertible for every P, and so φ is a homomorphism from SU_2 to $GL_3(\mathbb{R})$, as asserted. □

(3.13) **Lemma.** For any P, $\varphi(P) \in SO_3$. Hence $P \rightsquigarrow \varphi(P)$ defines a homomorphism $SU_2 \longrightarrow SO_3$.

Proof. One could prove this lemma using Formula (3.11). To prove it conceptually, we note that dot product on $\mathbb{R}^3$ carries over to a bilinear form on V with a nice expression in terms of the matrices. Using the notation of (3.6), we define $\langle A, A' \rangle = y_1y_1' + y_2y_2' + y_3y_3'$. Then

$$\langle A, A' \rangle = -\tfrac{1}{2}\operatorname{trace}(AA'). \tag{3.14}$$

This is proved by computation:

$$AA' = \begin{bmatrix} -(y_2y_2'+y_3y_3'+y_4y_4') & * \\ * & -(y_2y_2'+y_3y_3'+y_3y_3') \end{bmatrix},$$

and so $\operatorname{tr} AA' = -2\langle A, A' \rangle$.

This expression for dot product shows that it is preserved by conjugation by an element $P \in SU_2$:

$$\langle PAP^*, PA'P^* \rangle = -\tfrac{1}{2}\operatorname{trace}(PAP^*PA'P^*) = -\tfrac{1}{2}\operatorname{trace}(AA') = \langle A, A' \rangle.$$

Or, in terms of the coordinate vectors, $(\varphi(P)Y \cdot \varphi(P)Y') = (Y \cdot Y')$. It follows that $\varphi(P)$ lies in the orthogonal group $O_3 = O_3(\mathbb{R})$ [Chapter 4 (5.13)].

To complete the proof, let us verify that $\varphi(P)$ has determinant 1 for every $P \in SU_2$: Being a sphere, SU_2 is path connected. So only one of the two possible values ± 1 can be taken on by the continuous function det $\varphi(P)$. Since $\varphi(I_2) = I_3$ and det $I_3 = 1$, the value is always $+1$, and $\varphi(P) \in SO_3$, as required. □

(3.15) **Lemma.** $\ker \varphi = \{\pm I\}$.

Proof. The kernel of φ consists of the matrices $P \in SU_2$ which act trivially on V, meaning that $PAP^* = A$ for all skew-hermitian matrices of trace zero. Suppose that P has the property $PAP^* = A$, or $PA = AP$, for all $P \in V$. We test it on the basis (3.7). The test leads to $b = 0$, $a = \bar{a}$, which gives the two possibilities $P = \pm I$, and they are in the kernel. So $\ker \varphi = \{\pm I\}$, as claimed. □

(3.16) **Lemma.** The image of the map φ is SO_3.

Proof. We first compute $\varphi(P)$ explicitly on the subgroup T of diagonal matrices in SU_2. Let $z = y_3 + y_4 i$. Then

$$PAP^* = \begin{bmatrix} a & \\ & \bar{a} \end{bmatrix} \begin{bmatrix} y_2 i & z \\ -\bar{z} & -y_2 i \end{bmatrix} \begin{bmatrix} \bar{a} & \\ & a \end{bmatrix} = \begin{bmatrix} y_2 i & a^2 z \\ -\bar{a}^2 \bar{z} & -y_2 i \end{bmatrix}. \tag{3.17}$$

So $\varphi(P)$ fixes the first coordinate y_2 and it multiplies z by a^2. Since $|a| = 1$, we may write $a = e^{i\theta}$. Multiplication by $a^2 = e^{2i\theta}$ defines a rotation by 2θ of the complex z-plane. Therefore

$$\varphi(P) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2\theta & -\sin 2\theta \\ 0 & \sin 2\theta & \cos 2\theta \end{bmatrix}. \tag{3.18}$$

This shows that the image of φ in SO_3 contains the subgroup H of all rotations about the point $(1, 0, 0)^t$. This point corresponds to the matrix $E = \begin{bmatrix} i & \\ & -i \end{bmatrix}$. Since the unit sphere C is a conjugacy class, the operation of SU_2 is transitive. So if Y is any unit vector in $\mathbb{R}^3$, there is an element $Q \in SU_2$ such that $\varphi(Q)(1, 0, 0)^t = Y$, or in matrix notation, such that $QEQ^* = A$. The conjugate subgroup $\varphi(Q)H\varphi(Q)^*$ of rotations about Y is also in the image of φ. Since every element of SO_3 is a rotation, φ is surjective. □

The cosets making up the Hopf fibration which was mentioned at the end of last section, are the fibres of a continuous surjective map

$$\pi\colon S^3 \longrightarrow S^2 \tag{3.19}$$

from the 3-sphere to the 2-sphere. To define π, we interpret S^3 as the special unitary group SU_2, and S^2 as the conjugacy class C of trace-zero matrices, as above. We

set $E = \begin{bmatrix} i & \\ & -i \end{bmatrix}$, and we define $\pi(P) = PEP^*$, for $P \in SU_2$. The proof of the following proposition is left as an exercise.

(3.20) **Proposition.** The fibres of the map π are the left cosets QT of the group T of diagonal matrices in SU_2. □

4. THE SPECIAL LINEAR GROUP $SL_2(\mathbb{R})$

Since the special unitary group is a sphere, it is a compact set. As an example of a noncompact group, we will describe the special linear group $SL_2(\mathbb{R})$. To simplify notation, we denote $SL_2(\mathbb{R})$ by SL_2 in this section.

Invertible 2×2 matrices operate by left multiplication on the space $\mathbb{R}^2$ of column vectors, and we can look at the associated action on *rays* in $\mathbb{R}^2$. A ray is a half line $R = \{rX \mid r \geq 0\}$. The set of rays is in bijective correspondence with the points on the unit circle S^1, the ray R corresponding to the point $R \cap S^1$.

Our group SL_2 operates by left multiplication on the set of rays. Let us denote by H the stabilizer of the ray $R_1 = \{re_1\}$ in $SL_2(\mathbb{R})$. It consists of matrices

$$B = \begin{bmatrix} a & b \\ 0 & a^{-1} \end{bmatrix}, \tag{4.1}$$

where a is positive and b is arbitrary.

The rotation group SO_2 is another subgroup of SL_2, and it operates transitively on the set of rays.

(4.2) **Proposition.** The map $f\colon SO_2 \times H \longrightarrow SL_2$ defined by $f(Q,B) = QB$ is a homeomorphism (but not a group homomorphism).

Proof. Notice that $H \cap SO_2 = \{I\}$. Therefore f is injective [Chapter 2 (8.6)]. To prove surjectivity of f, let P be an arbitrary element of SL_2, and let R_1 be the ray $\{re_1 \mid r > 0\}$. Choose a rotation $Q \in SO_2$ such that $PR_1 = QR_1$. Then $Q^{-1}P$ is in the stabilizer H, say $Q^{-1}P = B$, or

$$P = QB. \tag{4.3}$$

Since f is defined by matrix multiplication, it is a continuous map. Also, in the construction of the inverse map, the rotation Q depends continuously on P because the ray PR_1 does. Then $B = Q^{-1}P$ also is a continuous function of P, and this shows that f^{-1} is continuous as well. □

Note that H can be identified by the rule $B \longleftrightarrow (a, b)$ with the product space (*positive reals*) $\times \mathbb{R}$. And the space of positive reals is homeomorphic by the log

function to the space $\mathbb{R}$ of all real numbers. Thus H is homeomorphic to $\mathbb{R}^2$. Since SO_2 is a circle, we find that

(4.4) *$SL_2(\mathbb{R})$ is homeomorphic to the product space $S^1 \times \mathbb{R}^2$.*

The special linear group can be related to the Lorentz group $O_{2,1}$ of two-dimensional space-time by a method analogous to that used in Section 3 for the orthogonal representation of SU_2. Let the coordinates in $\mathbb{R}^3$ be y_1, y_2, t, with the Lorentz form

$$y_1y_1' + y_2y_2' - tt', \tag{4.5}$$

and let W be the space of real trace-zero matrices. Using the basis

$$\begin{bmatrix} 1 & \\ & -1 \end{bmatrix}, \quad \begin{bmatrix} & 1 \\ 1 & \end{bmatrix}, \quad \begin{bmatrix} & 1 \\ -1 & \end{bmatrix}, \tag{4.6}$$

we associate to a coordinate vector $(y_1, y_2, t)^t$ the matrix

$$A = \begin{bmatrix} y_1 & y_2+t \\ y_2-t & -y_1 \end{bmatrix}. \tag{4.7}$$

We use this representation of trace-zero matrices because the Lorentz form (4.5) has a simple matrix interpretation on such matrices:

$$\langle A, A' \rangle = y_1y_1' + y_2y_2' - tt' = \tfrac{1}{2} \operatorname{trace}(AA'). \tag{4.8}$$

The group SL_2 acts on W by conjugation,

$$P, A \rightsquigarrow PAP^{-1}, \tag{4.9}$$

and this action preserves the Lorentz form on W, because

$$\operatorname{trace}(AA') = \operatorname{trace}((PAP^{-1})(PA'P^{-1})),$$

as in the previous section. Since conjugation is a linear operator on W, it defines a homomorphism φ: $SL_2 \longrightarrow GL_3(\mathbb{R})$. Since conjugation preserves the Lorentz form, the image $\varphi(P)$ of P is an element of $O_{2,1}$.

(4.10) **Theorem.** The kernel of the homomorphism φ is the subgroup $\{\pm I\}$, and the image is the path-connected component $O_{2,1}^0$ of $O_{2,1}$ containing the identity I. Therefore $O_{2,1}^0 \approx SL_2(\mathbb{R})/\{\pm I\}$.

It can be shown that the two-dimensional Lorentz group has four path-connected components.

The fact that the kernel of φ is $\{\pm I\}$ is easy to check, and the last assertion of the theorem follows from the others. We omit the proof that the image of φ is the subgroup $O_{2,1}^0$. □

5. ONE-PARAMETER SUBGROUPS

In Chapter 4, we defined the exponential of a matrix by the series

$$e^A = I + (1/1!)A + (1/2!)A^2 + (1/3!)A^3 + \cdots. \tag{5.1}$$

We will now use this function to describe the homomorphisms from the additive group of real numbers to the general linear group, which are differentiable functions of the variable $t \in \mathbb{R}$. Such a homomorphism is called a *one-parameter subgroup* of GL_n. (Actually, this use of the phrase "one-parameter subgroup" to describe such homomorphisms is a misnomer. The image of φ should be called the subgroup.)

(5.2) **Proposition.**

(a) Let A be an arbitrary real or complex matrix, and let GL_n denote $GL_n(\mathbb{R})$ *or* $GL_n(\mathbb{C})$, according to the case. The map φ: $\mathbb{R}^+ \longrightarrow GL_n$ defined by $\varphi(t) = e^{tA}$ is a group homomorphism.

(b) Conversely, let φ: $\mathbb{R}^+ \longrightarrow GL_n$ be a homomorphism which is a differentiable function of the variable $t \in \mathbb{R}^+$, and let A denote its derivative $\varphi'(0)$ at the origin. Then $\varphi(t) = e^{tA}$ for all t.

Proof. For any two real numbers r, s, the two matrices rA and sA commute. So Chapter 4 (7.13) tells us that

$$e^{(r+s)A} = e^{rA}e^{sA}. \tag{5.3}$$

This shows that $\varphi(t) = e^{tA}$ is a homomorphism. Conversely, let φ be a differentiable homomorphism $\mathbb{R}^+ \longrightarrow GL_n$. The assumption that φ is a homomorphism allows us to compute its derivative at any point. Namely, it tells us that $\varphi(t + \Delta t) = \varphi(\Delta t)\varphi(t)$ and $\varphi(t) = \varphi(0)\varphi(t)$. Thus

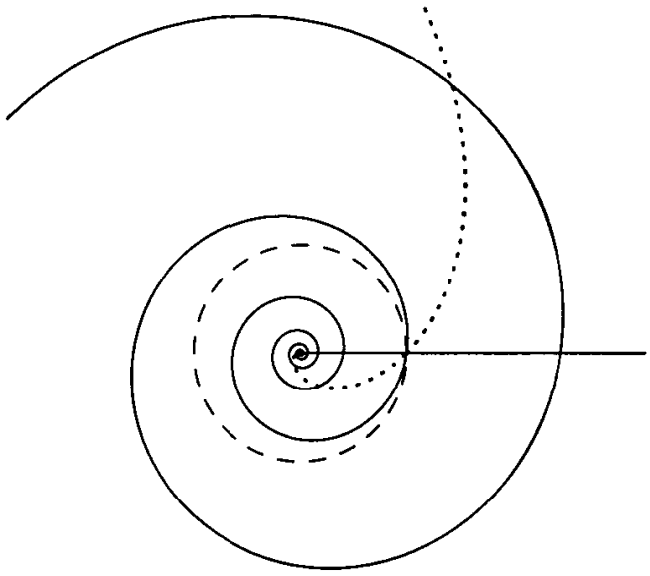
$$\frac{\varphi(t + \Delta t) - \varphi(t)}{\Delta t} = \frac{\varphi(\Delta t) - \varphi(0)}{\Delta t}\varphi(t). \tag{5.4}$$

Letting $\Delta t \longrightarrow 0$, we find $\varphi'(t) = \varphi'(0)\varphi(t) = A\varphi(t)$. Therefore $\varphi(t)$ is a matrix-valued function which solves the differential equation

$$\frac{d\varphi}{dt} = A\varphi. \tag{5.5}$$

The function e^{tA} is another solution, and both solutions take the value I at $t = 0$. It follows that $\varphi(t) = e^{tA}$ [see Chapter 4 (8.14)]. □

By the proposition we have just proved, the one-parameter subgroups all have the form $\varphi(t) = e^{tA}$. They are in bijective correspondence with $n \times n$ matrices.



(5.6) **Figure.** Some one-parameter subgroups of $\mathbb{C}^\times = GL_1(\mathbb{C})$.

Now suppose that a subgroup of G of GL_n is given. We may ask for one-parameter subgroups of G, meaning homomorphisms φ: $\mathbb{R}^+ \to G$, or, equivalently, homomorphisms to GL_n whose image is in G. Since a one-parameter subgroup of GL_n is determined by a matrix, this amounts to asking for the matrices A such that $e^{tA} \in G$ for all t. It turns out that linear groups of positive dimension always have one-parameter subgroups and that they are not hard to determine for a particular group.

(5.7) **Examples.**

(a) The usual parametrization of the unit circle in the complex plane is a one-parameter subgroup of U_1:

$$t \rightsquigarrow e^{ti} = \cos t + i \sin t .$$

(b) A related example is obtained for SO_2 by setting

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \text{ Then } e^{tA} = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}.$$

This is the standard parametrization of the rotation matrices.

In examples (a) and (b), the image of the homomorphism is the whole subgroup.

(c) Let A be the 2×2 matrix unit e_{12}. Then since $A^2 = 0$, all but two terms of the series expansion for the exponential vanish, and

$$e^{tA} = I + e_{12}t = \begin{bmatrix} 1 & t \\ & 1 \end{bmatrix}.$$

In this case the exponential map defines an isomorphism from $\mathbb{R}^+$ to its image, which is the group of triangular matrices with diagonal entries equal to 1.

(d) The one-parameter subgroups of SU_2 are the conjugates of the group of diagonal special unitary matrices, the longitudes described in (2.13). □

Instead of attempting to state a general theorem describing one-parameter subgroups of a group, we will determine them for the orthogonal and special linear groups as examples of the methods used. We will need to know that the exponential function on matrices has an inverse function.

(5.8) **Proposition.** The matrix exponential maps a small neighborhood S of 0 in $\mathbb{R}^{n\times n}$ homeomorphically to a neighborhood T of I.

Proof. This proposition follows from the Inverse Function Theorem, which states that a differentiable function f: $\mathbb{R}^k \longrightarrow \mathbb{R}^k$ has an inverse function at a point p if the Jacobian matrix $(\partial f_i/\partial x_j)(p)$ is invertible. We must check this for the matrix exponential at the zero matrix in $\mathbb{R}^{n\times n}$. This is a notationally unpleasant but easy computation. Let us denote a variable matrix by X. The Jacobian matrix is the $n^2 \times n^2$ matrix whose entries are $(\partial(e^X)_{\alpha\beta}/\partial X_{ij})|_{x=0}$. We use the fact that $d/dt\,(e^{tA})|_{t=0} = A$. It follows directly from the definition of the partial derivative that $(\partial e^X/\partial X_{ij})|_{X=0} = (de^{te_{ij}}/dt)|_{t=0} = e_{ij}$. Therefore $(\partial(e^X)_{\alpha\beta}/\partial X_{ij})|_{x=0} = 0$ if $\alpha, \beta \neq i, j$ and $(\partial(e^X)_{ij}/\partial X_{ij})|_{x=0} = 1$. The Jacobian matrix is the $n^2 \times n^2$ identity matrix. □

We will now describe one-parameter subgroups of the orthogonal group O_n. Here we are asking for the matrices A such that e^{tA} is orthogonal for all t.

(5.9) **Lemma.** If A is skew-symmetric, then e^A is orthogonal. Conversely, there is a neighborhood S' of 0 in $\mathbb{R}^{n\times n}$ such that if e^A is orthogonal and $A \in S'$, then A is skew-symmetric.

Proof. To avoid confusing the variable t with the symbol for the transpose matrix, we denote the transpose of the matrix A by A^* here. If A is skew-symmetric, then $e^{(A^*)} = e^{-A}$. The relation $e^{(A^*)} = (e^A)^*$ is clear from the definition of the exponential, and $e^{-A} = (e^A)^{-1}$ by Chapter 4 (8.10). Thus $(e^A)^* = e^{(A^*)} = e^{-A} = (e^A)^{-1}$. This shows that e^A is orthogonal. For the converse, we choose S' small enough so that if $A \in S'$, then $-A$ and A^* are in the neighborhood S of Proposition (5.8). Suppose that $A \in S'$ and that e^A is orthogonal. Then $e^{(A^*)} = e^{-A}$, and by Proposition (5.8), this means that A is skew-symmetric. □

(5.10) **Corollary.** The one-parameter subgroups of the orthogonal group O_n are the homomorphisms $t \rightsquigarrow e^{tA}$, where A is a real skew-symmetric matrix.

Proof. If A is skew-symmetric, tA is skew-symmetric for all t. So e^{tA} is orthogonal for all t, which means that e^{tA} is a one-parameter subgroup of O_n. Conversely, suppose that e^{tA} is orthogonal for all t. For sufficiently small ϵ, ϵA is in the neighborhood S' of the lemma, and $e^{\epsilon A}$ is orthogonal. Therefore ϵA is skew-symmetric, and this implies that A is skew-symmetric too. □

This corollary is illustrated by Example (5.7b).

Next, let us consider the special linear group $SL_n(\mathbb{R})$.

(5.11) **Proposition.** Let A be a matrix whose trace is zero. Then e^A has determinant 1. Conversely, there is a neighborhood S' of 0 in $\mathbb{R}^{n\times n}$ such that if $A \in S'$ and $\det e^A = 1$, then $\operatorname{trace} A = 0$.

Proof. The first assertion follows from the pretty formula

$$e^{\operatorname{tr}A} = \det e^A, \tag{5.12}$$

where $\operatorname{tr}A$ denotes the trace of the matrix. This formula follows in turn from the fact that if the eigenvalues of a complex matrix A are $\lambda_1,\dots,\lambda_n$, then the eigenvalues of e^A are $e^{\lambda_1},\dots,e^{\lambda_n}$. We leave the proof of this fact as an exercise. Using it, we find

$$e^{\operatorname{tr}A} = e^{\lambda_1+\cdots+\lambda_n} = e^{\lambda_1}\cdots e^{\lambda_n} = \det e^A.$$

For the converse, we note that if $|x| < 1$, $e^x = 1$ implies $x = 0$. We choose S' small enough so that $\operatorname{tr} A < 1$ if $A \in S'$. Then if $\det e^A = e^{\operatorname{tr}A} = 1$ and if $A \in S'$, $\operatorname{tr} A = 0$. □

(5.13) **Corollary.** The one-parameter subgroups of the special linear group $SL_n(\mathbb{R})$ are the homomorphisms $t \rightsquigarrow e^{tA}$, where A is a real $n \times n$ matrix whose trace is zero. □

The simplest one-parameter subgroup of $SL_2(\mathbb{R})$ is described in Example (5.7c).

6. THE LIE ALGEBRA

As always, we think of a linear group G as a subset of $\mathbb{R}^{n\times n}$ or of $\mathbb{C}^{n\times n}$. The space of vectors tangent to G at the identity matrix I, which we will describe in this section, is called the *Lie algebra* of the group.

We will begin by reviewing the definition of tangent vector. If $\varphi(t) = (\varphi_1(t),\dots,\varphi_k(t))$ is a differentiable path in $\mathbb{R}^k$, its velocity vector $v = \varphi'(t)$ is tangent to the path at the point $x = \varphi(t)$. This is the basic observation from which the definition of tangent vector is derived.

Suppose that we are given a subset S of $\mathbb{R}^k$. A vector v is said to be *tangent* to S at a point x if there is a differentiable path $\varphi(t)$ lying entirely in S, such that $\varphi(0) = x$ and $\varphi'(0) = v$.

If our subset S is the locus of zeros of one or more polynomial functions $f(x_1,\dots,x_k)$, it is called a *real algebraic set:*

$$S = \{x \mid f(x) = 0\}. \tag{6.1}$$

For example, the unit circle in $\mathbb{R}^2$ is a real algebraic set because it is the locus of zeros of the polynomial $f(x_1, x_2) = x_1^2 + x_2^2 - 1 = 0$.

The chain rule for differentiation provides a necessary condition for a vector to be tangent to a real algebraic set S. Let $\varphi(t)$ be a path in S, and let $x = \varphi(t)$ and

$v = \varphi'(t)$. Since the path is in S, the functions $f(\varphi(t))$ vanish identically; hence their derivatives also vanish identically:

$$(6.2) \qquad 0 = \frac{d}{dt} f(\varphi(t)) = \frac{\partial f}{\partial x_1} v_1 + \cdots + \frac{\partial f}{\partial x_k} v_k = (\nabla f(x) \cdot v),$$

where $\nabla f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_k}\right)$ is the gradient vector.

(6.3) **Corollary.** Let S be a real algebraic set in $\mathbb{R}^k$, defined as the locus of zeros of one or more polynomial functions $f(x)$. The tangent vectors to S at x are orthogonal to the gradients $\nabla f(x)$. □

For instance, if S is the unit circle and x is the point $(1,0)$, then the gradient vector $\nabla f(0)$ is $(2,0)$. Corollary (6.3) tells us that tangent vectors at $(1,0)$ have the form $(0,c)$, that is, that they are vertical, which is as it should be.

Computing tangent vectors by means of parametrized paths is clumsy because there are many paths with the same tangent. If we are interested only in the tangent vector, then we can throw out all of the information contained in a path except for the first-order term of its Taylor expansion. To do this systematically, we introduce a formal *infinitesimal element* ϵ. This means that we work algebraically with the rule

$$(6.4) \qquad \epsilon^2 = 0.$$

Just as with complex numbers, where the rule is $i^2 = -1$, we can use this rule to define a multiplication on the vector space

$$E = \{a + b\epsilon \mid a, b \in \mathbb{R}\}$$

of formal linear combinations of $(1, \epsilon)$ with real coefficients. The rule for multiplication is

$$(6.5) \qquad (a + b\epsilon)(c + d\epsilon) = ac + (bc + ad)\epsilon.$$

In other words, we expand formally, using the relations $\epsilon c = c\epsilon$ for all $c \in \mathbb{R}$ and $\epsilon^2 = 0$. As with complex numbers, addition is vector addition:

$$(a + b\epsilon) + (c + d\epsilon) = (a + c) + (b + d)\epsilon.$$

The main difference between $\mathbb{C}$ and E is that E is not a field, because ϵ has no multiplicative inverse. [It is a ring (see Chapter 10).]

Given a point x of $\mathbb{R}^k$ and a vector $v \in \mathbb{R}^k$, the sum $x + v\epsilon$ is a vector with entries in E which we interpret intuitively as an *infinitesimal change in x, in the direction of v*. Notice that we can evaluate a polynomial $f(x) = f(x_1, \ldots, x_k)$ at $x + v\epsilon$ using Taylor's expansion. Since $\epsilon^2 = 0$, the terms of degree ≥ 2 in ϵ drop out, and we are left with an element of E:

$$(6.6) \qquad f(x+v\epsilon) = f(x) + \left(\frac{\partial f}{\partial x_1} v_1 + \cdots + \frac{\partial f}{\partial x_k} v_k\right)\epsilon = f(x) + (\nabla f(x) \cdot v)\epsilon.$$

Working with rule (6.4) amounts to ignoring the higher-order terms in ϵ. Thus the dot product $(\nabla f(x) \cdot v)$ represents the infinitesimal change in f which results when we make an infinitesimal change in x in the direction of v.

Going back to a real algebraic set S defined by the polynomial equations $f(x) = 0$, let x be a point of S. Then $f(x) = 0$, so (6.6) tells us that

(6.7) $$f(x + v\epsilon) = 0 \text{ if and only if } (\nabla f(x) \cdot v) = 0,$$

which is the same as the condition we obtained in Corollary (6.3). This suggests the following definition: Let S be a real algebraic set, defined by the polynomial equations $f(x) = 0$. A vector v is called an *infinitesimal tangent* to S at x if

(6.8) $$f(x + v\epsilon) = 0.$$

(6.9) **Corollary.** Let x be a point of a real algebraic set S. Every tangent to S at x is an infinitesimal tangent. □

Notice that if we fix $x \in S$, the equations $(\nabla f(x) \cdot v) = 0$ are linear and homogeneous in v. So the infinitesimal tangent vectors to S at x form a subspace of the space of all vectors.

Actually, our terminology is slightly ambiguous. The definition of an infinitesimal tangent depends on the equations f, not only on the set S. *We must have particular equations in mind when speaking of infinitesimal tangents.*

For sets S which are sufficiently smooth, the converse of (6.9) is also true: Every infinitesimal tangent is a tangent vector. When this is the case, we can compute the space of tangent vectors at a point $x \in S$ by solving the linear equations $(\nabla f(x) \cdot v) = 0$ for v, which is relatively easy. However, this converse will not be true at "singular points" of the set S, or if the defining equations for S are chosen poorly. For example, let S denote the union of the two coordinate axes in $\mathbb{R}^2$. This is a real algebraic set defined by the single equation $x_1x_2 = 0$. It is clear that at the origin a tangent vector must be parallel to one of the two axes. On the other hand, $\nabla f = (x_2, x_1)$, which is zero when $x_1 = x_2 = 0$. Therefore *every* vector is an infinitesimal tangent to S at the origin.

This completes our general discussion of tangent vectors. We will now apply this discussion to the case that the set S is one of our linear groups G in $\mathbb{R}^{n\times n}$ or $\mathbb{C}^{n\times n}$. The tangent vectors to G will be n^2-dimensional vectors, and we will represent them by matrices too. As we said earlier, the vectors tangent to G at the identity I form the *Lie algebra* of the group.

The first thing to notice is that every one-parameter subgroup e^{tA} of our linear group G is a parametrized path. We already know that its velocity vector $(de^{tA}/dt)_{t=0}$ is A. So A represents a tangent vector to G at the identity—it is in the Lie algebra. For example, the unitary group U_1 is the unit circle in the complex plane, and e^{ti} is a one-parameter subgroup of U_1. The velocity vector of this one-parameter subgroup at $t = 0$ is the vector i, which is indeed a tangent vector to the unit circle at the point 1.

A matrix group G which is a real algebraic set in $\mathbb{R}^{n\times n}$ is called a *real algebraic group*. The classical linear groups such as $SL_n(\mathbb{R})$ and O_n are real algebraic, because their defining equations are polynomial equations in the matrix entries. For example, the group $SL_2(\mathbb{R})$ is defined by the single polynomial equation $\det P = 1$:

$$x_{11}x_{22} - x_{12}x_{21} - 1 = 0.$$

The orthogonal group O_3 is defined by nine polynomials f_{ij} expressing the condition $P^tP = I$:

$$f_{ij} = x_{1i}x_{1j} + x_{2i}x_{2j} + x_{3i}x_{3j} - \delta_{ij} = 0, \quad \delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}.$$

Complex groups such as the unitary groups can also be made into real algebraic groups in $\mathbb{R}^{2n\times n}$ by separating the matrix entries into their real and imaginary parts.

It is a fact that for every infinitesimal tangent A to a real algebraic group G at the identity, e^{tA} is a one-parameter subgroup of G. In other words, there is a one-parameter subgroup leading out from the identity in an arbitrary tangent direction. This is quite remarkable for a nonabelian group, but it is true with essentially no restriction. Unfortunately, though this fact is rather easy to check for a particular group, it is fairly hard to give a general proof. Therefore we will content ourselves with verifying particular cases.

Having an infinitesimal element available, we may work with matrices whose entries are in E. Such a matrix will have the form $A + B\epsilon$, where A,B are real matrices. Intuitively, $A + B\epsilon$ represents an infinitesimal change in A in the direction of the matrix B. The rule for multiplying two such matrices is the same as (6.5):

$$(A + B\epsilon)(C + D\epsilon) = AC + (AD + BC)\epsilon. \tag{6.10}$$

The product $B\epsilon D\epsilon$ is zero because $(b_{ij}\epsilon)(d_{kl}\epsilon) = 0$ for all values of the indices.

Let G be a real algebraic group. To determine its infinitesimal tangent vectors at the identity, we must determine the matrices A such that

$$I + A\epsilon, \tag{6.11}$$

which represents an infinitesimal change in I in the direction of the matrix A, satisfies the equations defining G. This is the definition (6.8) of an infinitesimal tangent.

Let us make this computation for the special linear group $SL_n(\mathbb{R})$. The defining equation for this group is $\det P = 1$. So A is an infinitesimal tangent vector if $\det(I + A\epsilon) = 1$. To describe this condition, we must calculate the change in the determinant when we make an infinitesimal change in I. The formula is nice:

$$\det(I + A\epsilon) = 1 + (\text{trace } A)\epsilon. \tag{6.12}$$

The proof of this formula is left as an exercise. Using it, we find that A is an infinitesimal tangent vector if and only if $\text{trace } A = 0$.

(6.13) **Proposition.** The following conditions on a real $n \times n$ matrix A are equivalent:

(i) $\operatorname{trace} A = 0$;
(ii) e^{tA} is a one-parameter subgroup of $SL_n(\mathbb{R})$;
(iii) A is in the Lie algebra of $SL_n(\mathbb{R})$;
(iv) A is an infinitesimal tangent to $SL_n(\mathbb{R})$ at I.

Proof. Proposition (5.11) tells us that (i) $\Rightarrow$ (ii). Since A is tangent to the path e^{tA} at $t = 0$, (ii) $\Rightarrow$ (iii). The implication (iii) $\Rightarrow$ (iv) is (6.9), and (iv) $\Rightarrow$ (i) follows from (6.12). □

There is a general principle at work here. We have three sets of matrices A: those such that e^{tA} is a one-parameter subgroup of G, those which are in the Lie algebra, and those which are infinitesimal tangents. Let us denote these three sets by $\operatorname{Exp}(G)$, $\operatorname{Lie}(G)$, and $\operatorname{Inf}(G)$. They are related by the following inclusions:

$$\operatorname{Exp}(G) \subset \operatorname{Lie}(G) \subset \operatorname{Inf}(G). \tag{6.14}$$

The first inclusion is true because A is the tangent vector to e^{tA} at $t = 0$, and the second holds because every tangent vector is an infinitesimal tangent. If $\operatorname{Exp}(G) = \operatorname{Inf}(G)$, then these two sets are also equal to $\operatorname{Lie}(G)$. Since the computations of $\operatorname{Exp}(G)$ and $\operatorname{Inf}(G)$ are easy, this gives us a practical way of determining the Lie algebra. A general theorem exists which implies that $\operatorname{Exp}(G) = \operatorname{Inf}(G)$ for every real algebraic group, provided that its defining equations are chosen properly. However, it isn't worthwhile proving the general theorem here.

We will now make the computation for the orthogonal group O_n. The defining equation for O_n is the matrix equation $P^tP = I$. In order for A to be an infinitesimal tangent at the identity, it must satisfy the relation

$$(I + A\epsilon)^t(I + A\epsilon) = I. \tag{6.15}$$

The left side of this relation expands to $I + (A^t + A)\epsilon$, so the condition that $I + A\epsilon$ be orthogonal is $A^t + A = 0$, or A is skew-symmetric. This agrees with the condition (5.10) for e^{tA} to be a one-parameter subgroup of O_n.

(6.16) **Proposition.** The following conditions on a real $n \times n$ matrix A are equivalent:

(i) A is skew-symmetric;
(ii) e^{tA} is a one-parameter subgroup of O_n;
(iii) A is in the Lie algebra of O_n;
(iv) A is an infinitesimal tangent to O_n at I. □

The Lie algebra of a linear group has an additional structure, an operation called the *Lie bracket*. The Lie bracket is the law of composition defined by the rule

$$[A,B] = AB - BA. \tag{6.17}$$

This law of composition is not associative. It does, however, satisfy an identity called the *Jacobi identity*,

$$[A,[B,C]] + [B,[C,A]] + [C,[A,B]] = 0, \tag{6.18}$$

which is a substitute for the associative law.

To show that the bracket is a law of composition on the Lie algebra, we must check that if A, B are in Lie(G), then $[A,B]$ is also in Lie(G). This can be done easily for any particular group. For the special linear group, the required verification is that if A, B have trace zero, then $AB - BA$ also has trace zero. This is true, because trace AB = trace BA. Or let $G = O_n$, so that the Lie algebra is the space of skew-symmetric matrices. We must verify that if A, B are skew, then $[A,B]$ is skew too:

$$[A,B]^t = (AB - BA)^t = B^tA^t - A^tB^t = BA - AB = -[A,B],$$

as required.

The bracket operation is important because it is the infinitesimal version of the commutator $PQP^{-1}Q^{-1}$. To see why this is so, we must work with two infinitesimals ϵ,δ, using the rules $\epsilon^2 = \delta^2 = 0$ and $\epsilon\delta = \delta\epsilon$. Note that the inverse of the matrix $I + A\epsilon$ is $I - A\epsilon$. So if $P = I + A\epsilon$ and $Q = I + B\delta$, the commutator expands to

$$(I + A\epsilon)(I + B\delta)(I - A\epsilon)(I - B\delta) = I + (AB - BA)\epsilon\delta. \tag{6.19}$$

Intuitively, the bracket is in the Lie algebra because the product of two elements in G, even infinitesimal ones, is in G, and therefore the commutator of two elements is also in G.

Using the bracket operation, we can also define the concept of Lie algebra abstractly.

(6.20) **Definition.** A *Lie algebra* V over a field F is a vector space together with a law of composition

$$V \times V \longrightarrow V$$
$$v\ ,\ w \rightsquigarrow [v,w]$$

called the *bracket*, having these properties:

(i) bilinearity: $[v_1 + v_2, w] = [v_1, w] + [v_2, w]$, $[cv, w] = c[v, w]$,
$[v, w_1 + w_2] = [v, w_1] + [v, w_2]$, $[v, cw] = c[v, w]$,

(ii) skew symmetry: $[v, v] = 0$,

(iii) Jacobi identity: $[u,[v,w]] + [v,[w,u]] + [w,[u,v]] = 0$,

for all $u, v, w \in V$ and all $c \in F$.

The importance of Lie algebras comes from the fact that, being vector spaces, they are much easier to work with than the linear groups themselves, and at the same time the classical groups are nearly determined by their Lie algebras. In other

words, the infinitesimal structure of the group at the identity element is almost enough to determine the group.

7. TRANSLATION IN A GROUP

We will use one more notion from topology in this section—the definition of manifold in $\mathbb{R}^k$. This definition is reviewed in the appendix [Definition (3.12)]. Do not be discouraged if you are not familiar with the concept of manifold. You can learn what is necessary without much trouble as we go along.

Let P be a fixed element of a matrix group G. We know that left multiplication by P is a bijective map from G to itself:

$$G \xrightarrow{m_P} G \tag{7.1}$$

$$X \rightsquigarrow PX,$$

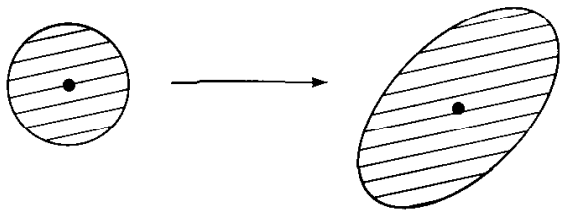
because it has the inverse function $m_{P^{-1}}$. The maps m_P and $m_{P^{-1}}$ are continuous, because matrix multiplication is continuous. Thus m_P is a homeomorphism from G to itself (not a homomorphism). It is also called *left translation* by P, in analogy with translation in the plane, which is left translation in the additive group $\mathbb{R}^{2+}$.

The important property of a group which is implied by the existence of these maps is *homogeneity*. Multiplication by P is a homeomorphism which carries the identity element I to P. So the topological structure of the group G is the same near I as it is near P, and since P is arbitrary, it is the same in the neighborhoods of any two points of the group. This is analogous to the fact that the plane looks the same at any two points.

Left multiplication in SU_2 happens to be defined by an *orthogonal* change of the coordinates (x_1, x_2, x_3, x_4), so it is a rigid motion of the 3-sphere. But multiplication by a matrix needn't be a rigid motion, so the sense in which any group is homogeneous is weaker. For example, let G be the group of real invertible diagonal 2×2 matrices, and let us identify the elements of G with the points (a, d) in the plane, which are not on the coordinate axes. Multiplication by the matrix

$$P = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \tag{7.2}$$

distorts the group G, but it does so continuously.



(7.3) **Figure.** Left multiplication in a group.

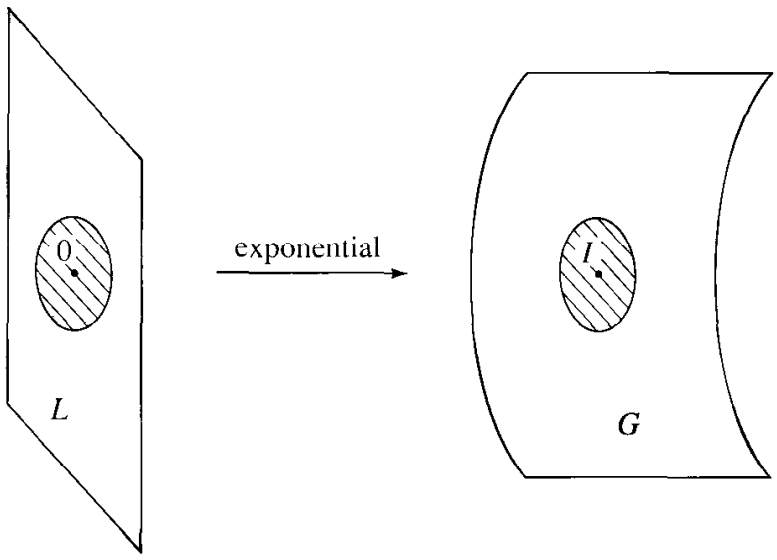
Now the only geometrically reasonable subsets of $\mathbb{R}^k$ which have this homogeneity property are manifolds. A manifold M of dimension d is a subset which is locally homeomorphic to $\mathbb{R}^d$ at any one of its points, meaning that every point $p \in M$ has a neighborhood homeomorphic to an open set in $\mathbb{R}^d$ [see Appendix (3.12)]. It isn't surprising that the classical groups, being homogeneous, are manifolds, though there are subgroups of GL_n which aren't. The group $GL_n(\mathbb{Q})$ of invertible matrices with rational coefficients, for example, is a rather ugly set when viewed geometrically, though it is an interesting group. The following theorem gives a satisfactory answer to the question of which linear groups are manifolds:

(7.4) **Theorem.** Let G be a subgroup of $GL_n(\mathbb{R})$ which is a closed set in $\mathbb{R}^{n\times n}$. Then G is a manifold.

Giving the proof of this theorem here would take us too far afield. Instead, we will illustrate the theorem by showing that the orthogonal groups O_n are manifolds. The proofs for other classical groups are similar.

(7.5) **Proposition.** The orthogonal group O_n is a manifold of dimension $\frac{1}{2}n(n-1)$.

Proof. Let us denote the group O_n by G and denote its Lie algebra, the space of skew-symmetric matrices, by L. Proposition (5.9) tells us that for matrices A near 0, $A \in L$ if and only if $e^A \in G$. Also, the exponential is a homeomorphism from a neighborhood of 0 in $\mathbb{R}^{n\times n}$ to a neighborhood of I. Putting these two facts together, we find that the exponential defines a homeomorphism from a neighborhood of 0 in L to a neighborhood of I in G. Since L is a vector space of dimension $\frac{1}{2}n(n-1)$, it is a manifold. This shows that the condition of being a manifold is satisfied by the orthogonal group at the identity. On the other hand, we saw above that any two points in G have homeomorphic neighborhoods. Therefore G is a manifold, as claimed. □



(7.6) **Figure.**

Here is another application of the principle of homogeneity:

(7.7) **Proposition.** Let G be a path-connected matrix group, and let $H \subset G$ be a subgroup which contains a nonempty open subset of G. Then $H = G$.

Proof. By hypothesis, H contains a nonempty open subset U of G. Since left multiplication by $g \in G$ is a homeomorphism, gU is also open in G. Each translate gU is contained in a single coset of H, namely in gH. Since the translates of U cover G, they cover each coset. In this way, each coset is a union of open subsets of G, and hence it is open itself. So G is partitioned into open subsets—the cosets of H. Now a path-connected set is not a disjoint union of proper open subsets [see Appendix, Proposition (3.11)]. Thus there can be only one coset, and $H = G$. □

We will now apply this proposition to determine the normal subgroups of SU_2.

(7.8) **Theorem.** The only proper normal subgroup of SU_2 is its center $\{\pm I\}$.

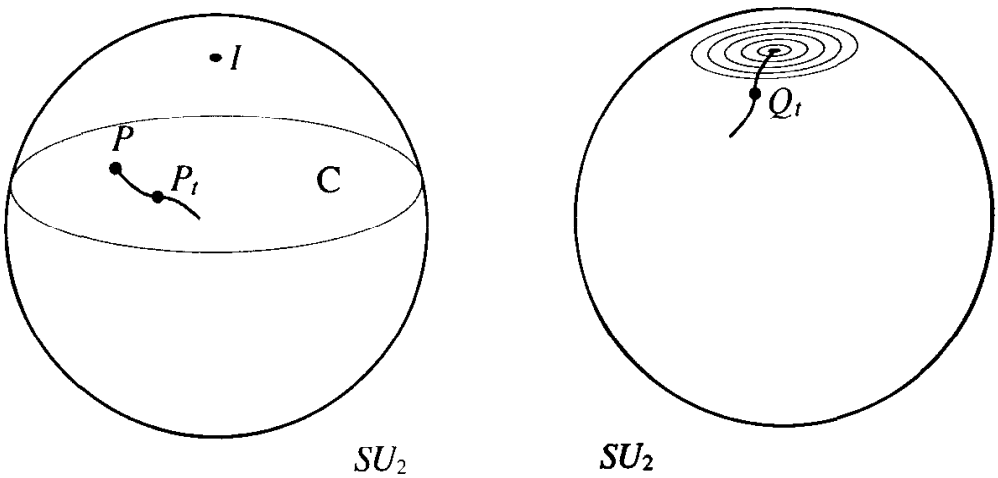
Since there is a surjective map φ: $SU_2 \longrightarrow SO_3$ whose kernel is $\{\pm I\}$, the rotation group is isomorphic to a quotient group of SU_2 [Chapter 2 (10.9)]:

$$SO_3 \approx SU_2/\{\pm I\}. \tag{7.9}$$

(7.10) **Corollary.** SO_3 is a simple group; that is, it has no proper normal subgroup.

Proof. The inverse image of a normal subgroup in SO_3 is a normal subgroup of SU_2 which contains $\{\pm I\}$ [Chapter 2 (7.4)]. Theorem (7.8) tells us that there are no proper ones. □

Proof of Theorem (7.8). It is enough to show that if N is a normal subgroup of SU_2 which is not contained in the center $\{\pm I\}$, then N is the whole group. Now since N is normal, it is a union of conjugacy classes [Chapter 6 (2.5)]. And we have



(7.11) **Figure.**

seen that the conjugacy classes are the latitudes, the 2-spheres (2.8). By assumption, N contains a matrix $P \neq \pm I$, so it contains the whole conjugacy class $C = C_P$, which is a 2-sphere. Intuitively, this set looks big enough to generate SU_2. For it has dimension 2 and is not a subgroup. So the set S of all products $P^{-1}Q$ with $P,Q \in C$ is larger than C. Therefore S ought to have dimension 3, which is the dimension of SU_2 itself, so it ought to contain an open set in the group.

To make this intuitive reasoning precise, we choose a nonconstant continuous map from the unit interval $[0, 1]$ to C such that $P_0 = P$ and $P_1 \neq P$. We form the path

$$Q_t = P^{-1}P_t. \tag{7.12}$$

Then $Q_0 = I$, and $Q_1 \neq I$, so this path leads out from I. Since P and P_t are in N, Q_t is in N for every $t \in [0, 1]$. We don't need to know anything else about the path Q_t.

Let $f(t)$ be the function trace Q_t. This is a continuous function on the interval $[0, 1]$. Note that $f(0) = 2$, while $f(1) = \tau < 2$ because $Q_1 \neq I$. By continuity, all values between τ and 2 are taken on by f in the interval.

Since N is normal, it contains the conjugacy class of Q_t for every t. So since trace Q_t takes on all values near 2, Proposition (2.9) tells us that N contains all matrices in SU_2 whose trace is sufficiently near to 2, and this includes all matrices sufficiently near to I. So N contains an open neighborhood of I in SU_2. Now SU_2, being a sphere, is path-connected, so Proposition (7.7) completes the proof. □

We can also apply translation in a group G to tangent vectors. If A is a tangent vector at the identity and if $P \in G$ is arbitrary, then PA is a tangent vector to G at the point P. Intuitively, this is because $P(I + A\epsilon) = P + PA\epsilon$ is the product of elements in G, so it lies in G itself. As always, this heuristic is easy to check for a particular group. We fix A, and associate the tangent vector PA to the element P of G. In this way we obtain what is called a *tangent vector field* on the group G. Since A is nonzero and P is invertible, this vector field does not vanish at any point. Now just the existence of a tangent vector field which is nowhere zero puts strong restrictions on the space G. For example, it is a theorem of topology that any vector field on the 2-sphere must vanish at some point. That is why the 2-sphere has no group structure. But the 3-sphere, being a group, has tangent vector fields which are nowhere zero.

8. SIMPLE GROUPS

Recall that a group G is called *simple* if it is not the trivial group and if it contains no proper normal subgroup (Chapter 6, Section 2). So far, we have seen two nonabelian simple groups: the icosahedral group $I \approx A_5$ [Chapter 6 (2.3)] and the rotation group SO_3 (7.10). This section discusses the classification of simple groups. We will omit most proofs.

Simple groups are important for two reasons. First of all, if a group G has a proper normal subgroup N, then the structure of G is partly described when we

know the structure of N and of the quotient group G/N. If N or G/N has a normal subgroup, we can further decompose the structure of these groups. In this way we may hope to describe a particular finite group G, by building it up inductively from simple groups.

Second, though the condition of being simple is a very strong restriction, simple groups often appear. The classical linear groups are almost simple. For example, we saw in the last section that SU_2 has center $\{\pm I\}$ and that $SU_2/\{\pm I\} \approx SO_3$ is a simple group. The other classical groups have similar properties.

In order to focus attention, we will restrict our discussion here to the complex groups. We will use the symbol Z to denote the center of any group. The following theorem would take too much time to prove here, but we will illustrate it in the special case of $SL_2(\mathbb{C})$.

(8.1) **Theorem.**

(a) The center Z of the special linear group $SL_n(\mathbb{C})$ is a cyclic group, generated by the matrix ζI where $\zeta = e^{2\pi i/n}$. The quotient group $SL_n(\mathbb{C})/Z$ is simple if $n \geq 2$.

(b) The center Z of the complex special orthogonal group $SO_n(\mathbb{C})$ is $\{\pm I\}$ if n is even, and is the trivial group $\{I\}$ if n is odd. The group SO_n/Z is simple if $n = 3$ or if $n \geq 5$.

(c) The center Z of the symplectic group $SP_{2n}(\mathbb{C})$ is $\{\pm I\}$, and $SP_{2n}(\mathbb{C})/Z$ is simple if $n \geq 1$. □

The group $SL_n(\mathbb{C})/Z$ is called the *projective group* and is denoted by $PSL_n(\mathbb{C})$:

$$PSL_n(\mathbb{C}) = SL_n(\mathbb{C})/Z, \quad \text{where } Z = \{\zeta I \mid \zeta^n = 1\}. \tag{8.2}$$

To illustrate Theorem (8.1), we will prove that $PSL_2(\mathbb{C}) = SL_2(\mathbb{C})/\{\pm I\}$ is simple. In fact, we will show that $PSL_2(F)$ is a simple group for almost all fields F.

(8.3) **Theorem.** Let F be a field which is not of characteristic 2 and which contains at least seven elements. Then the only proper normal subgroup of $SL_2(F)$ is the subgroup $\{\pm I\}$. Thus $PSL_2(F) = SL_2(F)/\{\pm I\}$ is a simple group.

Since the center of $SL_2(F)$ is a normal subgroup, it follows from the theorem that it is the group $\{\pm I\}$.

(8.4) **Corollary.** There are infinitely many nonabelian finite simple groups.

Proof of Theorem (8.3). The proof is algebraic, but it is closely related to the geometric proof given for the analogous assertion for SU_2 in the last section. Our procedure is to conjugate and multiply until the group is generated. To simplify notation, we will denote $SL_2(F)$ by SL_2. Let N be a normal subgroup of SL_2 which contains a matrix $A \neq \pm I$. We must show that $N = SL_2$. Since one possibility is that N

is the normal subgroup generated by A and its conjugates, we must show that the conjugates of this one matrix suffice to generate the whole group.

The first step in our proof will be to show that N contains a triangular matrix different from $\pm I$. Now if our given matrix A has eigenvalues in the field F, then it will be conjugate to a triangular matrix. But since we want to handle arbitrary fields, we can not make this step so easily. Though easy for the complex numbers, this step is the hardest part of the proof for a general field.

(8.5) **Lemma.** *N contains a triangular matrix $A \neq \pm I$.*

Proof. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a matrix in N which is different from $\pm I$. If $c = 0$, then A is the required matrix.

Suppose that $c \neq 0$. In this case, we will construct a triangular matrix out of A and its conjugates. We first compute the conjugate

$$\begin{bmatrix} 1 & x \\ & 1 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & -x \\ & 1 \end{bmatrix} = \begin{bmatrix} a+xc & * \\ c & d-xc \end{bmatrix} = A'.$$

Since $c \neq 0$, we may choose x so that $a + xc = 0$. The matrix A' is in N, so N contains a matrix whose upper left entry is zero. We replace A by this matrix, so that it has the form $A = \begin{bmatrix} & b \\ c & d \end{bmatrix}$. Unfortunately the zero is in the wrong place.

Note that since $\det A = 1$, $bc = -1$ in our new matrix A. We now compute the commutator $P^{-1}A^{-1}PA$ with a diagonal matrix:

$$P^{-1}A^{-1}PA = \begin{bmatrix} u & \\ & u^{-1} \end{bmatrix}\begin{bmatrix} d & -b \\ -c & \end{bmatrix}\begin{bmatrix} u^{-1} & \\ & u \end{bmatrix}\begin{bmatrix} & b \\ c & d \end{bmatrix} = \begin{bmatrix} u^2 & (1-u^2)bd \\ & u^{-2} \end{bmatrix}.$$

This matrix, which is in our normal subgroup N, is as required unless it is $\pm I$. If so, then $u^2 = \pm 1$ and $u^4 = 1$. But we are free to form the matrix P with an arbitrary element u in $F^\times$. We will show [Chapter 11 (1.8)] that the polynomial $x^4 - 1$ has at most four roots in any field. So there are at most four elements $u \in F$ with $u^4 - 1$. Our hypothesis is that $F^\times$ contains at least five elements. So we can choose $u \in F^\times$ with $u^4 \neq 1$. Then $P^{-1}A^{-1}PA$ is the required matrix. □

(8.6) **Lemma.** *N contains a matrix of the form $\begin{bmatrix} 1 & u \\ & 1 \end{bmatrix}$, with $u \neq 0$.*

Proof. By the previous lemma, N contains a triangular matrix $A = \begin{bmatrix} a & b \\ & d \end{bmatrix} \neq \pm I$. If $d \neq a$, let $A' = \begin{bmatrix} a & b' \\ & d \end{bmatrix}$ be its conjugate by the matrix $\begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}$. Then $b' = b + d - a$. Since $\det A = ad = 1$, the product

$$A'^{-1}A = \begin{bmatrix} d & -b' \\ & a \end{bmatrix}\begin{bmatrix} a & b \\ & d \end{bmatrix} = \begin{bmatrix} 1 & ad-d^2 \\ & 1 \end{bmatrix}$$

is the required matrix. If $a = d$, then $a = \pm 1$ because $\det A = 1$, and it follows that $b \neq 0$. In this case, one of the two matrices A or A^2 is as required. □

(8.7) **Lemma.** Let F be a field. The conjugacy class in SL_2 of the matrix $\begin{bmatrix} 1 & u \\ & 1 \end{bmatrix}$ contains the matrices $\begin{bmatrix} 1 & \\ -u & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & a^2u \\ & 1 \end{bmatrix}$, for all $a \neq 0$.

Proof.

$$\begin{bmatrix} & -1 \\ 1 & \end{bmatrix}\begin{bmatrix} 1 & u \\ & 1 \end{bmatrix}\begin{bmatrix} & 1 \\ -1 & \end{bmatrix} = \begin{bmatrix} 1 & \\ -u & 1 \end{bmatrix} \text{ and } \begin{bmatrix} a & \\ & a^{-1} \end{bmatrix}\begin{bmatrix} 1 & u \\ & 1 \end{bmatrix}\begin{bmatrix} a^{-1} & \\ & a \end{bmatrix} = \begin{bmatrix} 1 & a^2u \\ & 1 \end{bmatrix}. \square$$

(8.8) **Lemma.** Let F be a field of characteristic $\neq 2$. The additive group F^+ of the field is generated by the squares of elements of F.

Proof. We show that every element $x \in F$ can be written in the form $a^2 - b^2 = (a + b)(a - b)$, with $a, b \in F$. To do this, we solve the system of linear equations $a + b = 1$, $a - b = x$. This is where the assumption that the characteristic of F is not 2 is used. In characteristic 2, these equations need not have a solution. □

(8.9) **Lemma.** Let F be a field of characteristic $\neq 2$. If a normal subgroup N of $SL_2(F)$ contains a matrix $\begin{bmatrix} 1 & u \\ & 1 \end{bmatrix}$ with $u \neq 0$, then it contains all such matrices.

Proof. The set of x such that $\begin{bmatrix} 1 & x \\ & 1 \end{bmatrix} \in N$ is a subgroup of F^+, call it S. We want to show that $S = F^+$. Lemma (8.7) shows that if $u \in S$, then $a^2u \in S$ for all $a \in F$. Since the squares generate F^+, the set of elements $\{a^2u \mid a \in F\}$ generates the additive subgroup F^+u of F^+, and this subgroup is equal to F^+ because u is invertible. Thus $S = F^+$, as required. □

(8.10) **Lemma.** For every field F, the group $SL_2(F)$ is generated by the elementary matrices $\begin{bmatrix} 1 & u \\ & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & \\ u & 1 \end{bmatrix}$.

Proof. We perform row reduction on a matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL_2(F)$, using only the matrices of this form. We start work on the first column, reducing it to e_1. We eliminate the case $c = 0$ by adding the first row to the second if necessary. Then we add a multiple of the second row to the first to change a to 1. Finally, we clear out the entry c. At this point, the matrix has the form $A' = \begin{bmatrix} 1 & b' \\ 0 & d' \end{bmatrix}$. Then $d' = \det A' = \det A = 1$, and we can clear out the entry b', ending up with the identity matrix. Since we needed four operations or less to reduce to the identity, A is a product of at most four of these elementary matrices. □

The proof of Theorem (8.3) is completed by combining Lemmas (8.6), (8.7), (8.9), and (8.10). □

A famous theorem of Cartan asserts that the list (8.1) of simple groups is almost complete. Of course there are other simple groups; for instance, we have just proved that $PSL_2(F)$ is simple for most fields F. But if we restrict ourselves to complex algebraic groups, the list of simple groups becomes very short.

A subgroup G of $GL_n(\mathbb{C})$ is called a *complex algebraic group* if it is the set of solutions of a finite system of polynomial equations in the matrix entries. This is analogous to the concept of a real algebraic group introduced in Section 6. It will not be apparent why the property of being defined by polynomial equations is a reasonable one, but one thing is easy to see: Except for the unitary groups U_n and SU_n, all the complex classical groups are complex algebraic groups.

(8.11) **Theorem.**

(a) The groups $PSL_n(\mathbb{C}) = SL_n(\mathbb{C})/Z$, $SO_n(\mathbb{C})/Z$, and $SP_{2n}(\mathbb{C})/Z$ are path-connected complex algebraic groups.

(b) In addition to the isomorphism classes of these groups, there are exactly five isomorphism classes of simple, path-connected complex algebraic groups, called the *exceptional groups*.

Theorem (8.11) is too hard to prove here. It is based on a classification of the corresponding Lie algebras. What we should learn is that there are not many simple algebraic groups. This ought to be reassuring after the last chapter, where structures on a vector space were introduced one after the other, each with its own group of symmetries. There seemed to be no end. Now we see that we actually ran across most of the possible symmetry types, at least those associated to *simple* algebraic groups. It is no accident that these structures are important. □

A large project, the classification of the *finite* simple groups, was completed in 1980. The finite simple groups we have seen are the groups of prime order, the icosahedral group $I \approx A_5$ [Chapter 6 (2.3)], and the groups $PSL_2(F)$ where F is a finite field (8.3), but there are many more. The alternating groups A_n are simple for all $n \geq 5$.

Linear groups play a dominant role in the classification of the finite simple groups as well as of the complex algebraic groups. Each of the forms (8.11) leads to a whole series of finite simple groups when finite fields are substituted for the complex field. Also, some finite simple groups are analogous to the unitary groups. All of these finite linear groups are said to be of *Lie type*.

According to Theorem (8.3), $PSL_2(\mathbb{F}_7)$ is a finite simple group; its order is 168. This is the second smallest simple group; A_5 is the smallest. The orders of the smallest nonabelian simple groups are

(8.12) 60, 168, 360, 504, 660, 1092, 2448.

For each of these seven integers N, there is a single isomorphism class of simple groups of order N, and it is represented by $PSL_2(F)$ for a suitable finite field F. [The alternating group A_5 happens to be isomorphic to $PSL_2(\mathbb{F}_5)$.]

In addition to the groups of prime order, the alternating groups, and the groups of Lie type, there are exactly 26 finite simple groups called the *sporadic groups*. The smallest sporadic group is the *Matthieu group* M_{11}, whose order is 7920. The largest is called the *Monster*; its order is roughly 10^{53}. So the finite simple groups form a list which, though longer, is somewhat analogous to the list (8.11) of simple algebraic groups.

It seems unfair to crow about the successes of a theory
and to sweep all its failures under the rug.

Richard Brauer

EXERCISES

1. The Classical Linear Groups

1. **(a)** Find a subgroup of $GL_2(\mathbb{R})$ which is isomorphic to $\mathbb{C}^\times$.
   (b) Prove that for every n, $GL_n(\mathbb{C})$ is isomorphic to a subgroup of $GL_{2n}(\mathbb{R})$.
2. Show that $SO_2(\mathbb{C})$ is not a bounded set in $\mathbb{C}^4$.
3. Prove that $SP_2(\mathbb{R}) = SL_2(\mathbb{R})$, but that $SP_4(\mathbb{R}) \neq SL_4(\mathbb{R})$.
4. According to Sylvester's Law, every 2×2 real symmetric matrix is congruent to exactly one of six standard types. List them. If we consider the operation of $GL_2(\mathbb{R})$ on 2×2 matrices by $P, A \rightsquigarrow PAP^t$, then Sylvester's Law asserts that the symmetric matrices form six orbits. We may view the symmetric matrices as points in $\mathbb{R}^3$, letting (x,y,z) correspond to the matrix $\begin{bmatrix} x & y \\ y & z \end{bmatrix}$. Find the decomposition of $\mathbb{R}^3$ into orbits explicitly, and make a clear drawing showing the resulting geometric configuration.
5. A matrix P is orthogonal if and only if its columns form an orthonormal basis. Describe the properties that the columns of a matrix must have in order for it to be in the Lorentz group $O_{3,1}$.
6. Prove that there is no continuous isomorphism from the orthogonal group O_4 to the Lorentz group $O_{3,1}$.
7. Describe by equations the group $O_{1,1}$, and show that it has four connected components.
8. Describe the orbits for the operation of $SL_2(\mathbb{R})$ on the space of real symmetric matrices by $P, A \rightsquigarrow PAP^t$.
9. Let F be a field whose characteristic is not 2. Describe the orbits for the action $P, A \rightsquigarrow PAP^t$ of $GL_2(F)$ on the space of 2×2 symmetric matrices with coefficients in F.
10. Let $F = \mathbb{F}_2$. Classify the orbits of $GL_n(F)$ for the action on the space of symmetric $n \times n$ matrices by finding representatives for each congruence class.

11. Prove that the following matrices are symplectic, if the blocks are $n \times n$: $\begin{bmatrix} & -I \\ I & \end{bmatrix}$, $\begin{bmatrix} A^t & \\ & A^{-1} \end{bmatrix}$, $\begin{bmatrix} I & B \\ & I \end{bmatrix}$, where $B = B^t$ and A is invertible.
12. Prove that the symplectic group $SP_{2n}(\mathbb{R})$ operates transitively on $\mathbb{R}^{2n}$.
*13. Prove that $SP_{2n}(\mathbb{R})$ is path-connected, and conclude that every symplectic matrix has determinant 1.

2. The Special Unitary Group SU_2

1. Let P, Q be elements of SU_2, represented by the real vectors (x_1, x_2, x_3, x_4), (y_1, y_2, y_3, y_4). Compute the real vector which corresponds to the product PQ.
2. Prove that the subgroup SO_2 of SU_2 is conjugate to the subgroup T of diagonal matrices.
3. Prove that SU_2 is path-connected. Do the same for SO_3.
4. Prove that U_2 is homeomorphic to the product $S^3 \times S^1$.
5. Let G be the group of matrices of the form $\begin{bmatrix} x & y \\ & 1 \end{bmatrix}$, where $x, y \in \mathbb{R}$ and $x > 0$. Determine the conjugacy classes in G, and draw them in the (x, y)-plane.

*6. (a) Prove that every element P (2.4) of SU_2 can be written as a product: $P = DRD'$, where $D, D' \in T$ (2.13), and $R \in SO_2$ is a rotation through an angle θ with $0 \leq \theta \leq \pi/2$.
(b) Assume that the matrix entries a, b of P are not zero. Prove that this representation is unique, except that the pair D, D' can be replaced by $-D, -D'$.
(c) Describe the double cosets TPT, $P \in SU_2$. Prove that if the entries a, b of P are not zero, then the double coset is homeomorphic to a torus, and describe the remaining double cosets.

3. The Orthogonal Representation of SU_2

1. Compute the stabilizer H of the matrix $\begin{bmatrix} & 1 \\ 1 & \end{bmatrix}$ for the action of conjugation by SU_2, and describe $\varphi(P)$ for $P \in H$.
2. Prove that every great circle in SU_2 is a coset of one of the longitudes (2.14).
3. Find a subset of $\mathbb{R}^3$ which is homeomorphic to the space $S \times \Theta$ of (3.4).
4. Derive a formula for $\langle A, A \rangle$ in terms of the determinant of A.
5. The rotation group SO_3 may be mapped to the 2-sphere by sending a rotation matrix to its first column. Describe the fibres of this map.
6. Extend the map φ defined in this section to a homomorphism Φ: $U_2 \longrightarrow SO_3$, and describe the kernel of Φ.
7. Prove by direct computation that the matrix (3.11) is in SO_3.
*8. Describe the conjugacy classes in SO_3 carefully, relating them to the conjugacy classes of SU_2.
9. Prove that the operation of SU_2 on any conjugacy class other than $\{I\}$, $\{-I\}$ is by rotations of the sphere.
10. Find a bijective correspondence between elements of SO_3 and pairs (p, v) consisting of a point p on the unit 2-sphere S and a unit tangent vector v to S at p.

11. Prove Proposition (3.20).

*12. (a) Calculate left multiplication by a fixed matrix P in SU_2 explicitly in terms of the coordinates x_1, x_2, x_3, x_4. Prove that it is multiplication by a 4×4 orthogonal matrix Q, hence that it is a rigid motion of the unit 3-sphere S^3.
(b) Prove that Q is orthogonal by a method similar to that used in describing the orthogonal representation: Express dot product of the vectors (x_1, x_2, x_3, x_4), (x_1', x_2', x_3', x_4') corresponding to two matrices $P, P' \in SU_2$ in terms of matrix operations.
(c) Determine the matrix which describes the operation of conjugation by a fixed matrix P on SU_2.

*13. (a) Let H_i be the subgroup of SO_3 of rotations about the x_i-axis, $i = 1, 2, 3$. Prove that every element of SO_3 can be written as a product ABA', where $A, A' \in H_1$ and $B \in H_2$. Prove that this representation is unique unless $B = I$.
(b) Describe the double cosets H_1QH_1 geometrically.

*14. Let H_i be the subgroup of SO_3 of rotations about the x_i-axis. Prove that every element $Q \in SO_3$ can be written in the form $A_1A_2A_3$, with $A_i \in H_i$.

4. The Special Linear Group $SL_2(\mathbb{R})$

1. Let $G = SL_2(\mathbb{C})$. Use the operation of G on rays $\{rX\} \mid r \in \mathbb{R}, r > 0\}$ in $\mathbb{C}^2$ to prove that G is homeomorphic to the product $SU_2 \times H$, where H is the stabilizer of the ray $\{re_1\}$, and describe H explicitly.

2. (a) Prove that the rule $P, A \rightsquigarrow PAP^*$ defines an operation of $SL_2(\mathbb{C})$ on the space W of all hermitian matrices.
(b) Prove that the function $\langle A, A' \rangle = \det(A + A') - \det A - \det A'$ is a bilinear form on W, whose signature is $(3, 1)$.
(c) Use (a) and (b) to define a homomorphism φ: $SL_2(\mathbb{C}) \longrightarrow O_{3,1}$, whose kernel is $\{\pm I\}$.
*(d) Prove that the image of φ is the connected component of the identity in $O_{3,1}$.

3. Let P be a matrix in $SO_3(\mathbb{C})$.
(a) Prove that 1 is an eigenvalue of P.
(b) Let X_1, X_2 be eigenvectors for P, with eigenvalues λ_1, λ_2. Prove that $X_1{}^t X_2 = 0$, unless $\lambda_1 = \lambda_2^{-1}$.
(c) Prove that if X is an eigenvector with eigenvalue 1 and if $P \neq I$, then $X^t X \neq 0$.

4. Let $G = SO_3(\mathbb{C})$.
(a) Prove that left multiplication by G is a transitive operation on the set of vectors X such that $X^t X = 1$.
(b) Determine the stabilizer of e_1 for left multiplication by G.
(c) Prove that G is path-connected.

5. One-Parameter Subgroups

1. Determine the differentiable homomorphisms from $\mathbb{C}^+$ to $SL_n(\mathbb{C})$.
2. Describe all one-parameter subgroups of $\mathbb{C}^\times$.
3. Describe by equations the images of all one-parameter subgroups of the group of real 2×2 diagonal matrices, and make a neat drawing showing them.
4. Let φ: $\mathbb{R}^+ \longrightarrow GL_n(\mathbb{R})$ be a one-parameter subgroup. Prove that $\ker \varphi$ is either trivial, or the whole group, or else it is infinite cyclic.

5. Find the conditions on a matrix A so that e^{tA} is a one-parameter subgroup of the special unitary group SU_n, and compute the dimension of that group.
6. Let G be the group of real matrices of the form $\begin{bmatrix} x & y \\ & 1 \end{bmatrix}$, $x > 0$.
   (a) Determine the matrices A such that e^{tA} is a one-parameter subgroup of G.
   (b) Compute e^A explicitly for the matrices determined in (a).
   (c) Make a drawing showing the one-parameter subgroups in the (x, y)-plane.
7. Prove that the images of the one-parameter subgroups of SU_2 are the conjugates of T (see Section 3). Use this to give an alternative proof of the fact that these conjugates are the longitudes.
8. Determine the one-parameter subgroups of U_2.
9. Let $\varphi(t) = e^{tA}$ be a one-parameter subgroup of G. Prove that the cosets of im φ are matrix solutions of the differential equation $dX/dt = AX$.
10. Can a one-parameter subgroup of $GL_n(\mathbb{R})$ cross itself?
*11. Determine the differentiable homomorphisms from SO_2 to $GL_n(\mathbb{R})$.

6. The Lie Algebra

1. Compute $(A + B\epsilon)^{-1}$, assuming that A is invertible.
2. Compute the infinitesimal tangent vectors to the plane curve $y^2 = x^3$ at the point (1,1) and at the point (0,0).
3. (a) Sketch the curve C: $x_2^2 = x_1^3 - x_1^2$.
   (b) Prove that this locus is a manifold of dimension 1 if the origin is deleted.
   (c) Determine the tangent vectors and the infinitesimal tangents to C at the origin.
4. Let S be a real algebraic set defined by one equation $f = 0$.
   (a) Show that the equation $f^2 = 0$ defines the same locus S.
   (b) Show that $\nabla(f^2)$ vanishes at every point x of S, hence that every vector is an infinitesimal tangent at x, when the defining equation is taken to be $f^2 = 0$.
5. Show that the set defined by $xy = 1$ is a subgroup of the group of diagonal matrices $\begin{bmatrix} x & \\ & y \end{bmatrix}$, and compute its Lie algebra.
6. Determine the Lie algebra of the unitary group.
7. (a) Prove the formula $\det(I + A\epsilon) = 1 + \text{trace } A\epsilon$.
   (b) Let A be an invertible matrix. Compute $\det(A + B\epsilon)$.
8. (a) Show that O_2 operates by conjugation on its Lie algebra.
   (b) Show that the operation in (a) is compatible with the bilinear form $\langle A, B \rangle = \frac{1}{2}$ trace AB.
   (c) Use the operation in (a) to define a homomorphism $O_2 \longrightarrow O_2$, and describe this homomorphism explicitly.
9. Compute the Lie algebra of the following: (a) U_n; (b) SU_n; (c) $O_{3,1}$; (d) $SO_n(\mathbb{C})$. In each case, show that e^{tA} is a one-parameter subgroup if and only if $I + A\epsilon$ lies in the group.
*10. Determine the Lie algebra of $G = SP_{2n}(\mathbb{R})$, using block form $M = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$.
11. (a) Show that $\mathbb{R}^3$ becomes a Lie algebra if the bracket is defined to be the cross product $[X, Y] = X \times Y = (x_2y_3 - y_2x_3, x_3y_1 - y_1x_3, x_1y_2 - x_2y_1)$.
   (b) Show that this Lie algebra is isomorphic to the Lie algebra of SO_3.

12. Classify all complex Lie algebras of dimension ≤ 3.

*13. The *adjoint representation* of a linear group G is the representation by conjugation on its Lie algebra: $G \times L \longrightarrow L$ is defined to be $P, A \rightsquigarrow PAP^{-1}$. The form $\langle A, A' \rangle = \text{trace}(AA')$ on L is called the *Killing form*. For each of the following groups, verify that if $P \in G$ and $A \in L$, then $PAP^{-1} \in L$, and prove that the Killing form is symmetric and bilinear and that the operation is compatible with the form, i.e., that $\langle A, A' \rangle = \langle PAP^{-1}, PA'P^{-1} \rangle$.
(a) SO_n (b) SU_n (c) $O_{3,1}$ (d) $SO_n(\mathbb{C})$ (e) $SP_{2n}(\mathbb{R})$

14. Prove that the Killing form is negative definite on the Lie algebra of (a) SU_n and (b) SO_n.

15. Determine the signature of the Killing form on the Lie algebra of $SL_n(\mathbb{R})$.

16. (a) Use the adjoint representation of SU_n to define a homomorphism φ: $SU_n \longrightarrow SO_m$, where $m = n^2 - 1$.
(b) Show that when $n = 2$, this representation is equivalent to the orthogonal representation defined in Section 3.

17. Use the adjoint representation of $SL_2(\mathbb{C})$ to define an isomorphism $SL_2(\mathbb{C})/\{\pm I\} \approx SO_3(\mathbb{C})$.

7. Translation in a Group

1. Compute the dimensions of the following groups.
(a) SU_n (b) $SO_n(\mathbb{C})$ (c) $SP_{2n}(\mathbb{R})$ (d) $O_{3,1}$

2. Using the exponential, find all solutions near I of the equation $P^2 = I$.

3. Find a path-connected, nonabelian subgroup of $GL_2(\mathbb{R})$ of dimension 2.

4. (a) Show that every positive definite hermitian matrix A is the square of another positive definite hermitian matrix B.
(b) Show that B is uniquely determined by A.

*5. Let A be a nonsingular matrix, and let B be a positive definite hermitian matrix such that $B^2 = AA^*$.
(a) Show that A^*B^{-1} is unitary.
(b) Prove the *Polar decomposition:* Every nonsingular matrix A is a product $A = PU$, where P is positive definite hermitian and U is unitary.
(c) Prove that the Polar decomposition is unique.
(d) What does this say about the operation of left multiplication by the unitary group U_n on the group GL_n?

*6. State and prove an analogue of the Polar decomposition for real matrices.

*7. (a) Prove that the exponential map defines a bijection between the set of all hermitian matrices and the set of positive definite hermitian matrices.
(b) Describe the topological structure of $GL_2(\mathbb{C})$ using the Polar decomposition and (a).

8. Let B be an invertible matrix. Describe the matrices A such that $P = e^A$ is in the centralizer of B.

*9. Let S denote the set of matrices $P \in SL_2(\mathbb{R})$ with trace r. These matrices can be written in the form $\begin{bmatrix} x & y \\ z & r-x \end{bmatrix}$, where (x, y, z) lies on the quadric $x(r - x) - yz = 1$.
(a) Show that the quadric is either a hyperbola of one or two sheets, or else a cone, and determine the values of r which correspond to each type.
(b) In each case, determine the decomposition of the quadric into conjugacy classes.

(c) Extend the method of proof of Theorem (7.11) to show that the only proper normal subgroup of $SL_2(\mathbb{R})$ is $\{\pm I\}$.

10. Draw the tangent vector field P_A to the group $\mathbb{C}^\times$, when $A = 1 + i$.

8. Simple Groups

1. Which of the following subgroups of $GL_n(\mathbb{C})$ are complex algebraic groups?
(a) $GL_n(\mathbb{Z})$ (b) SU_n (c) upper trianguiar matrices

2. (a) Write the polynomial functions in the matrix entries which define $SO_n(\mathbb{C})$.
(b) Write out the polynomial equations which define the symplectic group.
(c) Show that the unitary group U_n can be defined by real polynomial equations in the real and imaginary parts of the matrix entries.

3. Determine the centers of the groups $SL_n(\mathbb{R})$ and $SL_n(\mathbb{C})$.

4. Describe isomorphisms (a) $PSL_2(\mathbb{F}_2) \approx S_3$ and (b) $PSL_2(\mathbb{F}_3) \approx A_4$.

5. Determine the conjugacy classes of $GL_2(\mathbb{F}_3)$.

6. Prove that $SL_2(F) = PSL_2(F)$ for any field F of characteristic 2.

7. (a) Determine all normal subgroups of $GL_2(\mathbb{C})$ which contain its center $Z = \{cI\}$.
(b) Do the same for $GL_2(\mathbb{R})$.

8. For each of the seven orders (8.12), determine the order of the field F such that $PSL_2(F)$ has order n.

***9.** Prove that there is a simple group of order 3420.

10. (a) Let Z be the center of $GL_n(\mathbb{C})$. Is $PSL_n(\mathbb{C})$ isomorphic to $GL_n(\mathbb{C})/Z$?
(b) Answer the same question as in (a), with $\mathbb{R}$ replacing $\mathbb{C}$.

11. Prove that $PSL_2(\mathbb{F}_5)$ is isomorphic to A_5.

***12.** Analyze the proof of Theorem (8.3) to prove that $PSL_2(F)$ is a simple group when F is a field of characteristic 2, except for the one case $F = \mathbb{F}_2$.

13. (a) Let P be a matrix in the center of SO_n, and let A be a skew-symmetric matrix. Prove that $PA = AP$ by differentiating the matrix function e^{At}.
(b) Prove that the center of SO_n is trivial if n is odd and is $\{\pm I\}$ if n is even and ≥ 4.

14. Compute the orders of the following groups.
(a) $SO_2(\mathbb{F}_3)$ and $SO_3(\mathbb{F}_3)$
(b) $SO_2(\mathbb{F}_5)$ and $SO_3(\mathbb{F}_5)$

***15.** (a) Consider the operation of $SL_2(\mathbb{C})$ by conjugation on the space V of complex 2×2 matrices. Show that with the basis $(e_{11}, e_{12}, e_{21}, e_{22})$ of V, the matrix of conjugation by $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ has the block form $\begin{bmatrix} aB & bB \\ cB & dB \end{bmatrix}$, where $B = (A^t)^{-1} = \begin{bmatrix} d & -c \\ -b & a \end{bmatrix}$.
(b) Prove that this operation defines a homomorphism φ: $SL_2(\mathbb{C}) \longrightarrow GL_4(\mathbb{C})$, and that the image of φ is isomorphic to $PSL_2(\mathbb{C})$.
(c) Prove that $PSL_2(\mathbb{C})$ is an algebraic group by finding polynomial equations in the entries y_{ij} of a 4×4 matrix whose solutions are precisely the matrices in im φ.

***16.** Prove that $PSL_n(\mathbb{C})$ is a simple group.

***17.** There is no simple group of order $2^5 \cdot 7 \cdot 11$. Assuming this, determine the next smallest order after 2448 for a nonabelian simple group.

Miscellaneous Exercises

1. *Quaternions* are expressions of the form $\alpha = a + bi + cj + dk$, where $a, b, c, d \in \mathbb{R}$. They can be added and multiplied using the rules of multiplication for the quaternion group [Chapter 2 (2.12)].
   (a) Let $\overline{\alpha} = a - bi - cj - dk$. Compute $\alpha\overline{\alpha}$.
   (b) Prove that every $\alpha \neq 0$ has a multiplicative inverse.
   (c) Prove that the set of quaternions α such that $a^2 + b^2 + c^2 + d^2 = 1$ forms a group under multiplication which is isomorphic to SU_2.
2. The *affine group* $A_n = A_n(\mathbb{R})$ is the group of coordinate changes in $(x_1, \ldots, x_n)$ which is generated by $GL_n(\mathbb{R})$ and by the group T_n of translations: $t_a(x) = x + a$. Prove that T_n is a normal subgroup of A_n and that A_n/T_n is isomorphic to $GL_n(\mathbb{R})$.
3. *Cayley Transform*: Let U denote the set of matrices A such that $I + A$ is invertible, and define $A' = (I - A)(I + A)^{-1}$.
   (a) Prove that if $A \in U$, then $A' \in U$, and prove that $A'' = A$.
   (b) Let V denote the vector space of real skew-symmetric $n \times n$ matrices. Prove that the rule $A \rightsquigarrow (I - A)(I + A)^{-1}$ defines a homeomorphism from a neighborhood of 0 in V to a neighborhood of I in SO_n.
   (c) Find an analogous statement for the unitary group.
   (d) Let $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$. Show that a matrix $A \in U$ is symplectic if and only if $A'^t J = -JA'$.

*4. Let $p(t) = t^2 - ut + 1$ be a quadratic polynomial, with coefficients in the field $F = \mathbb{F}_p$.
   (a) Prove that if p has two distinct roots in F, then the matrices with characteristic polynomial p form two conjugacy classes in $SL_2(F)$, and determine their orders.
   (b) Prove that if p has two equal roots, then the matrices with characteristic polynomial p form three conjugacy classes in $SL_n(F)$, and determine their orders.
   (c) Suppose that p has no roots in F. Determine the centralizer of the matrix $A = \begin{bmatrix} & -1 \\ 1 & u \end{bmatrix}$ in $SL_2(F)$, and compute the order of the conjugacy class of A.
   (d) Find the class equations of $SL_2(\mathbb{F}_3)$ and $SL_2(\mathbb{F}_5)$.
   (e) Find the class equations of $PSL_2(\mathbb{F}_3)$ and $PSL_2(\mathbb{F}_5)$, and reconcile your answer with the class equations of A_4 and A_5.
   (f) Compute the class equation for $SL_2(\mathbb{F}_7)$ and for $PSL_2(\mathbb{F}_7)$. Use the class equation for $PSL_2(\mathbb{F}_7)$ to show that this group is simple.