THE CLASSICAL GROUPS

This week's lecture will cover Chapter 7 in the book, but I will begin more generally by defining the socalled classical (matrix) groups. These will be subgroups of the groups $\mathrm{GL}_n(\mathbb{R})$, $\mathrm{GL}_n(\mathbb{C})$, and $\mathrm{GL}_n(\mathbb{H})$ of invertible $n \times n$-matrices with entries in real numbers, complex numbers, and quaternions, respectively.

If $k = (k, +, \cdot\,)$ is a ring, then we define the opposite ring $k^{\mathrm{op}} = (k, +, \star)$ to have the same set of elements and the same addition but the opposite multiplication

$$a \star b = b \cdot a.$$

If k is a division ring, then so is k^{op}.

Definition 1. Let k be a ring. A ring homomorphism

$$k \xrightarrow{\ \sigma\ } k^{\mathrm{op}}$$

is an antiinvolution, if $\sigma \circ \sigma = \mathrm{id}$.

In particular, an antiinvolution is an isomorphism. We remark that the identity map $\mathrm{id}_k \colon k \to k$ is an antiinvolution if and only if k is commutative. We will often write a^* or $\overline{a}$ instead of $\sigma(a)$.

Example 2. (1) If $k = \mathbb{R}$, then the identity map is an antiinvolution, and one can show that it is the only one.

(2) If $k = \mathbb{C}$, then the identity map and complex conjugation are antiinvolutions.

(3) If $k = \mathbb{H}$, then quaternionic conjugation, which is the map $\sigma \colon \mathbb{H} \to \mathbb{H}$ that to the quaternion $q = a + ib + jc + kd$ assigns the quaternion

$$q^* = a - ib - jc - kd$$

is an antiinvolution. The identity map $\mathrm{id}_{\mathbb{H}} \colon \mathbb{H} \to \mathbb{H}$ is not an antiinvolution.

Definition 3. Let k be a division ring, and let $\sigma \colon k \to k^{\mathrm{op}}$ be an antiinvolution. The adjoint matrix of $A = (a_{ij}) \in M_{m,n}(k)$ is $A^* = (a_{ji}^*) \in M_{n,m}(k)$.[1]

The number of rows in A^* is equal to the number of columns in A and vice versa. So it is only meaningful to ask whether $A = A^*$ if A is a square matrix. If k is a field and $\sigma \colon k \to k^{\mathrm{op}}$ is the identity map, then it is customary to call A^* the transpose matrix of A and to denote it by A^t instead of A^*.

Proposition 4. *Let k be a division ring, and let $\sigma \colon k \to k^{\mathrm{op}}$ be an antiinvolution. For all matrices A, B, and C of appropriate dimensions, the following hold:*

(I1) $(A + B)^* = A^* + B^*$
(I2) $(AB)^* = B^* A^*$
(I3) $E^* = E$
(I4) $(A^*)^* = A$

[1] The notation $A^\dagger$ for the adjoint matrix is also common, particularly in physics.

Proof. Let us prove (2). For the purpose of this proof, given $A \in M_{m,n}(k)$, we write $A^* = (a'_{ij}) \in M_{n,m}(k)$. So $a'_{ij} = a^*_{ji}$ by the definition of the adjoint matrix. We let $A \in M_{m,n}(k)$ and $B \in M_{n,p}(k)$ with product $C = AB \in M_{m,p}(k)$ and calculate

$$c'_{ik} = c^*_{ki} = (\sum_{j=1}^{m} a_{kj}b_{ji})^* = \sum_{j=1}^{m}(a_{kj}b_{ji})^* = \sum_{j=1}^{m} b^*_{ji}a^*_{kj} = \sum_{j=1}^{m} b'_{ij}a'_{jk}.$$

This proves (2), and the remaining identities are proved analogously. □

Definition 5. Let k be a division ring, and let $\sigma\colon k \to k^{\mathrm{op}}$ be an antiinvolution. A square matrix $A \in M_n(k)$ is hermitian, if $A^* = A$, and it is skew-hermitian, if $A^* = -A$.

If k is a field and $\sigma\colon k \to k^{\mathrm{op}}$ is the identity map, then it is customary to say that $A \in M_n(k)$ is symmetric, if $A^t = A$, and that A is skew-symmetric, if $A^t = -A$.

We will now consider vector spaces over the division ring k, and we will always consider right vector spaces in the sense that scalars multiply from the right.

Definition 6. Let k be a division ring, let $\sigma\colon k \to k^{\mathrm{op}}$ be an antiinvolution, and let V be a right k-vector space. A hermitian form on V is a map

$$V \times V \xrightarrow{\langle -,-\rangle} k$$

such that the following hold for all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$ and $a \in k$:

(H1) $\langle \boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z}\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle + \langle \boldsymbol{x}, \boldsymbol{z}\rangle$
(H2) $\langle \boldsymbol{x}, \boldsymbol{y} \cdot a\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle \cdot a$
(H3) $\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z}\rangle = \langle \boldsymbol{x}, \boldsymbol{z}\rangle + \langle \boldsymbol{y}, \boldsymbol{z}\rangle$
(H4) $\langle \boldsymbol{x} \cdot a, \boldsymbol{y}\rangle = a^* \cdot \langle \boldsymbol{x}, \boldsymbol{y}\rangle$
(H5) $\langle \boldsymbol{y}, \boldsymbol{x}\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle^*$

Example 7. Let k be a division ring, and let $\sigma\colon k \to k^{\mathrm{op}}$ be an antiinvolution. Let $k^n = M_{n,1}(k)$ be the right k-vector space of column n-matrices with entries in k. If $A \in M_n(k)$ is a hermitian matrix, then the map $\langle -,-\rangle\colon k^n \times k^n \to k$ defined by $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \boldsymbol{x}^*A\boldsymbol{y}$ is a hermitian form. Conversely, if $\langle -,-\rangle\colon k^n \times k^n \to k$ is a hermitian form, then the matrix $A = (a_{i,j}) \in M_n(k)$ with entries $a_{ij} = \langle \boldsymbol{e}_i, \boldsymbol{e}_j\rangle$ is a hermitian matrix.

If $k = \mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$, and if $\sigma\colon k \to k^{\mathrm{op}}$ is the identity map, complex conjugation, and quaternionic conjugation, respectively, then for all $a \in k$, $a^* = a$ if and only if $a \in \mathbb{R} \subset k$. In particular, if $\langle -,-\rangle$ is a hermitian form on a right real, complex, or quaternionic vector space V, then for all $\boldsymbol{x} \in V$, we have $\langle \boldsymbol{x}, \boldsymbol{x}\rangle \in \mathbb{R}$.

Definition 8. Let $k = \mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$, and let $\sigma\colon k \to k^{\mathrm{op}}$ be the identity map, complex conjugation, and quaternionic conjugation, respectively. A hermitian inner product on a right k-vector space V is a hermitian form $\langle -,-\rangle\colon V \times V \to k$ such that, in addition to (H1)–(H5), the following positivity property holds:

(P) For all $\boldsymbol{0} \neq \boldsymbol{x} \in V$, $\langle \boldsymbol{x}, \boldsymbol{x}\rangle > 0$.

Let $k = \mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$, and let $\sigma\colon k \to k^{\mathrm{op}}$ be the identity map, complex conjugation, and quaternionic conjugation, respectively. The standard hermitian inner product on the right k-vector space $k^n = M_{n,1}(k)$ of column n-vectors is defined to be the map $\langle -,- \rangle \colon k^n \times k^n \to$ given by the matrix product

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^* \boldsymbol{y},$$

which is meaningful, since $\boldsymbol{x}^* \in M_{1,n}(k)$ and $\boldsymbol{y} \in M_{n,1}(k)$.

Definition 9. Let $(U, \langle -,- \rangle_U)$ and $(V, \langle -,- \rangle_V)$ be right real, complex or quaternionic vector spaces with hermitian inner products. A k-linear map $f \colon V \to U$ is an isometry with respect to the given hermitian inner products if

$$\langle f(\boldsymbol{x}), f(\boldsymbol{y}) \rangle_U = \langle \boldsymbol{x}, \boldsymbol{y} \rangle_V$$

for all $\boldsymbol{x}, \boldsymbol{y} \in V$.

An isometry $f \colon U \to V$ is always injective, but it need not be an isomorphism. However, if it is an isomorphism, then the inverse map $f^{-1} \colon U \to V$ is automatically an isometry. In particular, an endomorphism $f \colon V \to V$ of a finite dimensional real, complex, or quaternionic vector space that is an isometry with respect to a given hermitian inner product is automatically an isometric isomorphism.

Definition 10. Let $(U, \langle -,- \rangle_U)$ and $(V, \langle -,- \rangle_V)$ be right real, complex or quaternionic vector spaces with hermitian inner products. Two k-linear maps $f \colon V \to U$ and $g \colon U \to V$ are adjoint with respect to the given hermitian inner products if

$$\langle \boldsymbol{x}, f(\boldsymbol{y}) \rangle_U = \langle g(\boldsymbol{x}), \boldsymbol{y} \rangle_V$$

for all $\boldsymbol{x} \in U$ and $\boldsymbol{y} \in V$.

If both $g \colon U \to V$ and $h \colon U \to V$ are adjoint to $f \colon V \to U$, then $g = h$, so if an adjoint of $f \colon V \to U$ exists, then it is unique. If U and V are finite dimensional, then an adjoint always exists.

Proposition 11. *Let $(U, \langle -,- \rangle_U)$ and $(V, \langle -,- \rangle_V)$ be finite dimensional right real, complex, or quaternionic vector spaces with hermitian inner products, and let $f \colon V \to U$ be a linear map. Let $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_m)$ and $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ be bases of U and V that are orthonormal with respect to $\langle -,- \rangle_U$ and $\langle -,- \rangle_V$, respectively.*[2]

1. *There exists a unique linear map $g \colon U \to V$ that is adjoint to $f \colon V \to U$ with respect to $\langle -,- \rangle_U$ and $\langle -,- \rangle_V$.*
2. *If the matrix $A \in M_{m,n}(k)$ represents $f \colon V \to U$ with respect to the given orthonormal bases, then the adjoint matrix $A^* \in M_{n,m}(k)$ represents $g \colon U \to V$ with respect to these bases.*

Proof. We claim that if $f \colon V \to U$ and $g \colon U \to V$ are the linear maps represented by $A \in M_{m,n}(k)$ and $A^* \in M_{n,m}(k)$ with respect to the given orthonormal bases, then these two maps are adjoint with respect to the given hermitian inner products. Indeed, let $\boldsymbol{u} \in U$ and $\boldsymbol{v} \in V$, and let $\boldsymbol{x} \in k^m$ and $\boldsymbol{y} \in k^n$ be their coordinates with respect to the given bases. Since the bases are orthonormal, we find

$$\langle \boldsymbol{u}, f(\boldsymbol{v}) \rangle_U = \boldsymbol{x}^* A \boldsymbol{y} = \boldsymbol{x}^* (A^*)^* \boldsymbol{y} = (A^* \boldsymbol{x})^* \boldsymbol{y} = \langle g(\boldsymbol{u}), \boldsymbol{v} \rangle_V.$$

This proves the proposition, since an adjoint map, if it exists, is unique. □

[2] This means that $\langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle_U = \delta_{ij}$ and $\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle_V = \delta_{ij}$.

Lemma 12. *Let $k - \mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$, and let $(U, \langle -,-\rangle)_U$ and $(V, \langle -,-\rangle_V)$ be right k-vector spaces with hermitian inner product. If $f\colon V \to U$ and $g\colon U \to V$ are adjoint with respect to $\langle -,-\rangle_U$ and $\langle -,-\rangle_V$, then $f\colon V \to U$ is a linear isometry with respect to $\langle -,-\rangle_U$ and $\langle -,-\rangle_V$ if and only if $g \circ f = \mathrm{id}_V$.*

Proof. We find that $f\colon V \to U$ is a linear isometry if and only if

$$\langle (g \circ f)(\boldsymbol{x}), \boldsymbol{y}\rangle_V = \langle \boldsymbol{x}, \boldsymbol{y}\rangle_V$$

for all $\boldsymbol{x}, \boldsymbol{y} \in V$. If $g \circ f = \mathrm{id}_V$, then this is certainly true, and conversely, we find, by setting $\boldsymbol{y} = (g \circ f)(\boldsymbol{x}) - \boldsymbol{x}$, that

$$\langle \boldsymbol{y}, \boldsymbol{y}\rangle_V = \langle (g \circ f)(\boldsymbol{x}) - \boldsymbol{x}, \boldsymbol{y}\rangle_V = \langle (g \circ f)(\boldsymbol{x}), \boldsymbol{y}\rangle_V - \langle \boldsymbol{x}, \boldsymbol{y}\rangle_V = 0,$$

which shows that $g \circ f = \mathrm{id}_V$, because $\langle -,-\rangle_V$ is an inner product. □

Theorem 13. *Let $k = \mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$, and let $(U, \langle -,-\rangle_U)$ and $(V, \langle -,-\rangle_V)$ be finite dimensional right k-vector spaces with hermitian inner products. Let $f\colon V \to U$ be a linear map, and let $A \in M_{m,n}(k)$ be the matrix that represents $f\colon V \to U$ with respect to bases $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_m)$ of U and $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ of V that are orthonormal with respect to the given hermitian inner products. The following (1)–(3) are equivalent.*

(1) *The map $f\colon V \to U$ is a linear isometry.*
(2) *The matrix identity $A^*A = E_n$ holds.*
(3) *The family $(\boldsymbol{a}_1, \dots, \boldsymbol{a}_n)$ of vectors in k^m consisting of the columns of A is orthonormal with respect to the standard hermitian inner product.*

In addition, the following (4)–(6) are equivalent.

(4) *The map $f\colon V \to U$ is an isometric isomorphism.*
(5) *The matrix A is invertible and $A^{-1} = A^*$.*
(6) *The family $(\boldsymbol{a}_1, \dots, \boldsymbol{a}_n)$ of columns of A is basis of k^m that is orthonormal with respect to the standard hermitian inner product.*

Proof. By Proposition 11, the adjoint map $g\colon U \to V$ is represented by the adjoint matrix $A^* \in M_{n,m}(k)$ with respect to the given bases, so the equilvalence of (1) and (2) follows from Lemma 12. The (i,j)th entry in A^*A is $\boldsymbol{a}_i^*\boldsymbol{a}_j$, which, by definition, is the standard hermitian inner product of $\boldsymbol{a}_i, \boldsymbol{a}_j \in k^m$, from which the equivalence of (2) and (3) follows. To prove the equivalence of (4) and (5), we note that $f\colon V \to U$ is an isomorphism if and only if A is invertible, in which case

$$A^{-1} = (A^*A)A^{-1} = A^*(AA^{-1}) = A^*.$$

Finally, the equivalence of (5) and (6) uses that an $n \times n$-matrix invertible if and only if the family consisting of its columns is a basis of k^n. □

Corollary 14. *Let $k = \mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$, and let $(V, \langle -,-\rangle)$ be a finite dimensional right k-vector space with hermitian inner product, and let $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ be a basis of V that is orthonormal with respect to $\langle -,-\rangle$. Let $f\colon V \to V$ be an endomorphism, and let $A \in M_n(k)$ be the matrix that represents $f\colon V \to V$ with $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$.*

(1) *The endomorphism $f\colon V \to V$ is an isometry with respect to $\langle -,-\rangle$ if and only if $A^*A = E_n$. If so, then A is invertible and $A^{-1} = A^*$.*
(2) *The endomorphism $f\colon V \to V$ is selfadjoint[3] with respect to $\langle -,-\rangle$ if and only if $A^* = A$.*

[3] This means that $f\colon V \to V$ and its adjoint $g\colon V \to V$ with respect to $\langle -,-\rangle$ are equal.

Proof. The statement (1) follows from Theorem 13 and from the fact that a square matrix that has a right inverse is invertible. This fact, in turn, is a consequence of Gauss elimination. The statement (2) follows from Proposition 11. □

Remark 15. A matrix $P \in \mathrm{GL}_n(k)$ such that $P^* = P^{-1}$ is said to be orthogonal, if $k = \mathbb{R}$, unitary, if $k = \mathbb{C}$, and quaternionic unitary, if $k = \mathbb{H}$. A matrix $A \in M_n(k)$ such that $A^* = A$ is said to be symmetric, if $k = \mathbb{R}$, hermitian, if $k = \mathbb{C}$, and quaternionic hermitian, if $k = \mathbb{H}$.

We now define the classical groups. The subgroups

$$\begin{aligned} O(n) &= \{Q \in \mathrm{GL}_n(\mathbb{R}) \mid Q^* = Q^{-1}\} \subset \mathrm{GL}_n(\mathbb{R}) \\ U(n) &= \{U \in \mathrm{GL}_n(\mathbb{C}) \mid U^* = U^{-1}\} \subset \mathrm{GL}_n(\mathbb{C}) \\ \mathrm{Sp}(n) &= \{S \in \mathrm{GL}_n(\mathbb{H}) \mid S^* = S^{-1}\} \subset \mathrm{GL}_n(\mathbb{H}) \end{aligned}$$

are called the orthogonal group, the unitary group, and the compact sympletic group. They are topological groups with respect to the subspace topology from the metric topology on $M_n(k)$, and they are all compact. In particular, we have

$$\begin{aligned} O(1) &= \{x \in \mathrm{GL}_1(\mathbb{R}) \mid x^*x = 1\} \subset \mathrm{GL}_1(\mathbb{R}) \\ U(1) &= \{z \in \mathrm{GL}_1(\mathbb{C}) \mid z^*z = 1\} \subset \mathrm{GL}_1(\mathbb{C}) \\ \mathrm{Sp}(1) &= \{q \in \mathrm{GL}_1(\mathbb{H}) \mid q^*q = 1\} \subset \mathrm{GL}_1(\mathbb{H}), \end{aligned}$$

so as topological spaces, these are the unit 0-sphere S^0, the unit 1-sphere S^1, and the unit 3-sphere S^3, respectively. If $A = Q \in O(n)$ or $A = U \in U(n)$, then

$$\det(A)^* = \det(A^*) = \det(A^{-1}) = \det(A)^{-1}$$

so $\det(Q) \in O(1)$ and $\det(U) \in U(1)$. The subgroups

$$\begin{aligned} SO(n) &= \{Q \in O(n) \mid \det(Q) = 1\} \subset O(n) \\ SU(n) &= \{U \in U(n) \mid \det(U) = 1\} \subset U(n) \end{aligned}$$

are called the special orthogonal group and the special unitary group, respectively. There is no useful determinant of quaternionic square matrices, because the division ring $\mathbb{H}$ is noncommutative.[4]

We embed $\mathbb{C}$ in $\mathbb{H}$ as the subfield $L \subset \mathbb{H}$ consisting of all quaternions of the form $q = a + ib$. The subfield $L \subset \mathbb{H}$ is a maximal subfield, and if also $L' \subset \mathbb{H}$ is a maximal subfield, then there exists $q \in \mathbb{H}$ such that $L' = qLq^{-1}$. So every maximal subfield of $\mathbb{H}$ is isomorphic to $\mathbb{C}$, but the embedding of $\mathbb{C}$ as a maximal subfield in $\mathbb{H}$ is only well-defined, up to conjugation. Left multiplication by $q = z_1 + jz_2 \in \mathbb{H}$ defines an L-linear map $\lambda(q) \colon \mathbb{H} \to \mathbb{H}$, and hence, a ring homomorphism

$$\mathbb{H} \xrightarrow{\lambda} \mathrm{End}_L(\mathbb{H}).$$

Since $\mathbb{H}$ is a division ring, the kernel of λ is either $\{0\}$ or $\mathbb{H}$, and since $\lambda(1) = \mathrm{id}_{\mathbb{H}} \neq 0$, we conclude that the kernel is $\{0\}$. Let us choose the basis $(1, j)$ of $\mathbb{H}$ as a right L-vector space. This defines a ring isomorphism

$$\mathrm{End}_L(\mathbb{H}) \xrightarrow{\mu} M_2(L)$$

[4] The best one has is the Dieudonné determinant in $K_1(\mathbb{H}) = (\mathbb{R}_{>0}, \cdot)$.

that to an L-linear map $f\colon \mathbb{H} \to \mathbb{H}$ assigns the matrix $A = \mu(f) \in M_2(L)$ that represents $f\colon \mathbb{H} \to \mathbb{H}$ with respect to the basis $(1, j)$. The calculation

$$\begin{aligned} q \cdot 1 &= (z_1 + jz_2) \cdot 1 = 1 \cdot z_1 + j \cdot z_2 \\ q \cdot j &= (z_1 + jz_2) \cdot j = j \cdot z_1^* - 1 \cdot z_2^* \end{aligned}$$

shows that the composite ring homomorphism

$$\mathbb{H} \xrightarrow{f=\mu\circ\lambda} M_2(L)$$

takes the quaternion $q = z_1 + jz_2$ to the matrix

$$f(q) = \begin{pmatrix} z_1 & -z_2^* \\ z_2 & z_1^* \end{pmatrix}.$$

A map between topological groups is an isomorphism if and only if it is both an isomorphism of groups and a homeomorphism of topological spaces.

Proposition 16. *The ring homomorphism $f\colon \mathbb{H} \to M_2(L)$ induces an isomorphism of topological groups $h\colon \mathrm{Sp}(1) \to SU(2)$.*

Proof. We have $q^* = (z_1 + jz_2)^* = z_1^* + z_2^* j^* = z_1^* - jz_2$. Therefore,

$$q^*q = (z_1^* - jz_2)(z_1 + jz_2) = z_1^* z_1 + jz_1z_2 - jz_2z_1 + z_2^*z_2 = z_1^*z_1 + z_2^*z_2,$$

which shows that $q \in \mathrm{Sp}(1)$ if and only if $f(q) \in SU(2)$. So the ring homomorphism $f\colon \mathbb{H} \to M_2(K)$ restricts to a group homomorphism $h\colon \mathrm{Sp}(1) \to SU(2)$, which is continuous because $f\colon \mathbb{H} \to M_2(K)$ is continuous. We wish to prove that h is both an isomorphism of groups and a homeomorphism of spaces, and to do so, it suffices to show that h is a bijection. Indeed, the inverse map of a bijective group homomorphism is automatically a group homomorphism, and the inverse map of a continuous bijection from a compact space such as $\mathrm{Sp}(1)$ to a Hausdorff space such as $SU(2)$ is automatically continuous. Now, the map h is injective, because the map f is injective, and the map h is surjective because, if

$$U = \begin{pmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{pmatrix} \in SU(2),$$

then $U = f(q)$ with $q = z_{11} + jz_{21}$. This completes the proof. □

Let $k = \mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$. We define the Hilbert–Schmidt norm of $A \in M_n(k)$ by

$$\|A\| = \sqrt{\mathrm{tr}(A^*A)}\,.$$

It satisfies $\|A + B\| \leq \|A\| + \|B\|$ and $\|AB\| \leq \|A\|\|B\|$ for all $A, B \in M_n(k)$, so in particular, the exponential series

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

converges absolutely. If $[A, B] = AB - BA = 0$, then

$$\exp(A + B) = \exp(A)\exp(B),$$

but in general the left-hand side and the right-hand side are different.[5] Hence, the matrix $\exp(A)$ is invertible with inverse $\exp(-A)$, so we get a map

$$M_n(k) \xrightarrow{\exp} \mathrm{GL}_n(k).$$

Locally on $M_n(k)$, this map is a diffeomorphism. For it is a smooth map (considered as map between open subsets of $\mathbb{R}^m$) with derivative $\mathrm{id}\colon M_n(k) \to M_n(k)$, so the inverse function theorem shows that it is a diffeomorphism locally on $M_n(k)$.

If $G \subset \mathrm{GL}_n(k)$ is one of the classical groups, then we define its Lie algebra to the be the subset $\mathfrak{g} \subset M_n(k)$ consisting of all matrices A such that $\exp(tA) \in G$, for all $t \in \mathbb{R}$. It is a real subspace of $M_n(k)$.

Proposition 17. *The Lie algebras of the classical groups are given by*

$$\begin{aligned}
\mathfrak{o}(n) &= \{ A \in M_n(\mathbb{R}) \mid A^* + A = 0 \} \\
\mathfrak{u}(n) &= \{ A \in M_n(\mathbb{C}) \mid A^* + A = 0 \} \\
\mathfrak{sp}(n) &= \{ A \in M_n(\mathbb{H}) \mid A^* + A = 0 \} \\
\mathfrak{so}(n) &= \{ A \in \mathfrak{o}(n) \mid \mathrm{tr}(A) = 0 \} \\
\mathfrak{su}(n) &= \{ A \in \mathfrak{u}(n) \mid \mathrm{tr}(A) = 0 \}
\end{aligned}$$

Proof. We prove the statements for $\mathfrak{u}(n)$ and $\mathfrak{su}(n)$; the proofs in the remaining cases are analogous. If $A \in \mathfrak{u}(n)$, then for all $t \in \mathbb{R}$, we have

$$\exp(tA^*) = \exp(tA)^* = \exp(tA)^{-1} = \exp(-tA),$$

and since the exponential map is a local diffeomorphism, this implies that $A^* = -A$. Similarly, if $A \in \mathfrak{su}(n)$, then we have in addition that for all $t \in \mathbb{R}$,

$$\exp(nt\,\mathrm{tr}(A)) = \exp(\mathrm{tr}(tA)) = \det(\exp(tA)) = 1.$$

Since the exponential map is a local diffeomorphism, this implies that $\mathrm{tr}(A) = 0$. □

Example 18. The Lie algebra $\mathfrak{sp}(1) \subset \mathbb{H}$ is the 3-dimensional real subspace of purely imaginary quaternions. One can show that $\exp\colon \mathfrak{sp}(1) \to \mathrm{Sp}(1)$ is given by

$$\exp(v) = \cos|v| + \frac{v}{|v|}\sin|v|,$$

where $|v| = \sqrt{v^*v}$.

Lemma 19. *Let $G \subset \mathrm{GL}_n(k)$ be one of the classical groups, and let $\mathfrak{g} \subset M_n(k)$ be its Lie algebra. If $g \in G$ and $A \in \mathfrak{g}$, then $gAg^{-1} \in \mathfrak{g}$.*

Proof. Indeed, for all $t \in \mathbb{R}$, we have

$$\exp(tgAg^{-1}) = \exp(gtAg^{-1}) = g\exp(tA)g^{-1},$$

so if $\exp(tA) \in G$ and $g \in G$, then also $\exp(tgAg^{-1}) \in G$. □

Definition 20. The adjoint representation of the classical group $G \subset \mathrm{GL}_n(k)$ on its Lie algebra $\mathfrak{g} \subset M_n(k)$ is the real representation

$$G \xrightarrow{\mathrm{Ad}} \mathrm{GL}(\mathfrak{g})$$

defined by $\mathrm{Ad}(g)(A) = gAg^{-1}$.

[5] The difference is given by the Baker–Campbell–Hausdorff formula.

We consider the adjoint representation

$$\mathrm{Sp}(1) \xrightarrow{\mathrm{Ad}} \mathrm{GL}(\mathfrak{sp}(1))$$

of the compact symplectic group $\mathrm{Sp}(1)$ on its Lie algebra $\mathfrak{sp}(1)$ of purely imaginary quaternions, or equivalently, the adjoint representation

$$SU(2) \xrightarrow{\mathrm{Ad}} \mathrm{GL}(\mathfrak{su}(2))$$

of the special unitary group $SU(2)$ on its Lie algebra $\mathfrak{su}(2)$ given by the real vector space of complex 2×2-matrices that are skew-hermitian and traceless. The map that to $v \in \mathfrak{sp}(1)$ assigns $|v| = \sqrt{v^*v}$ is a norm on the real vector space $\mathfrak{sp}(1)$, and it determines a real inner product $\langle -,- \rangle$ on $\mathfrak{sp}(1)$ given by[6]

$$\langle v, w \rangle = \frac{1}{2}(|v+w|^2 - |v|^2 - |w|^2).$$

We claim that the adjoint representation takes values in the subgroup

$$SO(\mathfrak{sp}(1)) \subset \mathrm{GL}(\mathfrak{sp}(1))$$

of linear isometries with respect to $\langle -,- \rangle$ that have determinant 1. To see this, we first note that since $\mathrm{Ad}(q)(v) = qvq^{-1} = qvq^*$, we have

$$(qvq^*)^*qvq^* = qv^*q^*qvq^* = qv^*vq^* = v^*v,$$

where the last identity holds, because v^*v is an element of the center $\mathbb{R}$ of $\mathbb{H}$. This shows that $\mathrm{Ad}(q)$ is a linear isometry with respect to $\langle -,- \rangle$. Therefore, the adjoint representation induces a group homomorphism

$$\mathrm{Sp}(1) \xrightarrow{\mathrm{Ad}} O(\mathfrak{sp}(1))$$

to the subgroup $O(\mathfrak{sp}(1)) \subset \mathrm{GL}(\mathfrak{sp}(1))$ of linear isometric isomorphisms. It is clearly a continuous map, and since $\mathrm{Sp}(1)$ is connected, its image is fully contained in one of the two components of $O(\mathfrak{sp}(1))$. But $\mathrm{Ad}(1)$ is the identity map of $\mathfrak{sp}(1)$, which has determinant 1, so we conclude that $\mathrm{Ad}(q)$ takes values in $SO(\mathfrak{sp}(1))$ as claimed.

Theorem 21. *The adjoint representation induces a group homomorphism*

$$\mathrm{Sp}(1) \xrightarrow{\mathrm{Ad}} SO(\mathfrak{sp}(1))$$

which is surjective with kernel $\{\pm 1\}$.

We first prove two lemmas. If V is a real vector space with norm $\|-\|$, then we write $S(V) = \{v \in V \mid \|v\| = 1\} \subset V$ for the unit sphere.

Lemma 22. *If $H \subset SO(\mathfrak{sp}(1))$ is a subgroup such that the restriction to H of the standard action by $SO(\mathfrak{sp}(1))$ on $S(\mathfrak{sp}(1))$ is transitive and such that there exists $u \in S(\mathfrak{sp}(1))$ with $SO(\mathfrak{sp}(1))_u \subset H$, then $H = SO(\mathfrak{sp}(1))$.*

Proof. Given $g \in SO(\mathfrak{sp}(1))$, we can find $h \in H$ such that $h \cdot u = g \cdot u$. But then $h^{-1}g \cdot u = u$, so $h^{-1}g \in SO(\mathfrak{sp}(1))_u \subset H$, and hence, $g = h \cdot h^{-1}g \in H$. □

Lemma 23. *For all $v \in S(\mathfrak{sp}(1))$, there exists $g \in \mathrm{Sp}(1)$ such that*

$$\mathrm{Ad}(g)(v) = i.$$

[6] Writing $v = ib + jc + kd$, we have $|v|^2 = b^2 + c^2 + d^2$.

Proof. We will use the spectral theorem for normal operators on finite dimensional complex vector spaces. The ring homomorphism $f\colon \mathbb{H} \to M_2(\mathbb{C})$ that we considered above induces isomorphisms $h\colon \mathrm{Sp}(1) \to SU(2)$ and $h'\colon \mathfrak{sp}(1) \to \mathfrak{su}(2)$. It maps $v \in \mathfrak{sp}(1)$ to $X = h'(v) \in \mathfrak{su}(2)$ with $\det(X) = v^*v = 1$. Since the matrix X is skew-hermitian, it is normal.[7] Therefore, by the spectral theorem for normal matrices, there exists $P \in U(2)$ such that $PXP^{-1} = \mathrm{diag}(\lambda_1, \lambda_2)$, where λ_1 and λ_2 are the eigenvalues of X. Since X is skew-hermitian and $\det(X) = 1$, one shows that $\lambda_1 = i$ and $\lambda_2 = -i$. So we have $P \in U(2)$ with

$$PXP^{-1} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = h'(i).$$

Since $P \in U(2)$, we have $\det(P) \in U(1)$, so we can choose $z \in U(1)$ such that $z^2 = \det(P)$. Then $U = z^{-1}P \in SU(2)$, and we still have $UXU^{-1} = h'(i)$. Hence, if $g \in \mathrm{Sp}(1)$ is the unique element with $h(g) = U$, then $\mathrm{Ad}(g)(v) = i$. □

Proof of Theorem 21. We apply Lemma 22 to the subgroup $H \subset SO(\mathfrak{sp}(1))$ given by the image of $\mathrm{Ad}\colon \mathrm{Sp}(1) \to SO(\mathfrak{sp}(1))$. Lemma 23 shows that H acts transitively on $S(\mathfrak{sp}(1))$, and we proceed to show that for $SO(\mathfrak{sp}(1))_i \subset H$. The matrix that represents a general element of the isotropy subgroup $SO(\mathfrak{sp}(1))_i$ with respect to the basis (i, j, k) of $\mathfrak{sp}(1)$ has the form

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

for some $\theta \in \mathbb{R}$. We calculate that the matrix that represent $\mathrm{Ad}(e^{it})$ with respect to the basis (i, j, k) of $\mathfrak{sp}(1)$ is given by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos 2t & -\sin 2t \\ 0 & \sin 2t & \cos 2t \end{pmatrix}.$$

This shows that $SO(\mathfrak{sp}(1))_i \subset H$, and therefore, we conclude from Lemma 22 that $H = SO(\mathfrak{sp}(1))$ as stated.

Finally, if $\mathrm{Ad}(g) = \mathrm{id}$, then, in particular, $\mathrm{Ad}(g) \in SO(\mathfrak{sp}(1))_i$, so $g = e^{it}$. But if $\mathrm{Ad}(e^{it}) = \mathrm{id}$, then $e^{it} = \pm 1$, so $\ker(\mathrm{Ad}) = \{\pm 1\}$ as stated. □

Corollary 24. *The map induced by the adjoint representation,*

$$\mathrm{Sp}(1)/\{\pm 1\} \xrightarrow{\overline{\mathrm{Ad}}} SO(\mathfrak{sp}(1)),$$

is an isomorphism of topological groups.

Proof. We have not explicitly specified the topologies on these groups before, so we do that now. We have identified both $\mathrm{Sp}(1)$ and $SO(\mathfrak{sp}(1))$ with subsets of $M_2(\mathbb{C})$, and we give both the respective subspace topologies induced from the metric topology on $M_2(\mathbb{C})$. Finally, we give $\mathrm{Sp}(1)/\{\pm 1\}$ the quotient topology induced from the topology on $\mathrm{Sp}(1)$. As a topological space, $\mathrm{Sp}(1)/\{\pm 1\}$ is compact, because $\mathrm{Sp}(1)$ is compact, and $SO(\mathfrak{sp}(1))$ is Hausdorff, because the metric topology on $M_2(\mathbb{C})$ is Hausdorff. So it suffices to show that $\overline{\mathrm{Ad}}$ is a group homomorphism and a continuous bijection. Theorem 21 shows that it is a group isomorphism, so it

[7] Indeed, $X^*X = (-X)X = X(-X) = XX^*$.

only remains to show that the map $\overline{\mathrm{Ad}}$ is continuous. By the universal property of the quotient topology, the map $\overline{\mathrm{Ad}}$ is continuous if and only if the map Ad is continuous. And by the universal property of the subspace topology, the map Ad is continuous if and only if the map

$$\mathrm{Sp}(1) \xrightarrow{\widetilde{\mathrm{Ad}}} \mathrm{End}_{\mathbb{R}}(M_2(\mathbb{C}))$$

defined by $\widetilde{\mathrm{Ad}}(g)(X) = h(g)Xh(g)^{-1}$ is continuous. This, in turn, follows from the definition of matrix multiplication and from Cramer's formula for the inverse of a matrix. □

If G is a topological group, then we write $\mathrm{Rep}_{\mathbb{C}}(G)$ for the category, whose objects are complex representations (V, π) of G such that $\pi\colon G \to \mathrm{GL}(V)$ is continuous, and whose morphisms are intertwining $\mathbb{C}$-linear maps. Restriction along the continuous group homomorphism $\mathrm{Ad}\colon \mathrm{Sp}(1) \to SO(\mathfrak{sp}(1))$ defines a functor

$$\mathrm{Rep}_{\mathbb{C}}(SO(\mathfrak{sp}(1))) \xrightarrow{\mathrm{Ad}^*} \mathrm{Rep}_{\mathbb{C}}(\mathrm{Sp}(1)),$$

and Corollary 24 shows that this functor is a fully faithful embedding and that its essential image are the continuous complex representations (V, π) of $\mathrm{Sp}(1)$ with the property that $\pi(-1) = \mathrm{id}_V$.

Another consequence of Corollary 24 is that, as a topological space, $SO(3)$ is homeomorphic to the real projective space $\mathbb{P}^3(\mathbb{R})$. Indeed, as a topological space $\mathrm{Sp}(1)$ is homeomorphic to S^3, and the action of the subgroup $\{\pm 1\} \subset \mathrm{Sp}(1)$ by left multiplication is free.