https://doi.org/10.1090//stml/079/04

Chapter 3

The orthogonal groups

In this chapter, we define and study what are probably the most important subgroups of the general linear groups. These are denoted $O(n)$, $SO(n)$, $U(n)$, $SU(n)$ and $Sp(n)$. In particular, the group $SO(3)$, which was previously described as the "positions of a globe," now receives a more rigorous definition. We will continue to study these groups throughout the remainder of the book.

1. The standard inner product on $\mathbb{K}^n$

The conjugate and norm of an element $q \in \mathbb{K}$ are defined as:

(1) If $q \in \mathbb{R}$, then $\overline{q} = q$ and $|q|$ means the absolute value of q.

(2) If $q = a + b\mathbf{i} \in \mathbb{C}$, then $\overline{q} = a - b\mathbf{i}$ and $|q| = \sqrt{a^2 + b^2}$.

(3) If $q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \in \mathbb{H}$, then $\overline{q} = a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}$ and $|q| = \sqrt{a^2 + b^2 + c^2 + d^2}$.

In all cases, it is a quick calculation to verify that for $q, q_1, q_2 \in \mathbb{K}$:

$$\overline{q_1 \cdot q_2} = \overline{q}_2 \cdot \overline{q}_1. \tag{3.1}$$

$$q \cdot \overline{q} = \overline{q} \cdot q = |q|^2. \tag{3.2}$$

These two equalities together imply that:

$$|q_1 \cdot q_2| = |q_1| \cdot |q_2|. \tag{3.3}$$

Definition 3.1. *The* ***standard inner product*** *on* $\mathbb{K}^n$ *is the function from* $\mathbb{K}^n \times \mathbb{K}^n$ *to* $\mathbb{K}$ *defined by:*

$$\langle (x_1, x_2, ..., x_n), (y_1, y_2, ..., y_n) \rangle_{\mathbb{K}} = x_1 \cdot \overline{y}_1 + x_2 \cdot \overline{y}_2 + \cdots + x_n \cdot \overline{y}_n.$$

It follows from Equation 3.2 that for all $X \in \mathbb{K}^n$, $\langle X, X \rangle_{\mathbb{K}}$ is a real number that is ≥ 0 and equal to zero only when $X = (0, ..., 0)$. This allows us to define:

Definition 3.2. *The* ***standard norm*** *on* $\mathbb{K}^n$ *is the function from* $\mathbb{K}^n$ *to the nonnegative real numbers defined by:*

$$|X|_{\mathbb{K}} = \sqrt{\langle X, X \rangle_{\mathbb{K}}}.$$

We will omit the $\mathbb{K}$-subscripts whenever there is no ambiguity.

Proposition 3.3. *For all* $X, Y, Z \in \mathbb{K}^n$ *and* $\lambda \in \mathbb{K}$,

(1) $\langle X, Y + Z \rangle = \langle X, Y \rangle + \langle X, Z \rangle$,
(2) $\langle X + Y, Z \rangle = \langle X, Z \rangle + \langle Y, Z \rangle$,
(3) $\langle \lambda X, Y \rangle = \lambda \langle X, Y \rangle$ *and* $\langle X, \lambda Y \rangle = \langle X, Y \rangle \overline{\lambda}$,
(4) $\overline{\langle X, Y \rangle} = \langle Y, X \rangle$.

Definition 3.4.

- *Vectors* $X, Y \in \mathbb{K}^n$ *are called* ***orthogonal*** *if* $\langle X, Y \rangle = 0$.
- *A basis* $\{X_1, ..., X_n\}$ *of* $\mathbb{K}^n$ *is called* ***orthonormal*** *if* $\langle X_i, X_j \rangle$ *equals* 1 *when* $i = j$ *and equals zero when* $i \neq j$ *(that is, the vectors have norm* 1 *and are mutually orthogonal).*
- *The* ***standard orthonormal basis*** *of* $\mathbb{K}^n$ *is:*

$$e_1 = (1, 0, ..., 0),\ e_2 = (0, 1, 0, ..., 0), ..., e_n = (0, ..., 0, 1).$$

When $\mathbb{K} = \mathbb{R}$, the standard inner product is the familiar "dot product," described geometrically in terms of the angle θ between $X, Y \in \mathbb{R}^n$:

$$\langle X, Y \rangle_{\mathbb{R}} = |X|_{\mathbb{R}} |Y|_{\mathbb{R}} \cos\theta. \tag{3.4}$$

When $\mathbb{K} = \mathbb{C}$, the standard inner product is also called the **hermitian inner product**. Since the hermitian inner product of two

vectors $X, Y \in \mathbb{C}^n$ is a complex number, we should separately interpret the geometric meanings of its real and imaginary parts. The cleanest such interpretation is in terms of the identification

$$f = f_n : \mathbb{C}^n \to \mathbb{R}^{2n}$$

from the previous chapter. It is easy to verify that for all $X, Y \in \mathbb{C}^n$,

$$\langle X, Y\rangle_{\mathbb{C}} = \langle f(X), f(Y)\rangle_{\mathbb{R}} + \mathbf{i}\langle f(X), f(\mathbf{i}Y)\rangle_{\mathbb{R}}, \tag{3.5}$$

$$|X|_{\mathbb{C}} = |f(X)|_{\mathbb{R}}. \tag{3.6}$$

Thus, if $X, Y \in \mathbb{C}^n$ are orthogonal, then two things are true:

$$\langle f(X), f(Y)\rangle_{\mathbb{R}} = 0 \text{ and } \langle f(X), f(\mathbf{i}Y)\rangle_{\mathbb{R}} = 0.$$

This observation leads to:

Proposition 3.5. *$\{X_1, ..., X_n\} \in \mathbb{C}^n$ is an orthonormal basis if and only if $\{f(X_1), f(\mathbf{i}X_1), ..., f(X_n), f(\mathbf{i}X_n)\}$ is an orthonormal basis of $\mathbb{R}^{2n}$.*

When $\mathbb{K} = \mathbb{H}$, the standard inner product is also called the **symplectic inner product**. For $X, Y \in \mathbb{H}^n$, the 1, $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ components of $\langle X, Y\rangle_{\mathbb{H}}$ are best interpreted geometrically in terms of the identification $h = f_{2n} \circ g_n : \mathbb{H}^n \to \mathbb{R}^{4n}$, as follows:

$$\begin{aligned}\langle X, Y\rangle_{\mathbb{H}} &= \langle h(X), h(Y)\rangle_{\mathbb{R}} + \mathbf{i}\langle h(X), h(\mathbf{i}Y)\rangle_{\mathbb{R}} \\ &\quad + \mathbf{j}\langle h(X), h(\mathbf{j}Y)\rangle_{\mathbb{R}} + \mathbf{k}\langle h(X), h(\mathbf{k}Y)\rangle_{\mathbb{R}}. \\ |X|_{\mathbb{H}} &= |h(X)|_{\mathbb{R}}.\end{aligned}$$

Proposition 3.6. *$\{X_1, ..., X_n\} \in \mathbb{H}^n$ is an orthonormal basis if and only if the following is an orthonormal basis of $\mathbb{R}^{4n}$:*

$$\{h(X_1), h(\mathbf{i}X_1), h(\mathbf{j}X_1), h(\mathbf{k}X_1), ..., h(X_n), h(\mathbf{i}X_n), h(\mathbf{j}X_n), h(\mathbf{k}X_n)\}.$$

The following inequality follows from Equation 3.4 when $\mathbb{K} = \mathbb{R}$:

Proposition 3.7 (Schwarz inequality)**.** *For all $X, Y \in \mathbb{K}^n$,*

$$|\langle X, Y\rangle| \leq |X| \cdot |Y|.$$

Proof. Let $X, Y \in \mathbb{K}^n$. Let $\alpha = \langle X, Y\rangle$. Assume that $X \neq 0$ (otherwise the proposition is trivial). For all $\lambda \in \mathbb{K}$, we have:

$$\begin{aligned} 0 &\leq |\lambda X + Y|^2 = \langle \lambda X + Y, \lambda X + Y\rangle \\ &= \lambda\langle X, X\rangle\overline{\lambda} + \lambda\langle X, Y\rangle + \langle Y, X\rangle\overline{\lambda} + \langle Y, Y\rangle \\ &= |\lambda|^2|X|^2 + \lambda\langle X, Y\rangle + \overline{\lambda\langle X, Y\rangle} + |Y|^2 \\ &= |\lambda|^2|X|^2 + 2\mathrm{Re}(\lambda\alpha) + |Y|^2. \end{aligned}$$

Choosing $\lambda = -\overline{\alpha}/|X|^2$ gives:

$$0 \leq |\alpha|^2/|X|^2 - 2|\alpha|^2/|X|^2 + |Y|^2,$$

which proves that $|\alpha| \leq |X| \cdot |Y|$ as desired. □

2. Several characterizations of the orthogonal groups

Definition 3.8. *The **orthogonal group over** $\mathbb{K}$,*

$$\mathcal{O}_n(\mathbb{K}) = \{A \in GL_n(\mathbb{K}) \mid \langle X \cdot A, Y \cdot A\rangle = \langle X, Y\rangle \text{ for all } X, Y \in \mathbb{K}^n\},$$

*... is denoted $O(n)$ and called the **orthogonal group** if $\mathbb{K} = \mathbb{R}$.*

*... is denoted $U(n)$ and called the **unitary group** if $\mathbb{K} = \mathbb{C}$.*

*... is denoted $Sp(n)$ and called the **symplectic group** if $\mathbb{K} = \mathbb{H}$.*

It is straightforward to see that $\mathcal{O}_n(\mathbb{K})$ is a subgroup of $GL_n(\mathbb{K})$. Its elements are called **orthogonal**, **unitary** or **symplectic** matrices. To describe their form, it is useful to denote the **conjugate-transpose** of $A \in M_n(\mathbb{K})$ as $A^* = (\overline{A})^T$, where $\overline{A}$ means the matrix obtained by conjugating all of the entries of A.

Proposition 3.9. *For $A \in GL_n(\mathbb{K})$ the following are equivalent.*

(1) $A \in \mathcal{O}_n(\mathbb{K})$.

(2) *R_A preserves orthonormal bases; i.e., if $\{X_1, ..., X_n\}$ is an orthonormal basis of $\mathbb{K}^n$, then so is $\{R_A(X_1), ..., R_A(X_n)\}$.*

(3) *The rows of A form an orthonormal basis of $\mathbb{K}^n$.*

(4) $A \cdot A^* = I$.

Proof. (1) $\Longrightarrow$ (2) is obvious. (2) $\Longrightarrow$ (3) because the rows of A equal $\{R_A(e_1), ..., R_A(e_n)\}$. To see that (3) $\Longleftrightarrow$ (4), notice that:

$$\begin{aligned}(A \cdot A^*)_{ij} &= (\text{row } i \text{ of } A) \cdot (\text{column } j \text{ of } A^*) \\ &= (\text{row } i \text{ of } A) \cdot (\text{row } j \text{ of } \overline{A})^T \\ &= \langle (\text{row } i \text{ of } A), (\text{row } j \text{ of } A) \rangle.\end{aligned}$$

Finally, we prove that (3) $\Longrightarrow$ (1). If the rows of A are orthonormal, then for all $X = (x_1, ..., x_n), Y = (y_1, ..., y_n) \in \mathbb{K}^n$,

$$\begin{aligned}&\langle R_A(X), R_A(Y) \rangle \\ &\quad = \left\langle \sum_{l=1}^{n} x_l (\text{row } l \text{ of } A), \sum_{s=1}^{n} y_s (\text{row } s \text{ of } A) \right\rangle \\ &\quad = \sum_{l,s=1}^{n} x_l \langle (\text{row } l \text{ of } A), (\text{row } s \text{ of } A) \rangle \overline{y}_s \\ &\quad = x_1\overline{y}_1 + \cdots + x_n\overline{y}_n = \langle X, Y \rangle.\end{aligned}$$

□

Geometrically, $O(n)$ is the group of matrices A for which the linear transformation $R_A : \mathbb{R}^n \to \mathbb{R}^n$ preserves dot products of vectors, and hence also norms of vectors. Such transformations should be visualized as "rigid motions" of $\mathbb{R}^n$ (we will be more precise about this in Section 5). The geometric meanings of $U(n)$ and $Sp(n)$ are best described in terms $O(n)$ by considering the homomorphisms from the previous chapter.

Proposition 3.10.

(1) $\rho_n(U(n)) = O(2n) \cap \rho_n(GL_n(\mathbb{C}))$.
(2) $\Psi_n(Sp(n)) = U(2n) \cap \Psi_n(GL_n(\mathbb{H}))$.
(3) $(\rho_{2n} \circ \Psi_n)(Sp(n)) = O(4n) \cap (\rho_{2n} \circ \Psi_n)(GL_n(\mathbb{H}))$.

Since $U(n)$ is isomorphic to its image, $\rho_n(U(n))$, part (1) says that $U(n)$ is isomorphic to the group of complex-linear real orthogonal matrices. In other words, $U(n)$ is isomorphic to the group of rigid motions of $\mathbb{R}^{2n}$ that preserve the standard complex structure. Similarly, part (3) says that $Sp(n)$ is isomorphic to the group of quaternionic-linear real orthogonal matrices.

Proof. We prove only (1), since (2) is similar and (3) follows from (1) and (2). The most straightforward idea is to use Equation 3.5. But a quicker approach is to first notice that for all $A \in M_n(\mathbb{C})$,

$$\rho_n(A^*) = \rho_n(A)^*.$$

If $A \in GL_n(\mathbb{C})$, then $\rho_n(A) \cdot \rho_n(A)^* = \rho_n(A) \cdot \rho_n(A^*) = \rho_n(A \cdot A^*)$, which shows that $A \in U(n)$ if and only if $\rho_n(A) \in O(2n)$. □

We said that $\mathcal{O}_n(\mathbb{K})$ is the group of matrices A for which R_A preserves inner products of vectors, and hence also norms of vectors. The next result says that if R_A preserves norms, then it automatically preserves inner products.

Proposition 3.11.

$$\mathcal{O}_n(\mathbb{K}) = \{A \in GL_n(\mathbb{K}) \mid |R_A(X)| = |X| \text{ for all } X \in \mathbb{K}^n\}.$$

Proof. To prove the case $\mathbb{K} = \mathbb{R}$, we show that the inner product is completely determined by the norm. Solving the equation

$$|X - Y|_{\mathbb{R}}^2 = \langle X - Y, X - Y\rangle_{\mathbb{R}} = \langle X, X\rangle_{\mathbb{R}} + \langle Y, Y\rangle_{\mathbb{R}} - 2\langle X, Y\rangle_{\mathbb{R}}$$

for $\langle X, Y\rangle_{\mathbb{R}}$ gives:

$$\langle X, Y\rangle_{\mathbb{R}} = \frac{1}{2}\left(|X|_{\mathbb{R}}^2 + |Y|_{\mathbb{R}}^2 - |X - Y|_{\mathbb{R}}^2\right). \tag{3.7}$$

From this, it is straightforward to show that if R_A preserves norms, then it also preserves inner products.

The above argument doesn't work for $\mathbb{K} \in \{\mathbb{C}, \mathbb{H}\}$ (why not?). Instead, we prove the case $\mathbb{K} = \mathbb{C}$ as a consequence of the real case. Suppose $A \in GL_n(\mathbb{C})$ is such that $R_A : \mathbb{C}^n \to \mathbb{C}^n$ is norm-preserving. Then $R_{\rho_n(A)} : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ also preserves norms, since for all $X \in \mathbb{C}^n$,

$$|R_{\rho_n(A)}(f_n(X))|_{\mathbb{R}} = |f_n(R_A(X))|_{\mathbb{R}} = |R_A(X)|_{\mathbb{C}} = |X|_{\mathbb{C}} = |f_n(X)|_{\mathbb{R}}.$$

Therefore $\rho_n(A) \in O(2n)$, which using Proposition 3.10 implies that $A \in U(n)$.

The $\mathbb{K} = \mathbb{H}$ case is proven from the real case in a similar fashion. □

3. The special orthogonal groups

In this section, we define important subgroups of the orthogonal groups, beginning with the observation that:

Proposition 3.12. *If $A \in \mathcal{O}_n(\mathbb{K})$, then $|\det(A)| = 1$.*

Proof. Since $A \cdot A^* = I$,

$$1 = \det(A \cdot A^*) = \det(A) \cdot \det(A^*) = \det(A) \cdot \overline{\det(A)} = |\det(A)|^2.$$

We used the fact that $\det(A^*) = \overline{\det(A)}$, which should be verified first for $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. The quaternionic case follows from the complex case because for quaternionic matrices, $\det(A)$ means $\det(\Psi_n(A))$, and $\Psi_n(A^*) = \Psi_n(A)^*$. □

The interpretation of Proposition 3.12 depends on $\mathbb{K}$:

- If $A \in O(n)$, then $\det(A) = \pm 1$.
- If $A \in U(n)$, then $\det(A) = e^{i\theta}$ for some $\theta \in [0, 2\pi)$.
- If $A \in Sp(n)$, then Proposition 2.10 implies $\det(A) = \pm 1$. We will see later that $\det(A) = 1$.

The subgroup

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

is called the **special orthogonal group**. The subgroup

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

is called the **special unitary group**. Both are clearly subgroups of the general linear group and in fact of the **special linear group**:

$$SL_n(\mathbb{K}) = \{A \in GL_n(\mathbb{K}) \mid \det(A) = 1\}.$$

Notice that $SO(n)$ comprises the orthogonal matrices whose determinants are one of two possibilities, while $SU(n)$ comprises the unitary matrices whose determinants are one of a circle's worth of possibilities. We will see later that the relationship of $SO(n)$ to $O(n)$ is very different from $SU(n)$ to $U(n)$.

4. Low dimensional orthogonal groups

In this section, we explicitly describe $\mathcal{O}_n(\mathbb{K})$ for small values of n. First, $O(1) = \{(1), (-1)\}$ and $SO(1) = \{(1)\}$ are isomorphic to the unique groups with 2 and 1 elements respectively.

Next, if $A \in O(2)$, then its two rows form an orthonormal basis of $\mathbb{R}^2$. Its first row is an arbitrary unit-length vector of $\mathbb{R}^2$, which can be written as $(\cos\theta, \sin\theta)$ for some θ. The second row is unit-length and orthogonal to the first, which leaves two choices: $(-\sin\theta, \cos\theta)$ or $(\sin\theta, -\cos\theta)$. For the first choice, $\det(A) = 1$, and for the second, $\det(A) = -1$. So we learn:

$$SO(2) = \left\{ \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \middle| \theta \in [0, 2\pi) \right\}, \tag{3.8}$$

$$O(2) = SO(2) \cup \left\{ \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \middle| \theta \in [0, 2\pi) \right\}.$$

$SO(2)$ is identified with the set of points on a circle; its group operation is addition of angles. $O(2)$ is a disjoint union of two circles. It is interesting that the disjoint union of two circles has a group operation.

Next, $SU(1) = \{(1)\}$ and $U(1) = \{(e^{i\theta}) \mid \theta \in [0, 2\pi)\}$, which is isomorphic to the circle-group $SO(2)$.

Next, $Sp(1) = \{(a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}) \mid a^2 + b^2 + c^2 + d^2 = 1\}$ is the **group of unit-length quaternions**, which is naturally identified with the three-dimensional sphere $S^3 \subset \mathbb{R}^4 \cong \mathbb{H}$. In fact, it follows from Equation 3.3 that the product of two unit-length quaternions is a unit-length quaternion. So we might have mentioned several pages ago the beautiful fact that *quaternionic multiplication provides a group operation on the three-dimensional sphere!* It turns out that S^0, S^1 and S^3 are the only spheres which are also groups.

We conclude this section by showing that $SU(2)$ is isomorphic to $Sp(1)$, and thus in some sense also has the shape of a three-dimensional sphere.

Proposition 3.13. *$SU(2)$ is isomorphic to $Sp(1)$.*

Proof. First notice that

$$\Psi_1(Sp(1)) = \left\{ \begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix} \mid z, w \in \mathbb{C} \text{ such that } |z|^2 + |w|^2 = 1 \right\}$$

is a subgroup of $U(2)$ by Proposition 3.10, namely, the quaternionic-linear 2-by-2 unitary matrices. Calculating the determinant of such matrices shows that $\Psi_1(Sp(1)) \subset SU(2)$. We wish to prove that $\Psi_1(Sp(1)) = SU(2)$, so that Ψ_1 determines an isomorphism between $Sp(1)$ and $SU(2)$.

Let $A = \begin{pmatrix} z_1 & w_1 \\ w_2 & z_2 \end{pmatrix} \in SU(2)$. An easily verified formula for the inverse of a 2-by-2 matrix is: $A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} z_2 & -w_1 \\ -w_2 & z_1 \end{pmatrix}$. In our case, $\det(A) = 1$, so $\begin{pmatrix} z_2 & -w_1 \\ -w_2 & z_1 \end{pmatrix} = A^{-1} = A^* = \begin{pmatrix} \overline{z}_1 & \overline{w}_2 \\ \overline{w}_1 & \overline{z}_2 \end{pmatrix}$, which tells us that $z_2 = \overline{z}_1$ and $w_2 = -\overline{w}_1$. It now follows that $SU(2) = \Psi_1(Sp(1))$. □

5. Orthogonal matrices and isometries

In this section, we describe $O(n)$ geometrically as the group of isometries of $\mathbb{R}^n$ that fix the origin and we discuss the difference between $SO(3)$ and $O(3)$.

The **distance** between points $X = (x_1, ..., x_n)$ and $Y = (y_1, ..., y_n)$ in $\mathbb{R}^n$ is measured as:

$$\text{dist}(X, Y) = |X - Y| = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}.$$

A function $f : \mathbb{R}^n \to \mathbb{R}^n$ is called an **isometry** if for all $X, Y \in \mathbb{R}^n$, $\text{dist}(f(X), f(Y)) = \text{dist}(X, Y)$.

Proposition 3.14.

(1) *If $A \in O(n)$, then $R_A : \mathbb{R}^n \to \mathbb{R}^n$ is an isometry.*

(2) *If $f : \mathbb{R}^n \to \mathbb{R}^n$ is an isometry with $f(0) = 0$, then $f = R_A$ for some $A \in O(n)$. In particular, f is linear.*

Proof. For part (1), if $A \in O(n)$, then for all $X, Y \in \mathbb{R}^n$,

$$\begin{aligned} \operatorname{dist}(R_A(X), R_A(Y)) &= |R_A(X) - R_A(Y)| = |R_A(X - Y)| \\ &= |X - Y| = \operatorname{dist}(X, Y), \end{aligned}$$

which proves that R_A is an isometry.

For part (2), suppose that $f : \mathbb{R}^n \to \mathbb{R}^n$ is an isometry for which $f(0) = 0$. Equation 3.7 can be re-expressed as a description of the inner product completely in terms of distances:

$$\langle X, Y \rangle = \frac{1}{2} \left(\operatorname{dist}(X, 0)^2 + \operatorname{dist}(Y, 0)^2 - \operatorname{dist}(X, Y)^2 \right).$$

Since f preserves distances and fixes the origin, it is straightforward to show using this that it must also preserve the inner product:

$$\langle f(X), f(Y) \rangle = \langle X, Y \rangle \text{ for all } X, Y \in \mathbb{R}^n.$$

Let A be the matrix whose i^{th} row is $f(e_i)$, so $f(e_i) = R_A(e_i)$ for all $i = 1, ..., n$. Notice that $A \in O(n)$, since its rows are orthonormal. We will prove that $f = R_A$ (and thus that f is linear) by showing that $g = (R_A)^{-1} \circ f$ is the identity function. Notice that g is an isometry with $g(0) = 0$ (so g preserves inner products, as above) and $g(e_i) = e_i$ for all $i = 1, ..., n$. Let $X \in \mathbb{R}^n$. Write $X = \sum a_i e_i$ and $g(X) = \sum b_i e_i$. Then,

$$b_i = \langle g(X), e_i \rangle = \langle g(X), g(e_i) \rangle = \langle X, e_i \rangle = a_i,$$

which proves $g(X) = X$, so g is the identity function. □

$O(n)$ is the group of isometries of $\mathbb{R}^n$ that fix the origin and that therefore map the sphere $S^{n-1} \subset \mathbb{R}^n$ to itself. For example, elements of $O(3)$ represent functions from the "globe" $S^2 \subset \mathbb{R}^3$ to itself. We will see next that elements of $SO(3)$ represent real physical motions of the globe, which justifies our characterization of $SO(3)$ as the group of positions of a globe (Section 1 of Chapter 1).

To understand the difference between $O(3)$ and $SO(3)$, we must discuss the **orientation** of $\mathbb{R}^3$. An ordered orthonormal basis of $\mathbb{R}^3$, like $\{X_1, X_2, X_3\}$, is called **right-handed** if $X_1 \times X_2 = X_3$, where "$\times$" denotes the vector cross product in $\mathbb{R}^3$. Visually, this means that if the fingers of your right hand are curled from X_1 towards X_2, then your thumb will point in the direction of X_3.

Proposition 3.15. *Let $A \in O(3)$. Then $A \in SO(3)$ if and only if the rows of A, $\{R_A(e_1), R_A(e_2), R_A(e_3)\}$, form a right-handed ordered orthonormal basis.*

Proof. Let $R_A(e_1) = (a, b, c)$ and $R_A(e_2) = (d, e, f)$ denote the first two rows of A. The third row is unit-length and orthogonal to both, which leaves two choices:

$$R_A(e_3) = \pm(R_A(e_1) \times R_A(e_2)) = \pm(bf - ce, cd - af, ae - bd).$$

A quick calculation shows that the "+" choice gives $\det(A) > 0$, while the "−" choice gives $\det(A) < 0$. □

Elements of $SO(3)$ correspond to "physically performable motions" of a globe. This statement is imprecise, but in Chapter 9 we give it teeth by proving that every element of $SO(3)$ is a rotation through some angle about some single axis. An element of $O(3)$ with negative determinant turns the globe inside-out. For example, $R_{\mathrm{diag}(-1,-1,-1)}$ maps each point of the globe to its antipode (its negative). This is not a physically performable motion.

6. The isometry group of Euclidean space

It is a straightforward exercise to show that

$$\mathrm{Isom}(\mathbb{R}^n) = \{f : \mathbb{R}^n \to \mathbb{R}^n \mid f \text{ is an isometry}\}$$

is a group under composition of functions. The subgroup of isometries that fix the origin is isomorphic to $O(n)$. An isometry, f, that does not fix the origin is not linear, so cannot equal R_A for any matrix A. In this case, let $V = f(0)$, so the function $X \mapsto f(X) - V$ is an isometry that fixes the origin and therefore equals R_A for some $A \in O(n)$. Therefore, an arbitrary isometry of $\mathbb{R}^n$ has the form

$$f(X) = R_A(X) + V$$

for some $A \in O(n)$ and $V \in \mathbb{R}^n$.

There is a clever trick for representing any isometry of $\mathbb{R}^n$ as a matrix, even ones that do not fix the origin. Graphics programmers use this trick to rotate *and translate* objects on the computer screen via matrices. We first describe the $n = 3$ case.

Let $A \in O(3)$ and $V = (v_1, v_2, v_3) \in \mathbb{R}^3$. We will represent the isometry $f(X) = R_A(X) + V$ by the matrix:

$$F = \begin{pmatrix} A & 0 \\ V & 1 \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & A_{13} & 0 \\ A_{21} & A_{22} & A_{23} & 0 \\ A_{31} & A_{32} & A_{33} & 0 \\ v_1 & v_2 & v_3 & 1 \end{pmatrix} \in GL_4(\mathbb{R}).$$

Let $X = (x_1, x_2, x_3) \in \mathbb{R}^3$. Denote $(X, 1) = (x_1, x_2, x_3, 1) \in \mathbb{R}^4$. Notice that

$$(X, 1) \cdot F = (f(X), 1) \in \mathbb{R}^4.$$

In this way, F represents f.

The composition of two isometries, like the ones represented by $F_1 = \begin{pmatrix} A_1 & 0 \\ V_1 & 1 \end{pmatrix}$ and $F_2 = \begin{pmatrix} A_2 & 0 \\ V_2 & 1 \end{pmatrix}$, is the isometry represented by the product:

$$\begin{pmatrix} A_1 & 0 \\ V_1 & 1 \end{pmatrix} \cdot \begin{pmatrix} A_2 & 0 \\ V_2 & 1 \end{pmatrix} = \begin{pmatrix} A_1 \cdot A_2 & 0 \\ R_{A_2}(V_1) + V_2 & 1 \end{pmatrix}.$$

Matrix multiplication is quite useful here. It allowed us to see immediately that the isometry $X \mapsto R_{A_1}(X) + V_1$ followed by the isometry $X \mapsto R_{A_2}(X) + V_2$ is the isometry $X \mapsto R_{(A_1 \cdot A_2)}(X) + R_{A_2}(V_1) + V_2$.

The above ideas also work for values of n other than 3. We conclude that $\mathrm{Isom}(\mathbb{R}^n)$ is isomorphic to the following subgroup of $GL_{n+1}(\mathbb{R})$:

$$\mathrm{Isom}(\mathbb{R}^n) \cong \left\{ \begin{pmatrix} A & 0 \\ V & 1 \end{pmatrix} \Big| A \in O(n) \text{ and } V \in \mathbb{R}^n \right\}.$$

Notice that the following subgroup of $\mathrm{Isom}(\mathbb{R}^n)$ is isomorphic to $(\mathbb{R}^n, +)$, which denotes $\mathbb{R}^n$ under the group-operation of vector-addition:

$$\mathrm{Trans}(\mathbb{R}^n) = \left\{ \begin{pmatrix} I & 0 \\ V & 1 \end{pmatrix} \Big| V \in \mathbb{R}^n \right\}.$$

This is the group of isometries of $\mathbb{R}^n$ that only translate and do not rotate. It is interesting that $(\mathbb{R}^n, +)$ is isomorphic to a matrix group!

7. Symmetry groups

The **symmetry group** of a subset $X \subset \mathbb{R}^n$ is the group of all isometries of $\mathbb{R}^n$ that carry X onto itself:

Definition 3.16. *$Symm(X) = \{f \in Isom(\mathbb{R}^n) \mid f(X) = X\}$.*

The statement "$f(X) = X$" means that each point of X is sent by f to a (possibly different) point of X.

For example, the symmetry group of the sphere $S^{n-1} \subset \mathbb{R}^n$ equals the group of isometries of $\mathbb{R}^n$ with no translational component, which is isomorphic to the orthogonal group:

$$\mathrm{Symm}(S^{n-1}) = \left\{ \begin{pmatrix} A & 0 \\ V & 1 \end{pmatrix} \Big| A \in O(n),\, V = (0, ..., 0) \right\} \cong O(n).$$

In an abstract algebra course, you probably encountered some important finite symmetry groups. For example, the symmetry group of a regular m-gon (triangle, square, pentagon, hexagon, etc.) centered at the origin in $\mathbb{R}^2$ is called the **dihedral group** of order $2m$, denoted D_m. The elements of D_m with determinant $+1$ are called rotations; they form a subgroup of index 2 that is isomorphic to the cyclic group $\mathbb{Z}_m$, of order m. The elements of D_m with determinant -1 are called flips.

The fact that half of the elements of D_m are rotations illustrates a general principal:

Definition 3.17. *$Symm(X) = Symm^+(X) \cup Symm^-(X)$, where the sets*

$$Symm^{\pm}(X) = \left\{ \begin{pmatrix} A & 0 \\ V & 1 \end{pmatrix} \in Symm(X) \Big| \det(A) = \pm 1 \right\}$$

*are called the **proper** and **improper** symmetry groups of X.*

Proposition 3.18. *For any $X \subset \mathbb{R}^n$, $Symm^+(X) \subset Symm(X)$ is a subgroup with index 1 or 2.*

The proof is left to the reader in Exercise 3.4. An example of a set $Y \subset \mathbb{R}^2$ whose proper symmetry group has index 1 (meaning all symmetries are proper) is illustrated in Figure 1 (right).

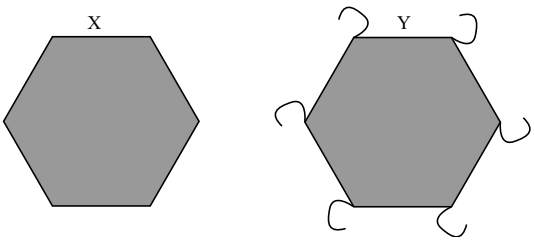

Figure 1. $\mathrm{Symm}(X) = D_6$, while $\mathrm{Symm}(Y) = \mathbb{Z}_6$.

Symmetry groups of subsets of $\mathbb{R}^2$ are useful for studying objects that are essentially 2-dimensional, like snowflakes and certain crystal structures. Many subsets of $\mathbb{R}^2$, like the wallpaper tilings of $\mathbb{R}^2$ illustrated in some M.C. Escher prints, have infinite symmetry groups. Chapter 28 of [**5**] describes the classification of such infinite "wallpaper groups". Perhaps surprisingly, the only *finite* symmetry groups in dimension 2 are D_m and $\mathbb{Z}_m$. The following theorem is attributed to Leonardo da Vinci (1452-1519):

Proposition 3.19. *For $X \subset \mathbb{R}^2$, if $Symm(X)$ is finite, then it is isomorphic to D_m or $\mathbb{Z}_m$ for some m.*

The proof involves two steps. First, when $\mathrm{Symm}(X)$ is finite, its elements must share a common fixed point, so it is isomorphic to a subgroup of $O(2)$. Second, D_m and $\mathbb{Z}_m$ are the only finite subgroups of $O(2)$.

Symmetry groups of subsets of $\mathbb{R}^3$ are even more interesting. In chemistry, the physical properties of a substance are intimately related to the symmetry groups of its molecules. In dimension 3, there are still very few possible finite symmetry groups:

Theorem 3.20. *For $X \subset \mathbb{R}^3$, if $Symm^+(X)$ is finite, then it is isomorphic to D_m, $\mathbb{Z}_m$, A_4, S_4 or A_5.*

Here, S_m denotes the group of permutations of a set with m elements, and $A_m \subset S_m$ denotes the subgroup of even permutations (called the **alternating group**). Like the $n = 2$ case, the proof involves verifying that all symmetries have a common fixed point and that the only finite subgroups of $SO(3)$ are D_m, $\mathbb{Z}_m$, A_4, S_4 and A_5.

The *regular solids* provide examples of sets whose proper symmetry groups equal A_4, S_4 and A_5. A **regular solid** (also called a "platonic solid" or a "regular polyhedra") is a polyhedra whose faces are mutually congruent regular polygons, at each of whose vertices the same number of edges meet. A famous classification theorem, attributed to Plato around 400 B.C., says that there are only five regular solids, pictured in Figure 2. The regular solids were once con-

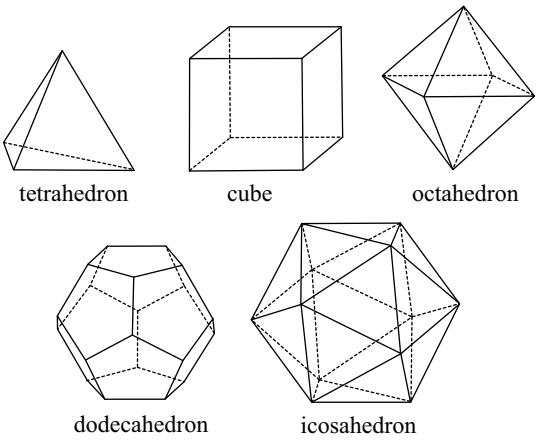

Figure 2. The five regular solids.

sidered to be sacred shapes, thought to represent fire, earth, air, the universe, and water.

The regular solids exemplify the last three possibilities enumerated in Theorem 3.20, which enhances one's sense that they are of universal importance. It turns out that A_4 is the proper symmetry group of a tetrahedron, S_4 is the proper symmetry group of a cube or an octahedron, and A_5 is the proper symmetry group of a dodecahedron or an icosahedron. See [**6**] for a complete calculation of these proper symmetry groups and a proof of Theorem 3.20. Since a cube has 6 faces, 12 edges, and 8 vertices, it may be surprising that its proper symmetry group is S_4. What does a cube have 4 of which get permuted by its proper symmetries? It has 4 diagonals (lines connecting antipodal pairs of vertices). This observation is the starting point of the calculation of its proper symmetry group.

8. Exercises

Ex. 3.1. Prove part (4) of Proposition 3.3.

Ex. 3.2. Prove Equations 3.5 and 3.6.

Ex. 3.3. Prove Proposition 3.5.

Ex. 3.4. Prove Proposition 3.18.

Ex. 3.5. Let $A \in GL_n(\mathbb{K})$. Prove that $A \in \mathcal{O}_n(\mathbb{K})$ if and only if the columns of A are an orthonormal basis of $\mathbb{K}^n$.

Ex. 3.6.

(1) Show that for every $A \in O(2) - SO(2)$, $R_A : \mathbb{R}^2 \to \mathbb{R}^2$ is a flip about some line through the origin. How is this line determined by the angle of A (as in Equation 3.8)?

(2) Let $B = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \in SO(2)$. Assume that θ is not an integer multiple of π. Prove that B does not commute with any $A \in O(2) - SO(2)$. *Hint: Show that R_{AB} and R_{BA} act differently on the line in $\mathbb{R}^2$ about which A is a flip.*

Ex. 3.7. Describe the product of two arbitrary elements of $O(2)$ in terms of their angles (as in Equation 3.8).

Ex. 3.8. Let $A \in O(n)$ have determinant -1. Prove that:

$$O(n) = SO(n) \cup \{A \cdot B \mid B \in SO(n)\}.$$

Ex. 3.9. Define a map $f : O(n) \to SO(n) \times \{+1, -1\}$ as follows:

$$f(A) = (\det(A) \cdot A, \det A).$$

(1) If n is odd, prove that f is an isomorphism.

(2) Assume that n is odd and that $X \subset \mathbb{R}^n$ is symmetric about the origin, which means that $-p \in X$ if and only if $p \in X$. Also assume that $\mathrm{Symm}(X) \subset O(n)$; in other words, X has no translational symmetries. Prove that $\mathrm{Symm}(X)$ is isomorphic to $\mathrm{Symm}^+(X) \times \{+1, -1\}$.
Comment: Four of the five regular solids are symmetric about the origin. The tetrahedron is not; its proper symmetry group is A_4 and its full symmetry group is S_4.

(3) Prove that $O(2)$ is not isomorphic to $SO(2) \times \{+1, -1\}$. *Hint: How many elements of order two are there?*

Ex. 3.10. Prove that $\mathrm{Trans}(\mathbb{R}^n)$ is a normal subgroup of $\mathrm{Isom}(\mathbb{R}^n)$.

Ex. 3.11. Prove that the **Affine group**,

$$\mathrm{Aff}_n(\mathbb{K}) = \left\{ \begin{pmatrix} A & 0 \\ V & 1 \end{pmatrix} \Big| A \in GL_n(\mathbb{K}) \text{ and } V \in \mathbb{K}^n \right\},$$

is a subgroup of $GL_{n+1}(\mathbb{K})$. Any $F \in \mathrm{Aff}_n(\mathbb{K})$ can be identified with the function $f(X) = R_A(X) + V$ from $\mathbb{K}^n$ to $\mathbb{K}^n$ as in Section 6. Prove that f sends translated lines in $\mathbb{K}^n$ to translated lines in $\mathbb{K}^n$. A *translated line* in $\mathbb{K}^n$ means a set of the form $\{v_0 + v | v \in W\}$, where $v_0 \in \mathbb{K}^n$, and $W \subset \mathbb{K}^n$ is a 1-dimensional $\mathbb{K}$-subspace.

Ex. 3.12. Is $\mathrm{Aff}_1(\mathbb{R})$ abelian? Explain algebraically and visually.

Ex. 3.13. Let $A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$.

(1) Calculate $R_A(x, y, z, w)$.

(2) Describe a subgroup, H, of $O(4)$ that is isomorphic to S_4 (S_4 = the group of permutations of a 4 element set).

(3) Describe a subgroup, H, of $O(n)$ that is isomorphic to S_n. What is $H \cap SO(n)$?

(4) Prove that every finite group is isomorphic to a subgroup of $O(n)$ for some integer n. *Hint: Use Cayley's Theorem, found in any abstract algebra textbook.*

Ex. 3.14. Let $\mathfrak{g}$ be a $\mathbb{K}$-subspace of $\mathbb{K}^n$ with dimension d. Let $\mathcal{B} = \{X_1, ..., X_d\}$ be an orthonormal basis of $\mathfrak{g}$. Let $f : \mathfrak{g} \to \mathfrak{g}$ be $\mathbb{K}$-linear. Let $A \in M_d(\mathbb{K})$ represent f in the basis $\mathcal{B}$. Prove that the following are equivalent:

(1) $A \in \mathcal{O}_d(\mathbb{K})$.

(2) $\langle f(X), f(Y) \rangle = \langle X, Y \rangle$ for all $X, Y \in \mathfrak{g}$.

Show by example that this is false when $\mathcal{B}$ is not orthonormal.

Ex. 3.15. Prove that the tetrahedron's symmetry group is S_4 and its proper symmetry group is A_4. *Hint: The symmetries permute the four vertices.*

Ex. 3.16. Think of $Sp(1)$ as the group of unit-length quaternions; that is, $Sp(1) = \{q \in \mathbb{H} \mid |q| = 1\}$.

(1) For every $q \in Sp(1)$, show that the conjugation map $C_q : \mathbb{H} \to \mathbb{H}$, defined as $C_q(V) = q \cdot V \cdot \overline{q}$, is an orthogonal linear transformation. Thus, with respect to the natural basis $\{1, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$ of $\mathbb{H}$, C_q can be regarded as an element of $O(4)$.

(2) For every $q \in Sp(1)$, verify that $C_q(1) = 1$ and therefore that C_q sends $\mathrm{Im}(\mathbb{H}) = \mathrm{span}\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ to itself. Conclude that the restriction $C_q\big|_{\mathrm{Im}(\mathbb{H})}$ can be regarded as an element of $O(3)$.

(3) Define $\varphi : Sp(1) \to O(3)$ as:
$$\varphi(q) = C_q\big|_{\mathrm{Im}(\mathbb{H})}.$$
Verify that φ is a group homomorphism.

(4) Verify that the kernel of φ is $\{1, -1\}$ and therefore that φ is two-to-one.

Comment: We will show later that the image of φ is $SO(3)$.

Ex. 3.17. Think of $Sp(1) \times Sp(1)$ as the group of pairs of unit-length quaternions.

(1) For every $q = (q_1, q_2) \in Sp(1) \times Sp(1)$, show that the map $F(q) : \mathbb{H} \to \mathbb{H}$ defined as $F(q)(V) = q_1 \cdot V \cdot \overline{q_2}$, is an orthogonal linear transformation. Thus, with respect to the natural basis $\{1, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$ of $\mathbb{H}$, $F(q)$ can be regarded as an element of $O(4)$.

(2) Show that this function $F : Sp(1) \times Sp(1) \to O(4)$ is a group homomorphism.

(3) Verify that the kernel of F is $\{(1, 1), (-1, -1)\}$ and therefore that F is two-to-one.

(4) How is F related to the function φ from the previous exercise?

Comment: We will show later that the image of F is $SO(4)$.

Ex. 3.18 (Gram-Schmidt)**.** For $m < n$, let $S = \{v_1, v_2, ..., v_m\} \subset \mathbb{K}^n$ be an orthonormal set.

(1) Prove that S can be completed to an orthonormal basis; that is, there exist vectors $v_{m+1}, ..., v_n$ such that $\{v_1, ..., v_n\}$ is an orthonormal basis of $\mathbb{K}^n$.

(2) Prove that the vectors in part (1) can be chosen such that the matrix, $M \in \mathcal{O}_n(\mathbb{K})$, whose rows are $\{v_1, ..., v_n\}$ (in that order) has determinant 1.

(3) Re-phrase the case $m = 1$ of part (2) as follows: For any unit-length vector $v \in \mathbb{K}^n$, there exists a matrix $M \in \mathcal{O}_n(\mathbb{K})$ with determinant 1 such that $R_M(e_1) = v$.