

12.6 The Shape of an Orthogonal Matrix

An orthogonal matrix is normal, so the real normal form of the previous section applies to it — and the extra information that an orthogonal matrix carries, that every eigenvalue has modulus 1, collapses that form into something very small. What comes out is a complete description of the rigid motions of $\mathbb{R}^n$ fixing the origin: a few axes reversed, and the rest of the space turned in independent two-dimensional planes. In dimension 3 this is the statement that every rotation has an axis, and we push it as far as an explicit formula for the rotation matrix in terms of its axis and angle.

Throughout, $\mathbb{R}^n$ carries the standard inner product, and $\mathbf{Q}$ is a real matrix.

12.6.1 From the real normal form to the canonical form

Two pieces of notation, both recalled. For $\theta \in \mathbb{R}$, Chapter 11 wrote

$$\mathbf{R}_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

for the matrix that rotates the plane by θ (Theorem 11.7.6). It is exactly the rotation-scaling block of Chapter 10 for a complex number on the unit circle: $\mathbf{R}_\theta = \mathbf{\Lambda}(e^{i\theta})$ in the notation of Definition 10.4.1, since $e^{i\theta} = \cos \theta + i \sin \theta$. And $\mathbf{A} \oplus \mathbf{B}$ is the block diagonal matrix with diagonal blocks $\mathbf{A}$ and $\mathbf{B}$.

Now the observation that starts everything. If $\mathbf{Q} \in O(n)$ then $\mathbf{Q}^\top \mathbf{Q} = \mathbf{I}_n = \mathbf{Q}\mathbf{Q}^\top$, so $\mathbf{Q}$ commutes with its own adjoint: **every orthogonal matrix is normal**. So Theorem 12.5.3 applies. All that remains is to say what the blocks can be, and for that we use the one thing an orthogonal matrix has that a general normal matrix does not.

Theorem 12.6.1. Canonical Form of an Orthogonal Matrix

Let $\mathbf{Q} \in O(n)$. Then there are integers $p, q, m \geq 0$ with $p + q + 2m = n$, and angles $\theta_1, \dots, \theta_m$ with $0 < \theta_j < \pi$, and a matrix $\mathbf{P} \in O(n)$, such that

$$\mathbf{P}^\top \mathbf{Q} \mathbf{P} = \mathbf{I}_p \oplus (-\mathbf{I}_q) \oplus \mathbf{R}_{\theta_1} \oplus \cdots \oplus \mathbf{R}_{\theta_m}.$$

Moreover $\det \mathbf{Q} = (-1)^q$, and p, q , and the multiset $\{\theta_1, \dots, \theta_m\}$ are determined by $\mathbf{Q}$.

Idea. $\mathbf{Q}$ is normal, so Theorem 12.5.3 already hands us a block diagonal matrix with 1×1 blocks $[a]$ and 2×2 blocks $\mathbf{\Lambda}(\lambda)$. The blocks are unconstrained for a general normal matrix; here they are not, because the eigenvalues of an isometry lie on the unit circle. A real eigenvalue on the unit circle is ± 1 , and a conjugate pair on the unit circle is $e^{\pm i\theta}$, which turns the scaling block into a pure rotation. Then sort the ± 1 s to the front.

Proof. The operator $T_{\mathbf{Q}}$ on $\mathbb{R}^n$ satisfies $T_{\mathbf{Q}}^* = T_{\mathbf{Q}^\top}$ by Theorem 11.6.3, since the standard basis is orthonormal, so $T_{\mathbf{Q}}^* T_{\mathbf{Q}} = T_{\mathbf{Q}^\top} T_{\mathbf{Q}} = \text{id}$ and likewise in the other order. Hence $T_{\mathbf{Q}}$ is normal (Definition 11.6.3), and by Theorem 12.5.3 there is an orthonormal basis $\mathcal{B}$ of $\mathbb{R}^n$ in

which

$$[T_{\mathbf{Q}}]_{\mathcal{B}}^{\mathcal{B}} = [a_1] \oplus \cdots \oplus [a_r] \oplus \mathbf{\Lambda}(\lambda_1) \oplus \cdots \oplus \mathbf{\Lambda}(\lambda_m),$$

where $a_i \in \mathbb{R}$ and $\lambda_j = c_j + d_j i$ with $d_j > 0$. Let $\mathbf{P}$ be the matrix whose columns are the vectors of $\mathcal{B}$. Its columns are orthonormal, so $\mathbf{P} \in O(n)$ by Theorem 11.7.3 (f), and $\mathbf{P}^T \mathbf{Q} \mathbf{P} = \mathbf{P}^{-1} \mathbf{Q} \mathbf{P} = [T_{\mathbf{Q}}]_{\mathcal{B}}^{\mathcal{B}}$ by Theorem 4.2.3.

The blocks. A block diagonal matrix is triangular by blocks, so the eigenvalues of $[T_{\mathbf{Q}}]_{\mathcal{B}}^{\mathcal{B}}$ over $\mathbb{C}$ are the a_i together with the roots of each $\det(x\mathbf{I}_2 - \mathbf{\Lambda}(\lambda_j)) = (x - c_j)^2 + d_j^2$, namely λ_j and $\overline{\lambda_j}$. Now $O(n) = U(n) \cap M_n(\mathbb{R})$ by Proposition 11.7.4 (a), so $\mathbf{Q}$ is unitary, and every complex eigenvalue λ of $\mathbf{Q}$ satisfies $|\lambda| = 1$ by Proposition 11.7.1 applied to $T_{\mathbf{Q}}$ on $\mathbb{C}^n$. Therefore $a_i = \pm 1$ for each i , and $c_j^2 + d_j^2 = 1$ for each j . Since $c_j^2 + d_j^2 = 1$ and $d_j > 0$, there is a unique $\theta_j \in (0, \pi)$ with $c_j = \cos \theta_j$ and $d_j = \sin \theta_j$, and then $\mathbf{\Lambda}(\lambda_j) = \mathbf{R}_{\theta_j}$.

Sorting. Let p be the number of i with $a_i = 1$ and q the number with $a_i = -1$, so $p + q + 2m = n$. Reordering the vectors of $\mathcal{B}$ so that the $+1$ blocks come first, then the -1 blocks, then the rotation blocks in pairs, leaves the basis orthonormal and permutes the diagonal blocks; the new $\mathbf{P}$ is the old one with its columns permuted, still orthogonal. This gives the displayed form.

The determinant and uniqueness. By Theorem 7.5.4 and Corollary 7.5.7, $\det \mathbf{Q} = \det(\mathbf{P}^T \mathbf{Q} \mathbf{P})$, and the determinant of a block diagonal matrix is the product of the determinants of its blocks (Theorem 7.8.4), which are 1, -1 , and $\det \mathbf{R}_{\theta_j} = \cos^2 \theta_j + \sin^2 \theta_j = 1$. Hence $\det \mathbf{Q} = (-1)^q$. Finally, similar matrices have the same eigenvalues with the same multiplicities, and reading the eigenvalues off the displayed form gives 1 with multiplicity p , -1 with multiplicity q , and the pairs $e^{\pm i\theta_j}$; so p , q , and the θ_j are recovered from $\mathbf{Q}$ alone. This proves the theorem. ■

So an orthogonal matrix is, up to an orthogonal change of coordinates, nothing but a list of angles together with a count of reversed axes. The number q of reversed axes is not itself an invariant of the *geometry* — a -1 block and a -1 block together make $-\mathbf{I}_2 = \mathbf{R}_{\pi}$, a rotation by π — which is why the theorem demands $\theta_j \neq \pi$ and keeps the -1 s separate. What the determinant sees is only the parity of q .

Quick check. What is the canonical form of a $\mathbf{Q} \in O(2)$ with $\det \mathbf{Q} = -1$? Compare with Theorem 11.7.6.

Solution. Here $p + q + 2m = 2$ and $(-1)^q = -1$, so q is odd, forcing $q = 1$, $p = 1$, $m = 0$. The canonical form is $\mathbf{I}_1 \oplus (-\mathbf{I}_1) = \text{diag}(1, -1)$. Chapter 11 found that such a $\mathbf{Q}$ is the reflection $\mathbf{M}_{\theta}$ in a line L , fixing L and negating $L^{\perp}$; in the orthonormal basis of eigenvectors of Theorem 11.7.6 (b) its matrix is exactly $\text{diag}(1, -1)$. Note what the canonical form drops: the angle of the mirror. All reflections of the plane are orthogonally similar to one another, and [Exercise \(C1: The plane, re-derived\)](#) makes this comparison precise.

12.6.2 Rotations of space

In dimension 3 the arithmetic $p + q + 2m = 3$ leaves very few possibilities, and they all look the same.

Corollary 12.6.2. Orthogonal Matrices in Dimension Three

Let $\mathbf{Q} \in \mathrm{O}(3)$. Then there are an orthonormal basis $(\mathbf{u}, \mathbf{w}_1, \mathbf{w}_2)$ of $\mathbb{R}^3$ and an angle $\theta \in [0, \pi]$ such that, in that basis, the matrix of $T_{\mathbf{Q}}$ is

$$[1] \oplus \mathbf{R}_{\theta} \quad \text{if } \det \mathbf{Q} = 1, \quad [-1] \oplus \mathbf{R}_{\theta} \quad \text{if } \det \mathbf{Q} = -1.$$

In particular every $\mathbf{Q} \in \mathrm{SO}(3)$ fixes the line $\mathrm{span}(\mathbf{u})$ pointwise and rotates the plane $\mathrm{span}(\mathbf{u})^{\perp}$ by θ .

Proof. Write the canonical form of Theorem 12.6.1, with $p + q + 2m = 3$ and $\det \mathbf{Q} = (-1)^q$.

Case $\det \mathbf{Q} = 1$. Then q is even, so $q \in \{0, 2\}$. If $q = 0$ then $p + 2m = 3$, giving $(p, m) = (3, 0)$ or $(1, 1)$; if $q = 2$ then $p = 1$ and $m = 0$. The three forms are

$$\mathbf{I}_3 = [1] \oplus \mathbf{R}_0, \quad [1] \oplus \mathbf{R}_{\theta_1}, \quad [1] \oplus (-\mathbf{I}_2) = [1] \oplus \mathbf{R}_{\pi},$$

and in every case $p \geq 1$, so the form is $[1] \oplus \mathbf{R}_{\theta}$ with $\theta \in [0, \pi]$.

Case $\det \mathbf{Q} = -1$. Then q is odd, so $q \in \{1, 3\}$. If $q = 1$ then $p + 2m = 2$, giving $(p, m) = (2, 0)$ or $(0, 1)$; if $q = 3$ then $p = m = 0$. The three forms are $[-1] \oplus \mathbf{I}_2 = [-1] \oplus \mathbf{R}_0$, then $[-1] \oplus \mathbf{R}_{\theta_1}$, and $-\mathbf{I}_3 = [-1] \oplus \mathbf{R}_{\pi}$.

In both cases the first basis vector $\mathbf{u}$ spans the ± 1 eigenline and $(\mathbf{w}_1, \mathbf{w}_2)$ is an orthonormal basis of $\mathrm{span}(\mathbf{u})^{\perp}$, which is what the last sentence asserts when $\det \mathbf{Q} = 1$. ■

Chapter 11 proved the $\mathrm{SO}(3)$ statement directly, as Theorem 11.7.7, by producing a real eigenvalue from the odd degree of the characteristic polynomial and then invoking the classification of $\mathrm{O}(2)$. We do not repeat that argument: the point here is that the same conclusion falls out of the canonical form with no work at all, and that the same three lines also settle $\det \mathbf{Q} = -1$. An orthogonal matrix of determinant -1 in dimension 3 is a **rotatory reflection**: a rotation about an axis followed by the reflection in the plane perpendicular to it. Its two extreme cases are the plane reflection ($\theta = 0$) and the antipodal map $-\mathbf{I}_3$ ($\theta = \pi$).

Remark. The angle θ is determined by $\mathbf{Q}$ only up to sign, because exchanging $\mathbf{w}_1$ and $\mathbf{w}_2$ replaces $\mathbf{R}_{\theta}$ by $\mathbf{R}_{-\theta}$. Restricting to $\theta \in [0, \pi]$, as above, makes it unique; to speak of a *signed* angle one must also fix which of the two unit vectors on the axis is called $\mathbf{u}$.

12.6.3 The cross-product matrix

The canonical form tells us that a rotation of $\mathbb{R}^3$ exists in a suitable basis. For computation we want it in the standard basis, written directly in terms of the axis. The bookkeeping is done by one matrix.

Recall from Chapter 7 that the cross product of $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ is

$$\mathbf{a} \times \mathbf{b} = (a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1).$$

For fixed $\mathbf{a}$ this is linear in $\mathbf{b}$, so it is given by a matrix.

Definition 12.6.1. Cross-product Matrix

For $\mathbf{n} = (n_1, n_2, n_3) \in \mathbb{R}^3$, the **cross-product matrix** of $\mathbf{n}$ is

$$\mathbf{K}_{\mathbf{n}} := \begin{pmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{pmatrix} \in M_3(\mathbb{R}),$$

so that $\mathbf{K}_{\mathbf{n}}\mathbf{v} = \mathbf{n} \times \mathbf{v}$ for every $\mathbf{v} \in \mathbb{R}^3$.

The last clause is a check of the three entries against the formula above, with $\mathbf{a} = \mathbf{n}$ and $\mathbf{b} = \mathbf{v}$. The map $\mathbf{n} \mapsto \mathbf{K}_{\mathbf{n}}$ is linear and injective, and its image is exactly the set of skew-symmetric 3×3 matrices: a skew matrix has zero diagonal and three free entries below it, and those three entries are $n_3, -n_2, n_1$. We record the facts we need, all of them for a **unit** vector.

Lemma 12.6.3. Properties of the Cross-product Matrix

Let $\mathbf{n} \in \mathbb{R}^3$ with $\|\mathbf{n}\| = 1$, and write $\mathbf{K} = \mathbf{K}_{\mathbf{n}}$. Then

- (a) $\mathbf{K}^T = -\mathbf{K}$ and $\mathbf{K}\mathbf{n} = \mathbf{0}$;
- (b) $\mathbf{K}^2 = \mathbf{nn}^T - \mathbf{I}_3$, so $\mathbf{K}^2$ is symmetric and $\mathbf{K}^2\mathbf{v} = -\mathbf{v}$ for every $\mathbf{v} \perp \mathbf{n}$;
- (c) $\mathbf{K}^3 = -\mathbf{K}$;
- (d) for every $\mathbf{v} \perp \mathbf{n}$: $\|\mathbf{K}\mathbf{v}\| = \|\mathbf{v}\|$, $\mathbf{K}\mathbf{v} \perp \mathbf{v}$, and $\mathbf{K}\mathbf{v} \perp \mathbf{n}$.

Idea. Everything follows from (b), and (b) is one entry of a matrix product together with $n_1^2 + n_2^2 + n_3^2 = 1$. Part (c) is (b) multiplied by $\mathbf{K}$, where the term $\mathbf{Knn}^T$ dies by (a). For (d), a norm is computed through its square, and $\|\mathbf{K}\mathbf{v}\|^2 = \mathbf{v}^T \mathbf{K}^T \mathbf{K} \mathbf{v} = -\mathbf{v}^T \mathbf{K}^2 \mathbf{v}$ turns (d) into (b) as well.

Proof.

- (a) Transposing $\mathbf{K}_{\mathbf{n}}$ changes the sign of each of the three off-diagonal pairs, and the diagonal is zero. For the second claim, $\mathbf{K}\mathbf{n} = \mathbf{n} \times \mathbf{n} = \mathbf{0}$ directly from the formula, each coordinate being of the form $n_i n_j - n_j n_i$.
- (b) The $(1, 1)$ entry of $\mathbf{K}^2$ is $0 \cdot 0 + (-n_3)n_3 + n_2(-n_2) = -n_2^2 - n_3^2 = n_1^2 - 1$, using $\|\mathbf{n}\|^2 = 1$; and $(\mathbf{nn}^T - \mathbf{I}_3)_{11} = n_1^2 - 1$. The $(1, 2)$ entry of $\mathbf{K}^2$ is $0(-n_3) + (-n_3)0 + n_2 n_1 = n_1 n_2 = (\mathbf{nn}^T)_{12}$, and the $(1, 3)$ entry is $0n_2 + (-n_3)(-n_1) + n_2 0 = n_1 n_3$. The remaining six entries are obtained by cycling the indices $1 \rightarrow 2 \rightarrow 3 \rightarrow 1$, under which both $\mathbf{K}$ and $\mathbf{nn}^T - \mathbf{I}_3$ are unchanged in form. Hence $\mathbf{K}^2 = \mathbf{nn}^T - \mathbf{I}_3$, which is symmetric since $(\mathbf{nn}^T)^T = \mathbf{nn}^T$. If $\mathbf{v} \perp \mathbf{n}$ then $\mathbf{n}^T \mathbf{v} = 0$, so $\mathbf{K}^2 \mathbf{v} = \mathbf{n}(\mathbf{n}^T \mathbf{v}) - \mathbf{v} = -\mathbf{v}$.
- (c) By (b) and (a), $\mathbf{K}^3 = \mathbf{K}(\mathbf{nn}^T - \mathbf{I}_3) = (\mathbf{K}\mathbf{n})\mathbf{n}^T - \mathbf{K} = -\mathbf{K}$.

(d) Let $\mathbf{v} \perp \mathbf{n}$. Using (a) and then (b),

$$\|\mathbf{K}\mathbf{v}\|^2 = (\mathbf{K}\mathbf{v})^\top (\mathbf{K}\mathbf{v}) = \mathbf{v}^\top \mathbf{K}^\top \mathbf{K} \mathbf{v} = -\mathbf{v}^\top \mathbf{K}^2 \mathbf{v} = \mathbf{v}^\top \mathbf{v} = \|\mathbf{v}\|^2.$$

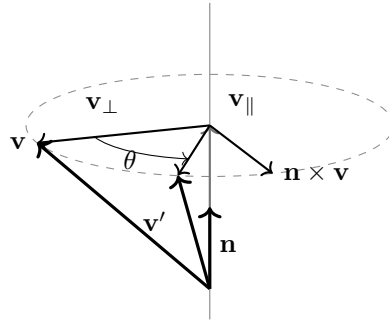
Next, $\langle \mathbf{K}\mathbf{v}, \mathbf{v} \rangle = \mathbf{v}^\top \mathbf{K} \mathbf{v}$ is a 1×1 matrix, hence equal to its own transpose $\mathbf{v}^\top \mathbf{K}^\top \mathbf{v} = -\mathbf{v}^\top \mathbf{K} \mathbf{v}$; a number equal to its own negative is 0, so $\mathbf{K}\mathbf{v} \perp \mathbf{v}$. Finally $\langle \mathbf{K}\mathbf{v}, \mathbf{n} \rangle = \mathbf{n}^\top \mathbf{K} \mathbf{v} = (\mathbf{K}^\top \mathbf{n})^\top \mathbf{v} = -(\mathbf{K}\mathbf{n})^\top \mathbf{v} = 0$ by (a). This proves the lemma. ■

Part (c) is the identity that makes every power of $\mathbf{K}$ reduce to $\mathbf{K}$ or $\mathbf{K}^2$, and it is the reason Rodrigues' formula has only three terms and no more.

12.6.4 Rodrigues' formula

Fix a unit vector $\mathbf{n}$ and an angle θ . We want a formula for “turn space by θ about the axis $\text{span}(\mathbf{n})$ ”. The construction is forced by the picture: split a vector into the part along the axis, which does not move, and the part perpendicular to the axis, which turns inside a single plane.

Write $\mathbf{v}_\parallel = \langle \mathbf{v}, \mathbf{n} \rangle \mathbf{n} = \mathbf{n} \mathbf{n}^\top \mathbf{v}$ for the component of $\mathbf{v}$ along the axis and $\mathbf{v}_\perp = \mathbf{v} - \mathbf{v}_\parallel$ for the component perpendicular to it; this is the splitting of Theorem 11.3.2 for $U = \text{span}(\mathbf{n})$, and $\mathbf{v}_\parallel = P_U \mathbf{v}$ by Theorem 11.3.3 (a) with the orthonormal basis $(\mathbf{n})$ of U . Write $\mathbf{v}'$ for the vector we want, the image of $\mathbf{v}$ under the turn.



the axis part stays, the perpendicular part turns

By Lemma 12.6.3 (d), the vector $\mathbf{K}_\mathbf{n} \mathbf{v}_\perp$ has the same length as $\mathbf{v}_\perp$ and is perpendicular to both $\mathbf{v}_\perp$ and $\mathbf{n}$: it is $\mathbf{v}_\perp$ turned by a quarter turn inside the plane $\mathbf{n}^\perp$. So turning $\mathbf{v}_\perp$ by θ means taking $\cos \theta \mathbf{v}_\perp + \sin \theta \mathbf{K}_\mathbf{n} \mathbf{v}_\perp$, and $\mathbf{v}' = \mathbf{v}_\parallel + \cos \theta \mathbf{v}_\perp + \sin \theta \mathbf{K}_\mathbf{n} \mathbf{v}_\perp$. Finally $\mathbf{K}_\mathbf{n} \mathbf{v}_\perp = \mathbf{K}_\mathbf{n} \mathbf{v}$, since $\mathbf{K}_\mathbf{n} \mathbf{v}_\parallel$ is a multiple of $\mathbf{K}_\mathbf{n} \mathbf{n} = \mathbf{0}$. That is the definition we take, and the theorem says it is the right one.

Theorem 12.6.4. Rodrigues' Rotation Formula

Let $\mathbf{n} \in \mathbb{R}^3$ be a unit vector, let $\theta \in \mathbb{R}$, and write $\mathbf{K} = \mathbf{K}_\mathbf{n}$. Define

$$\mathbf{R}_{\mathbf{n},\theta} := \mathbf{I}_3 + \sin \theta \mathbf{K} + (1 - \cos \theta) \mathbf{K}^2.$$

(a) For every $\mathbf{v} \in \mathbb{R}^3$,

$$\mathbf{R}_{\mathbf{n},\theta} \mathbf{v} = \langle \mathbf{v}, \mathbf{n} \rangle \mathbf{n} + \cos \theta (\mathbf{v} - \langle \mathbf{v}, \mathbf{n} \rangle \mathbf{n}) + \sin \theta (\mathbf{n} \times \mathbf{v}).$$

- (b) $\mathbf{R}_{\mathbf{n},\theta}\mathbf{n} = \mathbf{n}$, and for every unit $\mathbf{w} \perp \mathbf{n}$ the list $\mathcal{B} = (\mathbf{n}, \mathbf{w}, \mathbf{n} \times \mathbf{w})$ is an orthonormal basis of $\mathbb{R}^3$ with $[T_{\mathbf{R}_{\mathbf{n},\theta}}]_{\mathcal{B}}^{\mathcal{B}} = [1] \oplus \mathbf{R}_{\theta}$. In particular $\mathbf{R}_{\mathbf{n},\theta} \in \text{SO}(3)$.
- (c) Conversely, every $\mathbf{Q} \in \text{SO}(3)$ equals $\mathbf{R}_{\mathbf{n},\theta}$ for some unit vector $\mathbf{n}$ and some $\theta \in [0, \pi]$.

Idea. Part (a) is the paragraph above, read backwards: substitute $\mathbf{nn}^{\top} = \mathbf{K}^2 + \mathbf{I}_3$ from Lemma 12.6.3 (b) into the geometric expression and collect terms. Part (b) then needs only the images of three basis vectors, and every one of them is a two-line computation with (a) and the lemma; the matrix that comes out is the canonical form of Corollary 12.6.2, which also gives orthogonality and the determinant for free. Part (c) runs Corollary 12.6.2 in reverse: it hands us a basis $(\mathbf{u}, \mathbf{w}_1, \mathbf{w}_2)$, and the only question is whether $\mathbf{w}_2$ is $\mathbf{u} \times \mathbf{w}_1$ or its negative — and if it is the negative, we change the sign of the angle.

Proof.

- (a) By Lemma 12.6.3 (b), $\mathbf{nn}^{\top} = \mathbf{K}^2 + \mathbf{I}_3$. Hence, writing $\langle \mathbf{v}, \mathbf{n} \rangle \mathbf{n} = \mathbf{nn}^{\top} \mathbf{v}$ and $\mathbf{n} \times \mathbf{v} = \mathbf{Kv}$,

$$\begin{aligned} \mathbf{nn}^{\top} \mathbf{v} + \cos \theta (\mathbf{v} - \mathbf{nn}^{\top} \mathbf{v}) + \sin \theta \mathbf{Kv} \\ &= \cos \theta \mathbf{v} + (1 - \cos \theta) \mathbf{nn}^{\top} \mathbf{v} + \sin \theta \mathbf{Kv} \\ &= \cos \theta \mathbf{v} + (1 - \cos \theta) (\mathbf{K}^2 + \mathbf{I}_3) \mathbf{v} + \sin \theta \mathbf{Kv} \\ &= (\mathbf{I}_3 + \sin \theta \mathbf{K} + (1 - \cos \theta) \mathbf{K}^2) \mathbf{v}, \end{aligned}$$

since $\cos \theta + (1 - \cos \theta) = 1$. The right-hand side is $\mathbf{R}_{\mathbf{n},\theta} \mathbf{v}$.

- (b) First $\mathbf{Kn} = \mathbf{0}$ and $\mathbf{K}^2 \mathbf{n} = \mathbf{K}(\mathbf{Kn}) = \mathbf{0}$ by Lemma 12.6.3 (a), so $\mathbf{R}_{\mathbf{n},\theta} \mathbf{n} = \mathbf{n}$. Let $\mathbf{w}$ be a unit vector with $\mathbf{w} \perp \mathbf{n}$, and put $\mathbf{z} = \mathbf{Kw} = \mathbf{n} \times \mathbf{w}$. By Lemma 12.6.3 (d), $\|\mathbf{z}\| = 1$, $\mathbf{z} \perp \mathbf{w}$, and $\mathbf{z} \perp \mathbf{n}$, so $\mathcal{B} = (\mathbf{n}, \mathbf{w}, \mathbf{z})$ is an orthonormal list of three vectors in $\mathbb{R}^3$, hence an orthonormal basis (Theorem 11.1.4 and Theorem 1.6.4). Now compute the three images. We have $\mathbf{R}_{\mathbf{n},\theta} \mathbf{n} = \mathbf{n}$. Next, $\mathbf{K}^2 \mathbf{w} = -\mathbf{w}$ by Lemma 12.6.3 (b), since $\mathbf{w} \perp \mathbf{n}$, so

$$\mathbf{R}_{\mathbf{n},\theta} \mathbf{w} = \mathbf{w} + \sin \theta \mathbf{z} - (1 - \cos \theta) \mathbf{w} = \cos \theta \mathbf{w} + \sin \theta \mathbf{z}.$$

Finally $\mathbf{z} \perp \mathbf{n}$, so $\mathbf{K}^2 \mathbf{z} = -\mathbf{z}$, and $\mathbf{Kz} = \mathbf{K}^2 \mathbf{w} = -\mathbf{w}$; hence

$$\mathbf{R}_{\mathbf{n},\theta} \mathbf{z} = \mathbf{z} - \sin \theta \mathbf{w} - (1 - \cos \theta) \mathbf{z} = -\sin \theta \mathbf{w} + \cos \theta \mathbf{z}.$$

Reading off the columns (Definition 4.1.1), $[T_{\mathbf{R}_{\mathbf{n},\theta}}]_{\mathcal{B}}^{\mathcal{B}} = [1] \oplus \mathbf{R}_{\theta}$. This matrix is orthogonal, and it is the matrix of $T_{\mathbf{R}_{\mathbf{n},\theta}}$ in an orthonormal basis, so $T_{\mathbf{R}_{\mathbf{n},\theta}}$ is an isometry and $\mathbf{R}_{\mathbf{n},\theta} \in \text{O}(3)$ by Theorem 11.7.3 (f) and Definition 11.7.2. The two matrices are matrices of the same operator, hence similar (Theorem 4.2.4 (a)), so they have the same determinant (Corollary 7.5.7), namely 1; so $\mathbf{R}_{\mathbf{n},\theta} \in \text{SO}(3)$.

- (c) Let $\mathbf{Q} \in \text{SO}(3)$. By Corollary 12.6.2 there are an orthonormal basis $\mathcal{C} = (\mathbf{u}, \mathbf{w}_1, \mathbf{w}_2)$ of $\mathbb{R}^3$ and $\theta \in [0, \pi]$ with $[T_{\mathbf{Q}}]_{\mathcal{C}}^{\mathcal{C}} = [1] \oplus \mathbf{R}_{\theta}$. The vector $\mathbf{u} \times \mathbf{w}_1$ is a unit vector orthogonal to $\mathbf{u}$ and $\mathbf{w}_1$ by Lemma 12.6.3 (d), and so is $\mathbf{w}_2$; since $\text{span}(\mathbf{u}, \mathbf{w}_1)^{\perp}$ is a line (Theorem 11.3.2 (c)), its unit vectors are $\pm \mathbf{u} \times \mathbf{w}_1$, so $\mathbf{w}_2 = \varepsilon \mathbf{u} \times \mathbf{w}_1$ with $\varepsilon = \pm 1$.

Case $\varepsilon = 1$. Then (b), applied with $\mathbf{n} = \mathbf{u}$ and $\mathbf{w} = \mathbf{w}_1$, gives $[T_{\mathbf{R}_{\mathbf{u},\theta}}]_{\mathcal{C}}^{\mathcal{C}} = [1] \oplus \mathbf{R}_{\theta} = [T_{\mathbf{Q}}]_{\mathcal{C}}^{\mathcal{C}}$. Operators with the same matrix in one basis are equal, so $\mathbf{Q} = \mathbf{R}_{\mathbf{u},\theta}$.

Case $\varepsilon = -1$. Then $\mathbf{u} \times \mathbf{w}_1 = -\mathbf{w}_2$, and the matrix of $T_{\mathbf{Q}}$ in the basis $(\mathbf{u}, \mathbf{w}_1, -\mathbf{w}_2)$ is obtained from $[1] \oplus \mathbf{R}_{\theta}$ by changing the sign of the third basis vector, which changes the

sign of the two off-diagonal entries of the $\mathbf{R}_\theta$ block: it is $[1] \oplus \mathbf{R}_{-\theta}$. By the previous case, $\mathbf{Q} = \mathbf{R}_{\mathbf{u}, -\theta}$. Finally $\mathbf{K}_{-\mathbf{u}} = -\mathbf{K}_{\mathbf{u}}$, so

$$\mathbf{R}_{-\mathbf{u}, \theta} = \mathbf{I}_3 - \sin \theta \mathbf{K}_{\mathbf{u}} + (1 - \cos \theta) \mathbf{K}_{\mathbf{u}}^2 = \mathbf{R}_{\mathbf{u}, -\theta},$$

and $\mathbf{Q} = \mathbf{R}_{-\mathbf{u}, \theta}$ with $-\mathbf{u}$ a unit vector and $\theta \in [0, \pi]$. This proves the theorem. ■

Part (c) is worth restating on its own: **a rotation of $\mathbb{R}^3$ is nothing but a unit vector and an angle**. That is three numbers of data, one of them redundant, for an object that started life as nine matrix entries subject to six equations.

Quick check. Compute $\mathbf{R}_{\mathbf{n}, \pi}$ from Rodrigues' formula and identify the map geometrically.

Solution. With $\sin \pi = 0$ and $1 - \cos \pi = 2$, the formula gives $\mathbf{R}_{\mathbf{n}, \pi} = \mathbf{I}_3 + 2\mathbf{K}^2 = \mathbf{I}_3 + 2(\mathbf{nn}^\top - \mathbf{I}_3) = 2\mathbf{nn}^\top - \mathbf{I}_3$, using Lemma 12.6.3 (b). It fixes $\mathbf{n}$ and negates every vector perpendicular to $\mathbf{n}$: the half-turn about the axis. Equivalently $\mathbf{R}_{\mathbf{n}, \pi} = -(\mathbf{I}_3 - 2\mathbf{nn}^\top) = -\mathbf{H}_{\mathbf{n}}$, minus the Householder reflection of Definition 11.8.1 in the plane $\mathbf{n}^\perp$.

12.6.5 Reading off the axis and the angle

Rodrigues' formula runs one way: from $(\mathbf{n}, \theta)$ to $\mathbf{Q}$. The inverse direction is just as cheap, because the formula splits $\mathbf{Q}$ into a symmetric part and a skew part that carry the two pieces of data separately.

Theorem 12.6.5. Axis and Angle from the Matrix

Let $\mathbf{Q} = \mathbf{R}_{\mathbf{n}, \theta} \in \text{SO}(3)$ with $\mathbf{n}$ a unit vector and $\theta \in [0, \pi]$. Then

- (a) $\text{tr } \mathbf{Q} = 1 + 2 \cos \theta$, so $\cos \theta = \frac{1}{2}(\text{tr } \mathbf{Q} - 1)$ and θ is determined by $\mathbf{Q}$;
- (b) $\mathbf{Q} - \mathbf{Q}^\top = 2 \sin \theta \mathbf{K}_{\mathbf{n}}$;
- (c) if $0 < \theta < \pi$, then $\sin \theta > 0$ and

$$\mathbf{n} = \frac{1}{2 \sin \theta} (q_{32} - q_{23}, q_{13} - q_{31}, q_{21} - q_{12});$$

if $\theta = 0$ then $\mathbf{Q} = \mathbf{I}_3$ and the axis is arbitrary; and if $\theta = \pi$ then $\mathbf{Q} + \mathbf{I}_3 = 2\mathbf{nn}^\top$, whose nonzero columns are the multiples $2n_i \mathbf{n}$ of $\mathbf{n}$.

Proof. Write $\mathbf{K} = \mathbf{K}_{\mathbf{n}}$. By Lemma 12.6.3 (b), $\mathbf{K}^2 = \mathbf{nn}^\top - \mathbf{I}_3$, so $\text{tr } \mathbf{K}^2 = \text{tr}(\mathbf{nn}^\top) - 3 = \|\mathbf{n}\|^2 - 3 = -2$; and $\text{tr } \mathbf{K} = 0$, the diagonal of $\mathbf{K}$ being zero.

- (a) Taking traces in Rodrigues' formula, which is legitimate because the trace is linear (Theorem 0.9.6 (1)),

$$\text{tr } \mathbf{Q} = 3 + \sin \theta \cdot 0 + (1 - \cos \theta)(-2) = 1 + 2 \cos \theta.$$

Since $\cos$ is injective on $[0, \pi]$, this determines θ .

- (b) By [Lemma 12.6.3](#) (a) and (b), $\mathbf{K}^\top = -\mathbf{K}$ and $(\mathbf{K}^2)^\top = \mathbf{K}^2$. Transposing Rodrigues' formula therefore gives $\mathbf{Q}^\top = \mathbf{I}_3 - \sin \theta \mathbf{K} + (1 - \cos \theta) \mathbf{K}^2$, and subtracting, $\mathbf{Q} - \mathbf{Q}^\top = 2 \sin \theta \mathbf{K}$.
- (c) For $\theta \in (0, \pi)$ we have $\sin \theta > 0$, so (b) may be divided by $2 \sin \theta$, and the three independent entries of $\mathbf{K}\mathbf{n}$ are $(\mathbf{K}\mathbf{n})_{32} = n_1$, $(\mathbf{K}\mathbf{n})_{13} = n_2$, and $(\mathbf{K}\mathbf{n})_{21} = n_3$ by [Definition 12.6.1](#); the displayed formula reads these off $\mathbf{Q} - \mathbf{Q}^\top$. If $\theta = 0$ then Rodrigues' formula gives $\mathbf{Q} = \mathbf{I}_3$. If $\theta = \pi$ then $\mathbf{Q} = 2\mathbf{n}\mathbf{n}^\top - \mathbf{I}_3$, as computed in the Quick check above, so $\mathbf{Q} + \mathbf{I}_3 = 2\mathbf{n}\mathbf{n}^\top$, whose i -th column is $2n_i\mathbf{n}$; at least one $n_i \neq 0$, so at least one column is a nonzero multiple of $\mathbf{n}$. This proves the theorem. ■

The two degenerate angles are exactly the two at which the skew part $\mathbf{Q} - \mathbf{Q}^\top$ vanishes, and they are exactly the two at which the axis is not determined by a *signed* angle: at $\theta = 0$ there is no rotation to speak of, and at $\theta = \pi$ turning by $+\pi$ and $-\pi$ about the same axis give the same map. Everywhere else, part (c) is a two-subtraction recipe.

Example. Axis and Angle of a Space Rotation

Let

$$\mathbf{Q} = \frac{1}{9} \begin{pmatrix} 1 & -4 & 8 \\ 8 & 4 & 1 \\ -4 & 7 & 4 \end{pmatrix}.$$

Verify that $\mathbf{Q} \in \text{SO}(3)$, find its axis and angle, and confirm the answer with Rodrigues' formula.

Solution. Orthogonality. The columns are $\frac{1}{9}(1, 8, -4)$, $\frac{1}{9}(-4, 4, 7)$, and $\frac{1}{9}(8, 1, 4)$. Each has squared norm $\frac{1}{81}(1 + 64 + 16) = \frac{1}{81}(16 + 16 + 49) = \frac{1}{81}(64 + 1 + 16) = 1$, and the three pairwise inner products are

$$\frac{1}{81}(-4 + 32 - 28) = \frac{1}{81}(8 + 8 - 16) = \frac{1}{81}(-32 + 4 + 28) = 0.$$

So $\mathbf{Q} \in \text{O}(3)$ by Theorem 11.7.3 (f).

Determinant. Expanding along the first row, $729 \det \mathbf{Q} = 1(16 - 7) - (-4)(32 + 4) + 8(56 + 16) = 9 + 144 + 576 = 729$, so $\det \mathbf{Q} = 1$ and $\mathbf{Q} \in \text{SO}(3)$.

Angle. $\text{tr } \mathbf{Q} = \frac{1}{9}(1 + 4 + 4) = 1$, so $\cos \theta = \frac{1}{2}(1 - 1) = 0$ and $\theta = \pi/2$ by [Theorem 12.6.5](#) (a).

Axis. Here $2 \sin \theta = 2$. The three entries needed by [Theorem 12.6.5](#) (c) are

$$\begin{aligned} q_{32} - q_{23} &= \frac{7-1}{9} = \frac{6}{9}, \\ q_{13} - q_{31} &= q_{21} - q_{12} = \frac{8+4}{9} = \frac{12}{9}, \end{aligned}$$

so $\mathbf{n} = \frac{1}{2}(\frac{6}{9}, \frac{12}{9}, \frac{12}{9}) = \frac{1}{3}(1, 2, 2)$, which is a unit vector since $1 + 4 + 4 = 9$.

Check. $\mathbf{Q}(1, 2, 2) = \frac{1}{9}(1 - 8 + 16, 8 + 8 + 2, -4 + 14 + 8) = \frac{1}{9}(9, 18, 18) = (1, 2, 2)$, so the

axis is fixed. And Rodrigues' formula with $\theta = \pi/2$ reads $\mathbf{R}_{\mathbf{n},\pi/2} = \mathbf{I}_3 + \mathbf{K} + \mathbf{K}^2$, where

$$\mathbf{K} = \frac{1}{3} \begin{pmatrix} 0 & -2 & 2 \\ 2 & 0 & -1 \\ -2 & 1 & 0 \end{pmatrix}, \quad \mathbf{K}^2 = \mathbf{nn}^\top - \mathbf{I}_3 = \frac{1}{9} \begin{pmatrix} -8 & 2 & 2 \\ 2 & -5 & 4 \\ 2 & 4 & -5 \end{pmatrix}.$$

Adding $\mathbf{I}_3 + \mathbf{K} + \mathbf{K}^2$ entry by entry, for instance $1 + 0 - \frac{8}{9} = \frac{1}{9}$ in position $(1,1)$ and $0 - \frac{2}{9} + \frac{2}{9} = -\frac{4}{9}$ in position $(1,2)$, returns $\mathbf{Q}$.

12.6.6 Reflections generate

The canonical form is one way to say that an orthogonal matrix is built from simple pieces. Here is another, and the pieces are even simpler: mirrors. Recall from Chapter 11 the Householder reflection $\mathbf{H}_{\mathbf{w}} = \mathbf{I}_n - 2\mathbf{ww}^\top/\|\mathbf{w}\|^2$ determined by a nonzero $\mathbf{w}$ (Definition 11.8.1); it is orthogonal, equal to its own inverse, and has determinant -1 (Proposition 11.8.2).

Lemma 12.6.6. A Mirror Between Two Vectors of Equal Length

Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ with $\|\mathbf{x}\| = \|\mathbf{y}\|$ and $\mathbf{x} \neq \mathbf{y}$. Then $\mathbf{w} := \mathbf{x} - \mathbf{y}$ is nonzero and $\mathbf{H}_{\mathbf{w}}\mathbf{x} = \mathbf{y}$.

Proof. $\mathbf{w} \neq \mathbf{0}$ because $\mathbf{x} \neq \mathbf{y}$. Expanding the two inner products and using $\|\mathbf{y}\|^2 = \|\mathbf{x}\|^2$,

$$\begin{aligned} \|\mathbf{w}\|^2 &= \|\mathbf{x}\|^2 - 2\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2 = 2(\|\mathbf{x}\|^2 - \langle \mathbf{x}, \mathbf{y} \rangle), \\ \langle \mathbf{x}, \mathbf{w} \rangle &= \|\mathbf{x}\|^2 - \langle \mathbf{x}, \mathbf{y} \rangle, \end{aligned}$$

so $2\langle \mathbf{x}, \mathbf{w} \rangle = \|\mathbf{w}\|^2$. Hence, by the formula $\mathbf{H}_{\mathbf{w}}\mathbf{v} = \mathbf{v} - 2(\langle \mathbf{v}, \mathbf{w} \rangle/\|\mathbf{w}\|^2)\mathbf{w}$ of Definition 11.8.1,

$$\mathbf{H}_{\mathbf{w}}\mathbf{x} = \mathbf{x} - \mathbf{w} = \mathbf{x} - (\mathbf{x} - \mathbf{y}) = \mathbf{y},$$

as claimed. ■

This is Theorem 11.8.3 with the target chosen for geometry rather than for arithmetic: there, $\mathbf{y}$ was a prescribed multiple of $\mathbf{e}_1$ and the sign was picked to avoid cancellation; here $\mathbf{y}$ is whatever we need, and the mirror is the perpendicular bisector of the segment from $\mathbf{x}$ to $\mathbf{y}$.

Theorem 12.6.7. Reflections Generate the Orthogonal Group

Let $n \geq 1$ and $\mathbf{Q} \in O(n)$. Then $\mathbf{Q}$ is a product of at most n Householder reflections. (The identity is the empty product, of zero reflections.)

Idea. Induct on n , and make the induction bite by fixing one coordinate at a time. If $\mathbf{Q}$ already fixes $\mathbf{e}_1$, then it maps $\mathbf{e}_1^\perp$ into itself, and the inductive hypothesis applied inside that $(n-1)$ -dimensional space produces at most $n-1$ reflections, each of which extends to a reflection of $\mathbb{R}^n$ by taking the same $\mathbf{w}$. If $\mathbf{Q}$ does not fix $\mathbf{e}_1$, spend one reflection making it

do so: $\mathbf{Q}\mathbf{e}_1$ and $\mathbf{e}_1$ are two distinct vectors of the same length, and Lemma 12.6.6 is exactly the tool that swaps them.

Proof. Induction on n . For $n = 1$, a matrix $[a] \in O(1)$ satisfies $a^2 = [a]^\top[a] = 1$, so $(a-1)(a+1) = 0$ and $a = \pm 1$; thus $O(1) = \{[1], [-1]\}$, where $[1]$ is the empty product and $[-1] = \mathbf{H}_{\mathbf{e}_1}$, one reflection.

Let $n \geq 2$ and assume the statement in dimension $n-1$. Let $\mathbf{Q} \in O(n)$ and set $\mathbf{x} = \mathbf{Q}\mathbf{e}_1$, a unit vector since $T_{\mathbf{Q}}$ is an isometry (Theorem 11.7.3).

Case 1: $\mathbf{x} = \mathbf{e}_1$. Put $U = \text{span}(\mathbf{e}_1)^\perp$, of dimension $n-1$ by Theorem 11.3.2 (c). For $\mathbf{v} \in U$ we have $\langle \mathbf{Q}\mathbf{v}, \mathbf{e}_1 \rangle = \langle \mathbf{Q}\mathbf{v}, \mathbf{Q}\mathbf{e}_1 \rangle = \langle \mathbf{v}, \mathbf{e}_1 \rangle = 0$, the middle equality by Theorem 11.7.3 (b); so U is $T_{\mathbf{Q}}$ -invariant. Identify U with $\mathbb{R}^{n-1}$ through the orthonormal basis $(\mathbf{e}_2, \dots, \mathbf{e}_n)$. The restriction $T_{\mathbf{Q}}|_U$ is an isometry of U , so its matrix $\mathbf{Q}'$ in that basis lies in $O(n-1)$ (Theorem 11.7.3 (f)), and by the inductive hypothesis $\mathbf{Q}' = \mathbf{H}_{\mathbf{w}'_1} \cdots \mathbf{H}_{\mathbf{w}'_k}$ with $k \leq n-1$ and $\mathbf{w}'_i \in \mathbb{R}^{n-1}$. Let $\mathbf{w}_i \in U \subseteq \mathbb{R}^n$ be the vector with coordinates $\mathbf{w}'_i$ in the basis $(\mathbf{e}_2, \dots, \mathbf{e}_n)$. Then $\mathbf{H}_{\mathbf{w}_i}$ fixes $\mathbf{e}_1$, because $\mathbf{e}_1 \perp \mathbf{w}_i$ (Proposition 11.8.2 (d)), and on U it acts as $\mathbf{H}_{\mathbf{w}'_i}$, because the defining formula only involves inner products, which agree. Hence $\mathbf{Q}$ and $\mathbf{H}_{\mathbf{w}_1} \cdots \mathbf{H}_{\mathbf{w}_k}$ agree on $\mathbf{e}_1$ and on U , so they agree on $\mathbb{R}^n = \text{span}(\mathbf{e}_1) \oplus U$, and $\mathbf{Q}$ is a product of $k \leq n-1 \leq n$ reflections.

Case 2: $\mathbf{x} \neq \mathbf{e}_1$. Then $\|\mathbf{x}\| = 1 = \|\mathbf{e}_1\|$, so $\mathbf{w} := \mathbf{x} - \mathbf{e}_1$ is nonzero and $\mathbf{H}_{\mathbf{w}}\mathbf{x} = \mathbf{e}_1$ by Lemma 12.6.6. The matrix $\mathbf{H}_{\mathbf{w}}\mathbf{Q}$ is orthogonal, being a product of orthogonal matrices (Proposition 11.7.4 (a)), and $\mathbf{H}_{\mathbf{w}}\mathbf{Q}\mathbf{e}_1 = \mathbf{H}_{\mathbf{w}}\mathbf{x} = \mathbf{e}_1$. By Case 1, $\mathbf{H}_{\mathbf{w}}\mathbf{Q}$ is a product of at most $n-1$ reflections. Multiplying on the left by $\mathbf{H}_{\mathbf{w}}$ and using $\mathbf{H}_{\mathbf{w}}^2 = \mathbf{I}_n$ (Proposition 11.8.2 (b)),

$$\mathbf{Q} = \mathbf{H}_{\mathbf{w}}(\mathbf{H}_{\mathbf{w}}\mathbf{Q}),$$

a product of at most n reflections. This completes the induction. ■

Counting reflections recovers the determinant: each factor contributes -1 (Proposition 11.8.2 (e)), so a product of k reflections has determinant $(-1)^k$, and an element of $SO(n)$ needs an even number of them. The bound n is sharp — $-\mathbf{I}_n$ is a product of n reflections and no fewer, since a product of k reflections fixes the intersection of k hyperplanes, a subspace of dimension at least $n-k$, while $-\mathbf{I}_n$ fixes only $\mathbf{0}$.

Remark. The general theorem behind this one, due to Cartan and Dieudonné, says the same for the orthogonal group of any nondegenerate symmetric bilinear form over any field of characteristic $\neq 2$. Only the Euclidean case is proved here, and only the Euclidean case is used.

12.6.7 Composing rotations

Two rotations of the plane about the origin compose into a rotation, and the angles add: $\mathbf{R}_\alpha \mathbf{R}_\beta = \mathbf{R}_{\alpha+\beta}$, so $SO(2)$ is abelian (Chapter 11). The reason is visible in the canonical form: in dimension 2 there is only one rotation block, so a rotation is a single number, and numbers commute.

In dimension 3 the picture changes completely, because there is an axis to move as well as an angle to add.

Example. Two Quarter Turns in Two Orders

Let $\mathbf{A} = \mathbf{R}_{\mathbf{e}_1, \pi/2}$ and $\mathbf{B} = \mathbf{R}_{\mathbf{e}_3, \pi/2}$, the quarter turns about the first and third coordinate axes. Compute $\mathbf{AB}$ and $\mathbf{BA}$, and find the axis and angle of each.

Solution. From [Theorem 12.6.4](#) (b), $\mathbf{A}$ fixes $\mathbf{e}_1$ and turns the $(\mathbf{e}_2, \mathbf{e}_3)$ -plane by $\pi/2$, and $\mathbf{B}$ fixes $\mathbf{e}_3$ and turns the $(\mathbf{e}_1, \mathbf{e}_2)$ -plane by $\pi/2$:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Multiplying,

$$\mathbf{AB} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix}, \quad \mathbf{BA} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

These are different matrices, so $\text{SO}(3)$ is not abelian. Both have trace 0, so both are rotations by the angle θ with $\cos \theta = \frac{1}{2}(0 - 1) = -\frac{1}{2}$, that is $\theta = 2\pi/3$. Their axes differ. Solving $\mathbf{AB}\mathbf{v} = \mathbf{v}$ gives $-v_2 = v_1$, $-v_3 = v_2$, $v_1 = v_3$, so the axis of $\mathbf{AB}$ is $\text{span}((1, -1, 1))$; solving $\mathbf{BA}\mathbf{v} = \mathbf{v}$ gives $v_3 = v_1$, $v_1 = v_2$, $v_2 = v_3$, so the axis of $\mathbf{BA}$ is $\text{span}((1, 1, 1))$. Same angle, different axes: the order of two quarter turns changes where the resulting axis points.

That the product is *again* a rotation is not an accident of this example: $\text{SO}(3)$ is a group (Proposition 11.7.4), so any product of rotations is orthogonal with determinant 1, and [Corollary 12.6.2](#) makes it a rotation. What no formula makes obvious is which one; [Exercise \(C2: Products of rotations\)](#) pursues this.

Warning. In dimension 4 and above, $\det \mathbf{Q} = 1$ does not mean “rotation about an axis”. Take

$$\mathbf{Q} = \mathbf{R}_{\pi/2} \oplus \mathbf{R}_{\pi/3} \in \text{O}(4), \quad \det \mathbf{Q} = 1 \cdot 1 = 1.$$

Its canonical form is itself, with $p = q = 0$ and $m = 2$, so 1 is not an eigenvalue and $\mathbf{Q}$ fixes no nonzero vector at all: there is no axis. Worse, the two angles are different, so $\mathbf{Q}$ is not a rotation “in a plane” either — it turns two orthogonal planes at two different speeds, and no single angle describes it. Chapter 11 made the same point with $-\mathbf{I}_4 \in \text{SO}(4)$; the canonical form explains it, since for even n the counts $p = q = 0$ are allowed. The clean statement “determinant 1 means a rotation about an axis” is a feature of dimension 3, where $p + q + 2m = 3$ is odd and forces $p \geq 1$ when q is even.

12.6.8 Exercises

12.6.8.1 A. Check your understanding

Exercise. A1

- (a) State the canonical form of an element of $O(n)$, saying what the blocks are and which angles are allowed.
- (b) Which hypothesis on $\mathbf{Q}$ turns the rotation-scaling blocks of the real normal form into rotation blocks, and why?
- (c) Write down the cross-product matrix $\mathbf{K}_{\mathbf{n}}$ and state the identity $\mathbf{K}_{\mathbf{n}}^3 = ?$ for a unit $\mathbf{n}$.
- (d) Given $\mathbf{Q} \in SO(3)$, how do you find its angle, and how do you find its axis? Name the two degenerate cases.
- (e) True or false: if $\mathbf{Q} \in O(n)$ and $\det \mathbf{Q} = 1$, then $\mathbf{Q}$ fixes a nonzero vector. Justify your answer.
- (f) How many Householder reflections can be needed to write an element of $O(n)$, and what does the parity of that number say?

Solution.

- (a) By [Theorem 12.6.1](#), $\mathbf{P}^\top \mathbf{Q} \mathbf{P} = \mathbf{I}_p \oplus (-\mathbf{I}_q) \oplus \mathbf{R}_{\theta_1} \oplus \cdots \oplus \mathbf{R}_{\theta_m}$ for some $\mathbf{P} \in O(n)$, with $p + q + 2m = n$ and $0 < \theta_j < \pi$. The angles 0 and π are excluded because $\mathbf{R}_0 = \mathbf{I}_2$ and $\mathbf{R}_\pi = -\mathbf{I}_2$ are already accounted for by the ± 1 blocks.
- (b) Orthogonality. An orthogonal matrix is unitary as a complex matrix ([Proposition 11.7.4 \(a\)](#)), so each of its eigenvalues has modulus 1 ([Proposition 11.7.1](#)). A block $\mathbf{A}(\lambda)$ has eigenvalues $\lambda, \bar{\lambda}$, so $|\lambda| = 1$ and the scaling factor is 1.
- (c) $\mathbf{K}_{\mathbf{n}}$ has rows $(0, -n_3, n_2), (n_3, 0, -n_1), (-n_2, n_1, 0)$ ([Definition 12.6.1](#)), and $\mathbf{K}_{\mathbf{n}}^3 = -\mathbf{K}_{\mathbf{n}}$ when $\|\mathbf{n}\| = 1$ ([Lemma 12.6.3 \(c\)](#)).
- (d) The angle from the trace, $\cos \theta = \frac{1}{2}(\text{tr } \mathbf{Q} - 1)$ with $\theta \in [0, \pi]$; the axis from the skew part, $\mathbf{Q} - \mathbf{Q}^\top = 2 \sin \theta \mathbf{K}_{\mathbf{n}}$ ([Theorem 12.6.5](#)). The degenerate cases are $\theta = 0$, where $\mathbf{Q} = \mathbf{I}_3$ and there is no axis to find, and $\theta = \pi$, where the skew part vanishes and one reads $\mathbf{n}$ off $\mathbf{Q} + \mathbf{I}_3 = 2\mathbf{n}\mathbf{n}^\top$ instead.
- (e) False for even n . The matrix $\mathbf{R}_{\pi/2} \oplus \mathbf{R}_{\pi/3} \in SO(4)$ has no eigenvalue 1. It is true for odd n : then $p + q + 2m = n$ is odd and q is even, so $p = n - q - 2m$ is odd, hence $p \geq 1$, and the $\mathbf{I}_p$ block supplies a fixed vector.
- (f) At most n ([Theorem 12.6.7](#)). Each reflection has determinant -1 , so the number used has the same parity in every such factorization as the sign of $\det \mathbf{Q}$ dictates: even for $\det \mathbf{Q} = 1$, odd for $\det \mathbf{Q} = -1$.

12.6.8.2 B. Practice**Exercise. B1: Axis and angle**

For each of the following, decide whether it lies in $SO(3)$, $O(3) \setminus SO(3)$, or neither. For those in $SO(3)$, find the axis and the angle; for those in $O(3)$ with determinant -1 , give the canonical form $[-1] \oplus \mathbf{R}_\theta$ and say what the map is geometrically.

(a) $\frac{1}{9} \begin{pmatrix} 4 & -8 & 1 \\ 4 & 1 & -8 \\ 7 & 4 & 4 \end{pmatrix}.$

$$(b) \frac{1}{9} \begin{pmatrix} -7 & 4 & 4 \\ 4 & -1 & 8 \\ 4 & 8 & -1 \end{pmatrix}.$$

$$(c) \frac{1}{9} \begin{pmatrix} 7 & -4 & -4 \\ -4 & 1 & -8 \\ -4 & -8 & 1 \end{pmatrix}.$$

$$(d) \frac{1}{9} \begin{pmatrix} 4 & -8 & 1 \\ 4 & 1 & -8 \\ 7 & 4 & 5 \end{pmatrix}.$$

Solution.

- (a) The columns $\frac{1}{9}(4, 4, 7)$, $\frac{1}{9}(-8, 1, 4)$, $\frac{1}{9}(1, -8, 4)$ have squared norms $\frac{1}{81}(16 + 16 + 49) = \frac{1}{81}(64 + 1 + 16) = \frac{1}{81}(1 + 64 + 16) = 1$ and pairwise inner products $\frac{1}{81}(-32 + 4 + 28) = \frac{1}{81}(4 - 32 + 28) = \frac{1}{81}(-8 - 8 + 16) = 0$, so the matrix is orthogonal. Expanding along the first row, $729 \det = 4(4 + 32) + 8(16 + 56) + 1(16 - 7) = 144 + 576 + 9 = 729$, so it lies in $SO(3)$. Its trace is $\frac{1}{9}(4 + 1 + 4) = 1$, so $\cos \theta = 0$ and $\theta = \pi/2$. For the axis, $q_{32} - q_{23} = \frac{1}{9}(4 + 8) = \frac{12}{9}$, $q_{13} - q_{31} = \frac{1}{9}(1 - 7) = -\frac{6}{9}$, and $q_{21} - q_{12} = \frac{1}{9}(4 + 8) = \frac{12}{9}$; dividing by $2 \sin \theta = 2$ gives $\mathbf{n} = \frac{1}{3}(2, -1, 2)$. Check: $\mathbf{Q}(2, -1, 2) = \frac{1}{9}(8 + 8 + 2, 8 - 1 - 16, 14 - 4 + 8) = (2, -1, 2)$.

- (b) The matrix is symmetric. Each column has entries $\pm 7, \pm 4, \pm 4$ in some order, so each has squared norm $\frac{1}{81}(49 + 16 + 16) = 1$, and the pairwise inner products are $\frac{1}{81}(-28 - 4 + 32) = \frac{1}{81}(-28 + 32 - 4) = \frac{1}{81}(16 - 8 - 8) = 0$; so it is orthogonal. Expanding along the first row, $729 \det = -7(1 - 64) - 4(-4 - 32) + 4(32 + 4) = 441 + 144 + 144 = 729$, so it lies in $SO(3)$. Its trace is $\frac{1}{9}(-7 - 1 - 1) = -1$, so $\cos \theta = -1$ and $\theta = \pi$: a half-turn.

The skew part is $\mathbf{O}$, so use $\mathbf{Q} + \mathbf{I}_3 = \frac{1}{9} \begin{pmatrix} 2 & 4 & 4 \\ 4 & 8 & 8 \\ 4 & 8 & 8 \end{pmatrix} = 2\mathbf{nn}^\top$ with $\mathbf{n} = \frac{1}{3}(1, 2, 2)$. Check:

$$\mathbf{Q}(1, 2, 2) = \frac{1}{9}(-7 + 8 + 8, 4 - 2 + 16, 4 + 16 - 2) = (1, 2, 2).$$

- (c) This is the negative of (b), so it is orthogonal with $\det = (-1)^3 \cdot 1 = -1$. Its trace is 1, and for the form $[-1] \oplus \mathbf{R}_\theta$ the trace is $-1 + 2 \cos \theta$, so $\cos \theta = 1$ and $\theta = 0$: the canonical form is $[-1] \oplus \mathbf{I}_2 = \text{diag}(-1, 1, 1)$, a reflection. Indeed it equals $\mathbf{I}_3 - 2\mathbf{nn}^\top = \mathbf{H}_\mathbf{n}$ with $\mathbf{n} = \frac{1}{3}(1, 2, 2)$, the reflection in the plane $x_1 + 2x_2 + 2x_3 = 0$.
- (d) Neither. This is (a) with the $(3, 3)$ entry changed from 4 to 5. The third column $\frac{1}{9}(1, -8, 5)$ has squared norm $\frac{1}{81}(1 + 64 + 25) = \frac{90}{81} = \frac{10}{9} \neq 1$, so the columns are not orthonormal and the matrix is not orthogonal (Theorem 11.7.3 (f)). Its determinant is $\frac{85}{81}$, neither 1 nor -1 , which settles it a second way (Proposition 11.7.4). Note that checking the determinant alone is never enough: $\det = 1$ does not make a matrix orthogonal.

Exercise. B2: Mirrors

- (a) Write the matrix $\mathbf{M} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ as a Householder reflection $\mathbf{H}_\mathbf{w}$, and say which plane is its mirror.

- (b) Following the proof of [Theorem 12.6.7](#), write $\mathbf{Q} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ as a product of Householder reflections, and check the count against $\det \mathbf{Q}$.

Solution.

- (a) $\mathbf{M}$ exchanges $\mathbf{e}_1$ and $\mathbf{e}_3$ and fixes $\mathbf{e}_2$. By [Lemma 12.6.6](#) with $\mathbf{x} = \mathbf{e}_1$ and $\mathbf{y} = \mathbf{e}_3$, take $\mathbf{w} = \mathbf{e}_1 - \mathbf{e}_3 = (1, 0, -1)$; then $\|\mathbf{w}\|^2 = 2$ and

$$\mathbf{H}_{\mathbf{w}} = \mathbf{I}_3 - \mathbf{w}\mathbf{w}^\top = \mathbf{I}_3 - \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} = \mathbf{M}.$$

The mirror is $\text{span}(\mathbf{w})^\perp$, the plane $x_1 = x_3$.

- (b) $\mathbf{Q}\mathbf{e}_1 = \mathbf{e}_2 \neq \mathbf{e}_1$, so we are in Case 2 with $\mathbf{w}_1 = \mathbf{Q}\mathbf{e}_1 - \mathbf{e}_1 = (-1, 1, 0)$ and $\|\mathbf{w}_1\|^2 = 2$:

$$\mathbf{H}_{\mathbf{w}_1} = \mathbf{I}_3 - \mathbf{w}_1\mathbf{w}_1^\top = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then $\mathbf{H}_{\mathbf{w}_1}\mathbf{Q} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$, which fixes $\mathbf{e}_1$ and exchanges $\mathbf{e}_2$ and $\mathbf{e}_3$; the computation of (a), with the indices shifted, shows that it is $\mathbf{H}_{\mathbf{w}_2}$ for $\mathbf{w}_2 = \mathbf{e}_2 - \mathbf{e}_3 = (0, 1, -1)$. Hence $\mathbf{Q} = \mathbf{H}_{\mathbf{w}_1}\mathbf{H}_{\mathbf{w}_2}$, a product of two reflections. The count is even, matching $\det \mathbf{Q} = (-1)^2 = 1$; and indeed $\mathbf{Q}$ is the rotation by $2\pi/3$ about $(1, 1, 1)$ found in [Example \(Two Quarter Turns in Two Orders\)](#).

Exercise. B3: Rodrigues in an example

- (a) Compute $\mathbf{R}_{\mathbf{e}_3, \theta}$ from Rodrigues' formula and check that it is $\mathbf{R}_\theta \oplus [1]$.
 (b) Compute $\mathbf{R}_{\mathbf{n}, 2\pi/3}$ for $\mathbf{n} = \frac{1}{\sqrt{3}}(1, 1, 1)$, and identify the resulting matrix.

Solution.

- (a) With $\mathbf{n} = \mathbf{e}_3$ we have $\mathbf{K} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ and $\mathbf{K}^2 = \mathbf{e}_3\mathbf{e}_3^\top - \mathbf{I}_3 = \text{diag}(-1, -1, 0)$. Hence

$$\mathbf{R}_{\mathbf{e}_3, \theta} = \mathbf{I}_3 + \sin \theta \mathbf{K} + (1 - \cos \theta) \mathbf{K}^2 = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

the diagonal entries being $1 - (1 - \cos \theta) = \cos \theta$ twice and 1 once. This is $\mathbf{R}_\theta \oplus [1]$, the rotation of the $(\mathbf{e}_1, \mathbf{e}_2)$ -plane, as it must be.

- (b) Here $\cos \theta = -\frac{1}{2}$ and $\sin \theta = \frac{\sqrt{3}}{2}$. Writing $\mathbf{1} = (1, 1, 1)$, so that $\mathbf{n} = \mathbf{1}/\sqrt{3}$ and

$$\mathbf{n}\mathbf{n}^\top = \frac{1}{3}\mathbf{1}\mathbf{1}^\top,$$

$$\sin \theta \mathbf{K} = \frac{\sqrt{3}}{2} \cdot \frac{1}{\sqrt{3}} \mathbf{K}_1 = \frac{1}{2} \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix},$$

$$(1 - \cos \theta) \mathbf{K}^2 = \frac{3}{2} \left(\frac{1}{3} \mathbf{1}\mathbf{1}^\top - \mathbf{I}_3 \right) = \frac{1}{2} \mathbf{1}\mathbf{1}^\top - \frac{3}{2} \mathbf{I}_3.$$

Adding $\mathbf{I}_3$ to these two gives $\frac{1}{2}\mathbf{1}\mathbf{1}^\top - \frac{1}{2}\mathbf{I}_3$ plus the skew part. The first piece has 0 on the diagonal and $\frac{1}{2}$ elsewhere, and adding the skew part turns each off-diagonal pair $\frac{1}{2}, \frac{1}{2}$ into 0, 1:

$$\mathbf{R}_{\mathbf{n}, 2\pi/3} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

This is the cyclic shift $\mathbf{e}_1 \mapsto \mathbf{e}_2 \mapsto \mathbf{e}_3 \mapsto \mathbf{e}_1$, which is the matrix $\mathbf{B}\mathbf{A}$ of [Example \(Two Quarter Turns in Two Orders\)](#): same axis $(1, 1, 1)$, same angle $2\pi/3$.

12.6.8.3 C. Going deeper

Exercise. C1: The plane, re-derived

- (a) Use [Theorem 12.6.1](#) to prove that every $\mathbf{Q} \in \mathbf{O}(2)$ is either orthogonally similar to $\mathbf{R}_\theta$ for a unique $\theta \in [0, \pi]$, or orthogonally similar to $\text{diag}(1, -1)$.
- (b) Chapter 11 proved more: $\mathbf{Q}$ equals $\mathbf{R}_\theta$ or $\mathbf{M}_\theta$ (Theorem 11.7.6). Explain exactly what information (a) loses, and why the canonical form must lose it.
- (c) Deduce that any two reflections of $\mathbb{R}^2$ are orthogonally similar, and that $\mathbf{R}_\alpha$ and $\mathbf{R}_\beta$ are orthogonally similar if and only if $\beta = \pm\alpha$.

Solution.

- (a) With $n = 2$, $p + q + 2m = 2$. If $m = 1$ then $p = q = 0$ and the form is $\mathbf{R}_{\theta_1}$ with $\theta_1 \in (0, \pi)$. If $m = 0$ then $(p, q) \in \{(2, 0), (1, 1), (0, 2)\}$, giving $\mathbf{I}_2 = \mathbf{R}_0$, $\text{diag}(1, -1)$, and $-\mathbf{I}_2 = \mathbf{R}_\pi$. So the possibilities are $\mathbf{R}_\theta$ with $\theta \in [0, \pi]$ and $\text{diag}(1, -1)$, and they are distinguished by the determinant. Uniqueness of θ follows from [Theorem 12.6.1](#), or directly from $\text{tr } \mathbf{R}_\theta = 2 \cos \theta$ and the injectivity of $\cos$ on $[0, \pi]$.
- (b) Part (a) determines $\mathbf{Q}$ only up to orthogonal similarity, that is, up to a change of orthonormal coordinates. It therefore cannot see *which* line is the mirror of a reflection, nor the sign of the angle of a rotation. It must lose the mirror, because the mirror moves: if $\mathbf{P}$ is the rotation by φ then $\mathbf{P}\mathbf{M}_\theta\mathbf{P}^\top = \mathbf{M}_{\theta+2\varphi}$, so all the $\mathbf{M}_\theta$ are conjugate. It must lose the sign of the angle, because conjugating by $\text{diag}(1, -1)$ sends $\mathbf{R}_\theta$ to $\mathbf{R}_{-\theta}$.
- (c) Every reflection has determinant -1 , so by (a) each is orthogonally similar to $\text{diag}(1, -1)$, and orthogonal similarity is an equivalence relation, so any two of them are similar to each other. For rotations: if $\beta = \pm\alpha$ then $\mathbf{R}_\beta = \mathbf{R}_\alpha$ or $\mathbf{R}_\beta = \text{diag}(1, -1)\mathbf{R}_\alpha\text{diag}(1, -1)$, so they are orthogonally similar. Conversely, similar matrices have equal traces (Theorem 4.2.5), so $\cos \alpha = \cos \beta$, which for real angles means $\beta \equiv \pm\alpha$ modulo 2π .

Exercise. C2: Products of rotations

- (a) Prove that the product of two rotations of $\mathbb{R}^3$ is a rotation of $\mathbb{R}^3$, and explain which theorem supplies the axis.
- (b) Let $\mathbf{A}$ and $\mathbf{B}$ be the half-turns about two **perpendicular** axes $\text{span}(\mathbf{a})$ and $\text{span}(\mathbf{b})$, with $\mathbf{a} \perp \mathbf{b}$ unit vectors. Prove that $\mathbf{AB}$ is the half-turn about the axis $\text{span}(\mathbf{a} \times \mathbf{b})$.
- (c) Deduce that every rotation of $\mathbb{R}^3$ is a product of two half-turns.

Hint for (c): use Theorem 12.6.7, or the description of a half-turn in the Quick check after Theorem 12.6.4.

Solution.

- (a) $\text{SO}(3)$ is a group (Proposition 11.7.4 (c)), so if $\mathbf{A}, \mathbf{B} \in \text{SO}(3)$ then $\mathbf{AB} \in \text{SO}(3)$. By Corollary 12.6.2, $\mathbf{AB}$ fixes a line pointwise and rotates the perpendicular plane; that corollary is what supplies the axis, and it does so through the count $p \geq 1$ in the canonical form.
- (b) Extend $(\mathbf{a}, \mathbf{b})$ to an orthonormal basis $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ of $\mathbb{R}^3$, where we may take $\mathbf{c} = \mathbf{a} \times \mathbf{b}$: it is a unit vector orthogonal to $\mathbf{a}$ and $\mathbf{b}$ by Lemma 12.6.3 (d). By the Quick check after Theorem 12.6.4, $\mathbf{A} = 2\mathbf{a}\mathbf{a}^\top - \mathbf{I}_3$ fixes $\mathbf{a}$ and negates $\mathbf{a}^\perp$, so in the basis $(\mathbf{a}, \mathbf{b}, \mathbf{c})$, $\mathbf{A}$ is $\text{diag}(1, -1, -1)$ and $\mathbf{B}$ is $\text{diag}(-1, 1, -1)$. Their product is $\text{diag}(-1, -1, 1)$, which fixes $\mathbf{c}$ and negates $\mathbf{c}^\perp$: the half-turn about $\text{span}(\mathbf{c}) = \text{span}(\mathbf{a} \times \mathbf{b})$. (Its trace is -1 , confirming $\theta = \pi$ through Theorem 12.6.5 (a).)
- (c) Let $\mathbf{Q} = \mathbf{R}_{\mathbf{n}, \theta} \in \text{SO}(3)$, with $\mathbf{n}$ a unit vector and $\theta \in [0, \pi]$ (Theorem 12.6.4 (c)). Choose a unit $\mathbf{w} \perp \mathbf{n}$, and work throughout in the orthonormal basis $\mathcal{B} = (\mathbf{n}, \mathbf{w}, \mathbf{K}_{\mathbf{n}}\mathbf{w})$ of Theorem 12.6.4 (b), in which $\mathbf{Q}$ has matrix $[1] \oplus \mathbf{R}_\theta$. Put $\mathbf{w}' = \mathbf{R}_{\mathbf{n}, \theta/2}\mathbf{w}$, again a unit vector perpendicular to $\mathbf{n}$, since $\mathbf{R}_{\mathbf{n}, \theta/2}$ fixes $\mathbf{n}$ and is an isometry; by Theorem 12.6.4 (b) its coordinate vector in $\mathcal{B}$ is $\mathbf{d} = (0, \cos \frac{\theta}{2}, \sin \frac{\theta}{2})$. Let $\mathbf{A}$ and $\mathbf{B}$ be the half-turns about $\text{span}(\mathbf{w}')$ and $\text{span}(\mathbf{w})$, so that in $\mathcal{B}$ their matrices are $2\mathbf{d}\mathbf{d}^\top - \mathbf{I}_3$ and $\text{diag}(-1, 1, -1)$ respectively, by the Quick check after Theorem 12.6.4. Using $2\cos^2 \frac{\theta}{2} - 1 = \cos \theta$, $2\sin^2 \frac{\theta}{2} - 1 = -\cos \theta$, and $2\cos \frac{\theta}{2} \sin \frac{\theta}{2} = \sin \theta$,

$$2\mathbf{d}\mathbf{d}^\top - \mathbf{I}_3 = [-1] \oplus \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}.$$

Multiplying the two matrices blockwise, $\mathbf{AB}$ has matrix

$$[1] \oplus \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = [1] \oplus \mathbf{R}_\theta,$$

which is the matrix of $\mathbf{Q}$ in the same basis. Hence $\mathbf{Q} = \mathbf{AB}$.

Exercise. C3: Moving one point of the sphere to another

Let $S = \{\mathbf{x} \in \mathbb{R}^3 : \|\mathbf{x}\| = 1\}$.

- (a) Prove that for all $\mathbf{u}, \mathbf{v} \in S$ there is $\mathbf{Q} \in \text{SO}(3)$ with $\mathbf{Q}\mathbf{u} = \mathbf{v}$. (One says that $\text{SO}(3)$ **acts transitively** on S .)
- (b) Prove that $\{\mathbf{Q} \in \text{SO}(3) : \mathbf{Q}\mathbf{u} = \mathbf{u}\} = \{\mathbf{R}_{\mathbf{u}, \theta} : \theta \in \mathbb{R}\}$, and that this set is an abelian

subgroup of $SO(3)$.

- (c) Is the analogue of (a) true for $SO(2)$ acting on the unit circle of $\mathbb{R}^2$? And for $SO(4)$ acting on the unit sphere of $\mathbb{R}^4$? Justify your answers.

Solution.

- (a) Extend $\mathbf{u}$ to an orthonormal basis $(\mathbf{u}, \mathbf{a}_2, \mathbf{a}_3)$ of $\mathbb{R}^3$ (Corollary 11.2.3), and likewise extend $\mathbf{v}$ to $(\mathbf{v}, \mathbf{b}_2, \mathbf{b}_3)$. Let $\mathbf{A} = (\mathbf{u} \mid \mathbf{a}_2 \mid \mathbf{a}_3)$ and $\mathbf{B} = (\mathbf{v} \mid \mathbf{b}_2 \mid \mathbf{b}_3)$; both are orthogonal, since their columns are orthonormal (Theorem 11.7.3 (f)). Replacing $\mathbf{a}_3$ by $-\mathbf{a}_3$ changes the sign of $\det \mathbf{A}$ and keeps the basis orthonormal, so we may assume $\det \mathbf{A} = 1$, and likewise $\det \mathbf{B} = 1$. Put $\mathbf{Q} = \mathbf{B}\mathbf{A}^\top$. Then $\mathbf{Q} \in O(3)$ and $\det \mathbf{Q} = \det \mathbf{B} \det \mathbf{A}^\top = 1$ by Theorem 7.5.4 and Theorem 7.4.5, so $\mathbf{Q} \in SO(3)$; and $\mathbf{A}^\top \mathbf{u} = \mathbf{e}_1$, because the entries of $\mathbf{A}^\top \mathbf{u}$ are the inner products of $\mathbf{u}$ with the columns of $\mathbf{A}$, so $\mathbf{Q}\mathbf{u} = \mathbf{B}\mathbf{e}_1 = \mathbf{v}$.
- (b) $(\supseteq)$ Each $\mathbf{R}_{\mathbf{u},\theta}$ lies in $SO(3)$ and fixes $\mathbf{u}$, by Theorem 12.6.4 (b). $(\subseteq)$ If $\mathbf{Q} \in SO(3)$ fixes $\mathbf{u}$, write $\mathbf{Q} = \mathbf{R}_{\mathbf{n},\theta}$ with $\theta \in [0, \pi]$ (Theorem 12.6.4 (c)). If $\theta = 0$ then $\mathbf{Q} = \mathbf{I}_3 = \mathbf{R}_{\mathbf{u},0}$. Otherwise the fixed space of $\mathbf{Q}$ is exactly the line $\text{span}(\mathbf{n})$, since in the basis of Theorem 12.6.4 (b) the matrix is $[1] \oplus \mathbf{R}_\theta$ and $\mathbf{R}_\theta$ fixes no nonzero vector for $\theta \in (0, \pi]$; so $\mathbf{u} = \pm \mathbf{n}$, and $\mathbf{R}_{-\mathbf{n},\theta} = \mathbf{R}_{\mathbf{n},-\theta}$ gives $\mathbf{Q} = \mathbf{R}_{\mathbf{u},\pm\theta}$. For the group claim: all these matrices are simultaneously $[1] \oplus \mathbf{R}_\theta$ in the *same* orthonormal basis $(\mathbf{u}, \mathbf{w}, \mathbf{K}_\mathbf{u}\mathbf{w})$, and $\mathbf{R}_\alpha \mathbf{R}_\beta = \mathbf{R}_{\alpha+\beta} = \mathbf{R}_\beta \mathbf{R}_\alpha$, so the set is closed under products and inverses and is abelian.
- (c) Yes for $SO(2)$: given unit $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$, write $\mathbf{u} = (\cos \alpha, \sin \alpha)$ and $\mathbf{v} = (\cos \beta, \sin \beta)$ and take $\mathbf{Q} = \mathbf{R}_{\beta-\alpha}$, which is in $SO(2)$ and sends $\mathbf{u}$ to $\mathbf{v}$. Yes for $SO(4)$ as well, by the same argument as (a): extend $\mathbf{u}$ and $\mathbf{v}$ to orthonormal bases of $\mathbb{R}^4$, fix the determinants to 1 by flipping the sign of the last vector if necessary, and take $\mathbf{Q} = \mathbf{B}\mathbf{A}^\top$. Nothing in that argument used the dimension 3 — transitivity holds for $SO(n)$ with $n \geq 2$. What is special to $n = 3$ is the classification of the maps themselves, not their ability to move one unit vector to another.