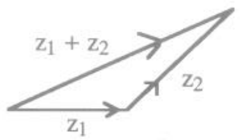
By a result of elementary geometry, the sum of lengths of two sides of a triangle equals at least the length of the third side. Since the sides of a triangle may represent vectors and hence complex numbers we have the first inequality stated below.



THEOREM 9.15 (The triangle inequality for complex numbers.) For complex numbers z_1 and z_2 we have

(a) $|z_1 + z_2| \le |z_1| + |z_2|$, (b) $||z_1| - |z_2|| \le |z_1 + z_2|$.

Proof (a) A direct proof can be given using the definition and properties of the modulus but we will be content with having deduced it from elementary geometry. In fact it is a special case of a result for n-space (Remark 7.12). We can now infer (b) by the argument $|z_1| = |(z_1 + z_2) - z_2| \le |z_1 + z_2| + |-z_2|$, for then $|z_1| - |z_2|$ and similarly $|z_2| - |z_1|$ do not exceed $|z_1 + z_2|$, and part (b) follows.

REMARKS 9.16 (1) *(Choosing the e in $e^{i\theta}$.)* The fact that $\cos\theta + \mathbf{i}\sin\theta$ behaves like a number raised to a power suggests we call it $d^{i\theta}$ for some real number d, but how should we choose d? Defining the derivative $[f(\theta) + \mathbf{i}g(\theta)]' = f'(\theta) + \mathbf{i}g'(\theta)$ for f, g real, we obtain $(\cos\theta + \mathbf{i}\sin\theta)' = -\sin\theta + \mathbf{i}\cos\theta = \mathbf{i}(\cos\theta + \mathbf{i}\sin\theta)$. But for real functions, if $d^{c\theta}$ has derivative $cd^{c\theta}$ then d is uniquely e, the base of 'natural logarithms ln(x)'. This may be deduced from the formula $(d^{c\theta})' = cd^{c\theta}$ (see a basic calculus text such as Swokowski, 1979). The result is that writing $\cos\theta + \mathbf{i}\sin\theta$ as $e^{i\theta}$ leads to a nice extension of differential calculus to complex numbers and to new techniques for solving real differential equations.

Some readers may find the series expansion $e^x = 1 + x + x^2/2! + x^3/3! + \ldots$ with $x = \mathbf{i}\theta$ more illuminating for (given that this is valid), by separating out the real and imaginary parts we obtain $e^{i\theta} = (1 - \theta^2/2! + \theta^4/4! + \ldots) + \mathbf{i}(\theta - \theta^3/3! + \theta^5/5! + \ldots)$, and these series are the expansions for $\cos\theta$ and $\sin\theta$.

(2) *(Solving equations.)* We mentioned at the start of this chapter that polynomial equations of degree less than five have a formula for their solution. Of course the quadratic equation $ax^2 + bx + c = 0$ has solutions $x = (-b \pm \sqrt{(b^2 - 4ac)})/2a$, which reduces in the case $x^2 + 2hx + c = 0$ to $x = -h \pm \sqrt{(h^2 - c)}$. The formulae for degrees 3,4 are somewhat more complicated and seem not often to be used, perhaps because Newton's iterative method is so simple to carry out (Example 13.22). However it is interesting to see these formulae and know that they exist, and it seems likely that they will be useful from time to time for equation-solving in computer graphics.

Degree 3. Dividing through by the coefficient of x^3 and setting $z = x + a_2/3$ (a_2 being the new coefficient of x^2) we reduce the cubic to the form (a) below, whose solutions are the cube roots in (b). The six values coincide in pairs.

(a) $z^3 + pz + q = 0$, (b) $z = [-q/2 \pm \sqrt{q^2/4 + p^3/27}\,]^{1/3}$.

Degree 4. As before, we eliminate the coefficient of the second highest power, after dividing through by the leading coefficient, setting now $z = x + a_3/4$. We obtain the form (c) below, whose solutions are those of the two quadratics shown in (d), where u is any solution of the cubic equation in (e) (complex cube roots are obtained by using polar forms, see Section 9.1.2).

(c) $z^4 + pz^2 + qz + r = 0$, (d) $z^2 + u/2 \pm (Az-B) = 0$,

(e) $A = \sqrt{u-p}$, $B = q/2A$, and $(u^2 - 4r)(u - p) - q^2 = 0$.

For more on the practicalities of solving polynomial equations of degree three and higher, see Schwarze et al (1990), and Pross et al (1988). We take the theory of complex numbers further for the purpose of Mandelbrot, Julia sets, and others, in Chapter 16.

EXERCISE Solve the cubic equation $z^3 + 3z + 2 = 0$. It may be helpful to express your answer in terms of $a = (1 + \sqrt{3})^{1/3}$ and $b = (-1 + \sqrt{2})^{1/3}$, noting that $ab = 1$. **or** By first squaring both sides, prove the triangle inequality for complex numbers, using especially the properties $|z|^2 = z\bar{z}$ and $\mathrm{Re}(z) \le |z|$.

9.2 The Quaternions

In this section we set up the foundations of quaternions as a preparation to their application to rotations in three dimensions.

9.2.1 Basics of quaternions

A quaternion **a** may be viewed as the extension of a complex number $a_0 + a_1\mathbf{i}$ to an entity with four components:

$$\mathbf{a} = a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}, \tag{9.14}$$

where the real numbers a_1, a_2, a_3 are written formally as coefficients of the symbols **i**, **j**, **k**. Quaternions are added, subtracted, multiplied by real numbers and by each other as algebraic expressions, subject to the following laws of multiplication

$$\begin{array}{lllll} \mathbf{ij} = \mathbf{k} = -\mathbf{ji}, & \mathbf{i}^2 = -1, \\ \mathbf{jk} = \mathbf{i} = -\mathbf{kj}, & \mathbf{j}^2 = -1, \\ \mathbf{ki} = \mathbf{j} = -\mathbf{ik}, & \mathbf{k}^2 = -1. \end{array} \tag{9.15}$$

Note that the cyclic rotation $\mathbf{i} \to \mathbf{j} \to \mathbf{k} \to \mathbf{i}$ transforms each row of (9.15) into another. We say accordingly that these laws are *invariant* under rotation of **i**, **j**, **k** (cf. (7.25)). They can be written more compactly as $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1$, once we have established that quaternion multiplication is associative: $(\mathbf{ab})\mathbf{c} = \mathbf{a}(\mathbf{bc})$, in Section 9.2.2.

EXAMPLES 9.17 (1) $2 \times 3\mathbf{j} = 6\mathbf{j} = 3\mathbf{j} \times 2$; multiplication by a real number may be done in either order.

(2) $\mathbf{i}(2\mathbf{j}-1) = 2\mathbf{ij} - \mathbf{i} = 2\mathbf{k}-\mathbf{i}$, but $(2\mathbf{j}-1)\mathbf{i} = 2\mathbf{ji} - \mathbf{i} = -2\mathbf{k} - \mathbf{i}$, a different answer. Thus quaternions do not commute, as is already clear from (9.15). But notice the steps in the multiplications.

(3)
$$\begin{aligned} (\mathbf{i}+2\mathbf{j})^2 &= (\mathbf{i}+2\mathbf{j})(\mathbf{i}+2\mathbf{j}) \\ &= \mathbf{i}(\mathbf{i}+2\mathbf{j}) + 2\mathbf{j}(\mathbf{i}+2\mathbf{j}) \\ &= \mathbf{i}^2 + 2\mathbf{ij} + 2\mathbf{ji} + 4\mathbf{j}^2 \\ &= -5, \quad \text{since } \mathbf{i}^2 = \mathbf{j}^2 = -1 \text{ and } \mathbf{ij} = -\mathbf{ji}. \end{aligned}$$

(4) $(2\mathbf{i}+3)(1+\mathbf{j}+\mathbf{k}) = 2\mathbf{i}(1+\mathbf{j}+\mathbf{k}) + 3(1+\mathbf{j}+\mathbf{k})$
$= 2\mathbf{i}+2\mathbf{k}-2\mathbf{j} + 3+3\mathbf{j}+3\mathbf{k}$
$= 3+2\mathbf{i}+\mathbf{j}+5\mathbf{k}.$

EXERCISE Show that $\mathbf{a}(\mathbf{bc}) = (\mathbf{ab})\mathbf{c}$ in case $\mathbf{a} = \mathbf{i}$, $\mathbf{b} = 3\mathbf{j}-\mathbf{k}$, $\mathbf{c} = 2-5\mathbf{j}+6\mathbf{k}$.

NOTATION 9.18 $\mathbb{H}$ is the set of all quaternions and, based on (9.14): we say **a** has *scalar* or *real* part $S\mathbf{a} = a_0$, *vector* or *pure* part $V\mathbf{a}$ (or $\acute{\mathbf{a}}$) $= a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$, and *conjugate* $\overline{\mathbf{a}} = S\mathbf{a}-V\mathbf{a} = a_0 - \acute{\mathbf{a}} = a_0 - a_1\mathbf{i} - a_2\mathbf{j} - a_3\mathbf{k}$. We call the quaternion **a** *pure* or a *vector* if it has zero scalar part, i.e. if $a_0 = 0$.

EXAMPLES 9.19 (1) The quaternion $\mathbf{a} = 3 + 2\mathbf{i} - 3\mathbf{j} + \mathbf{k}$ has real part $S\mathbf{a} = 3$, pure part $\acute{\mathbf{a}} = 2\mathbf{i}-3\mathbf{j}+\mathbf{k}$, and conjugate $\overline{\mathbf{a}} = 3-2\mathbf{i}+3\mathbf{j}-\mathbf{k}$.

(2)
$$\begin{aligned}(a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k})^2 &= a_1\mathbf{i}^2 + a_2\mathbf{j}^2 + a_3\mathbf{k}^2 \\ &\quad + a_1a_2\mathbf{ij} + a_2a_1\mathbf{ji} \\ &\quad + a_2a_3\mathbf{jk} + a_3a_2\mathbf{kj} \\ &\quad + a_3a_1\mathbf{ki} + a_1a_3\mathbf{ik} \\ &= -(a_1^2 + a_2^2 + a_3^2), \quad \text{by e.g. } \mathbf{ij} = -\mathbf{ji}.\end{aligned}$$

(3)
$$\begin{aligned}\mathbf{a}\overline{\mathbf{a}} &= (a_0 + \acute{\mathbf{a}})(a_0 - \acute{\mathbf{a}}) = a_0^2 - a_0\acute{\mathbf{a}} + \acute{\mathbf{a}}a_0 - \acute{\mathbf{a}}\acute{\mathbf{a}} \\ &= a_0^2 - \acute{\mathbf{a}}\acute{\mathbf{a}} \\ &= a_0^2 + a_1^2 + a_2^2 + a_3^2, \quad \text{by (2) above,} \\ &= \overline{\mathbf{a}}\mathbf{a}.\end{aligned}$$

NORM AND CONJUGATES Example (3) above shows that we can usefully define the *norm* or *modulus* $|\mathbf{a}|$ of a quaternion **a** by analogy with the complex numbers since it has the crucial property of being positive unless **a** is zero. Thus we define $|\mathbf{a}| \geq 0$ by

$$|\mathbf{a}|^2 = \mathbf{a}\overline{\mathbf{a}} = \overline{\mathbf{a}}\mathbf{a} = a_0^2 + a_1^2 + a_2^2 + a_3^2, \quad \textit{then} \tag{9.16}$$

$$|\mathbf{a}| \geq 0 \textit{ for all quaternions } \mathbf{a}, \textit{ and } |\mathbf{a}| = 0 \Leftrightarrow \mathbf{a} = 0, \tag{9.17A}$$

$$\textit{If } \lambda \textit{ is a positive number then } |\lambda\mathbf{a}| = \lambda|\mathbf{a}|. \tag{9.17B}$$

We call **a** a *unit quaternion* (cf. unit vectors) if $|\mathbf{a}| = 1$. In due course we will derive the famous and deeper property $|\mathbf{ab}| = |\mathbf{a}||\mathbf{b}|$, for it has important consequences (see Theorem 9.30). The immediately verifiable properties of conjugates are:

$$\overline{\mathbf{a}+\mathbf{b}} = \overline{\mathbf{a}} + \overline{\mathbf{b}}, \qquad \overline{\lambda\mathbf{a}} = \overline{\lambda}\overline{\mathbf{a}} \quad (\lambda \in \mathbf{C}) \tag{9.18}$$

$$\mathbf{a}+\overline{\mathbf{a}} = 2S\mathbf{a}, \qquad \mathbf{a}-\overline{\mathbf{a}} = 2V\mathbf{a} \tag{9.19}$$

$$\mathbf{a} \text{ is real} \Leftrightarrow \mathbf{a} = \overline{\mathbf{a}} \qquad \mathbf{a} \text{ is pure} \Leftrightarrow \mathbf{a} = -\overline{\mathbf{a}} \tag{9.20}$$

The expected property of conjugates, $\overline{\mathbf{ab}} = \overline{\mathbf{a}}\,\overline{\mathbf{b}}$, is more problematical. Indeed it fails to hold. However the situation is saved by the fact that $\overline{\mathbf{ab}} = \overline{\mathbf{b}}\,\overline{\mathbf{a}}$ (Theorem 9.24).

PURE QUATERNIONS AS VECTORS Without even a change of notation the pure quaternion $a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ represents a vector (a_1, a_2, a_3) in 3-space, the basis being $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$, $\mathbf{k} = (0, 0, 1)$. The modulus is the same from both viewpoints, $|\mathbf{a}| =$

$\sqrt{(a_1^2 + a_2^2 + a_3^2)}$, and the calculation in Example 9.19(2) gives the following result, which we state as a theorem because of its centrality.

THEOREM 9.20 The square of a pure quaternion $\mathbf{u}$ *is real. Specifically,* $\mathbf{u}^2 = -|\mathbf{u}|^2$. *In particular, every unit pure quaternion is a square root of -1.*

This is something quite new after complex numbers, where the equation $z^2 = -1$ can have only two solutions, $\pm\mathbf{i}$. We are saying that if we extend the permitted range of solutions to quaternions then every one of the infinitely many points on the unit sphere in 3-space, $S^2 = \{\mathbf{x} \in \mathbf{R}^3 : |\mathbf{x}| = 1\}$, yields such a square root. But does this help? The answer is *yes*, because these solutions enable us to represent a rotation about any axis in $\mathbf{R}^3$ by a quaternion. The mechanism is the *polar form*, introduced in Section 9.2.3 after some groundwork in 9.2.2, where we relate the three ways now available to combine two vectors: $\mathbf{ab}$, $\mathbf{a.b}$, $\mathbf{a} \times \mathbf{b}$.

EXAMPLE 9.21 Could there be square roots of -1 which are *not* pure? We'll clear this up right now. Let $\mathbf{a}^2 = -1$ $(\mathbf{a} \in \mathbf{H})$. Then in the usual notation $(a_0 + \acute{\mathbf{a}})^2 = a_0^2 + 2a_0\acute{\mathbf{a}} + \acute{\mathbf{a}}\acute{\mathbf{a}} = a_0^2 - (a_1^2 + a_2^2 + a_3^2) + 2a_0\acute{\mathbf{a}}$, by Theorem 9.4. Equating this to -1 gives two equations (i) $a_0^2-(a_1^2 + a_2^2 + a_3^2) = -1$, (ii) $2a_0\acute{\mathbf{a}} = 0$. From the second, either $a_0 = 0$ or $\acute{\mathbf{a}} = 0$. But $\acute{\mathbf{a}} = 0$ is impossible by (i) since it would involve $a_0^2 = -1$, and the square of a real number cannot be negative. Therefore $a_0 = 0$ and **a** is pure.

EXERCISE Calculate $(1-\mathbf{i}+2\mathbf{j}+5\mathbf{k})^2$ with the help of Theorem 9.20 on pure quaternions.

9.2.2 Theorems on quaternion multiplication

We can't go much further without showing that one thing the quaternions do share with complex numbers is associativity.

THEOREM 9.22 Let **a**, **b**, **c** *be quaternions. Then*
(i) $\mathbf{a}(\mathbf{b}+\mathbf{c}) = \mathbf{ab}+\mathbf{ac}$, $(\mathbf{b}+\mathbf{c})\mathbf{a} = \mathbf{ba}+\mathbf{ca}$ *(bilinearity)*
(ii) $(\mathbf{ab})\mathbf{c} = \mathbf{a}(\mathbf{bc})$ *(associativity)*

Proof (i) Note that the multiplication law for **i**,**j**,**k** extends to quaternions generally by bilinearity. That is, (i) holds by definition when **a**,**b**,**c** are individual terms of a quaternion, such as $a_1\mathbf{i}$. We sketch the verification that (i) holds as expected when **a**,**b**,**c** are general quaternions. The two parts of (i) are similar. Expanding $\mathbf{a}(\mathbf{b}+\mathbf{c})$ we obtain

$$(a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k})\{(b_0+c_0) + (b_1+c_1)\mathbf{i} + (b_2+c_2)\mathbf{j} + (b_3+c_3)\mathbf{k}\}$$

$= a_0(b_0+c_0) + a_0(b_1+c_1)\mathbf{i} + a_0(b_2+c_2)\mathbf{j} + a_0(b_3+c_3)\mathbf{k}$
+ three sets of terms in which a_0 is replaced respectively by $a_1\mathbf{i}$, $a_2\mathbf{j}$, $a_3\mathbf{k}$

$= a_0(b_v + b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}) + a_0(c_0 + c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k})$
+ (as above) $= \mathbf{ab} + \mathbf{ac}$.

Proof (ii) Expanding $(\mathbf{ab})\mathbf{c}$ with the four terms in each quaternion we obtain $4 \times 4 \times 4 =$

64 terms (pq)r, which we must show equals p(qr). When at least one of p, q, r is real there is nothing to prove, so we are left with proving that (pq)r = p(qr) when p,q,r run through the values **i**,**j**,**k**, a total of 3 x 3 x 3 = 27 cases. However the set of rules (9.15) for multiplying these factors is invariant under rotation of **i**,**j**,**k** so we need verify only the nine cases shown for readability as a matrix equality below, indexed by the first two factors (the rows correspond to the first factor and the columns to the second).

$$\begin{bmatrix} (\mathbf{ii})\mathbf{i} & (\mathbf{ij})\mathbf{i} & (\mathbf{ik})\mathbf{i} \\ (\mathbf{ji})\mathbf{i} & (\mathbf{jj})\mathbf{i} & (\mathbf{jk})\mathbf{i} \\ (\mathbf{ki})\mathbf{i} & (\mathbf{kj})\mathbf{i} & (\mathbf{kk})\mathbf{i} \end{bmatrix} = \begin{bmatrix} -\mathbf{i} & \mathbf{j} & \mathbf{k} \\ -\mathbf{j} & -\mathbf{i} & -1 \\ -\mathbf{k} & 1 & -\mathbf{i} \end{bmatrix}$$

The same matrix of answers is obtained if each product is bracketed the other way. For example, instead of (**ji**)**i** we calculate **j**(**ii**) = **j**(-1) = -**j**, obtaining the same answer. With these verifications the proof is complete.

EXERCISE Explain via rules (9.15) why the second matrix of Proof (ii) above is skew symmetric apart from the main diagonal.

INVERSES AND CANCELLATION If **ab** = 1 we say that **a** is a *left inverse* of **b** and that **b** is a *right inverse* of **a**. If **ab** = **ba** = 1 we call **b** a *2-sided inverse* of **a**, or simply an *inverse* of **a**. Fortunately the formula $\overline{\mathbf{a}} / |\mathbf{a}|^2$ for the inverse of a complex number works for the quaternions also, and for the same reason: $\mathbf{a}\overline{\mathbf{a}} = \overline{\mathbf{a}}\mathbf{a} = |\mathbf{a}|^2$, so that $\mathbf{a}(\overline{\mathbf{a}} / |\mathbf{a}|^2) = 1 = (\overline{\mathbf{a}} / |\mathbf{a}|^2)\mathbf{a}$.

Every nonzero quaternion **a** *has a unique inverse* $\mathbf{a}^{-1} = \overline{\mathbf{a}} / |\mathbf{a}|^2$ *and if* $\lambda \neq 0$ *is real then* $(\lambda\mathbf{a})^{-1} = \lambda^{-1}\mathbf{a}^{-1}$. (9.21)

Uniqueness of the inverse is a consequence of associativity, for if **b**,**c** are both inverses of **a** then **b** = **b**(**ac**) = (**ba**)**c** = 1**c** = **c**. An intriguing question remains. Could there be a left or a right inverse which is not a full 2-sided inverse? The simple answer is *no* (cf. the similar situation for square matrices). In fact every left or right inverse equals $\overline{\mathbf{a}} / |\mathbf{a}|^2$(denoted $\mathbf{a}^{-1}$). For example if **b** is a left inverse, i.e. 1 = **ba**, then, multiplying on the right by $\mathbf{a}^{-1}$, we have $\mathbf{a}^{-1} = (\mathbf{ba})\mathbf{a}^{-1} = \mathbf{b}(\mathbf{aa}^{-1}) = \mathbf{b}(1) = \mathbf{b}$. Similarly for 1 = **ab**. More generally, multiplying by $\mathbf{a}^{-1}$ gives us:

The cancellation laws for quaternions. Let $\mathbf{a} \neq 0$ $(\mathbf{a},\mathbf{b},\mathbf{c} \in \mathbf{H})$.
If **ab** = **ac** *or* **ba** = **ca** *then* **b** = **c**. (9.22)

The quaternions have no divisors of zero. That is, if **ab** = **0** *then* **a** = **0** *or* **b** = **0**. (9.23)

EXERCISE Write down the inverses of **k** and of 2 - **i** + **k**, and check that they work. Verify in this case the formula $(\mathbf{ab})^{-1} = \mathbf{b}^{-1}\mathbf{a}^{-1}$.

THREE WAYS TO MULTIPLY TWO VECTORS We can now combine a pair of vectors **x** and **y** as follows, and the next theorem gives an important relation between these products.

(1) $\mathbf{x}.\mathbf{y} = x_1y_1 + x_2y_2 + x_3y_3$, the *scalar product*,

(2) $\mathbf{x} \times \mathbf{y} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix}$, the *vector product*,

(3) **xy** the *quaternion product*.

THEOREM 9.23 Let **x**, **y** *be pure quaternions. Then* **xy** *has scalar part* -**x**.**y** *and vector part* **x** x **y**, *in fact*

$$\mathbf{xy} = -\mathbf{x}.\mathbf{y} + \mathbf{x} \times \mathbf{y} = \overline{\mathbf{yx}} \quad (S\mathbf{x} = S\mathbf{y} = 0) \tag{9.24}$$

Proof Since **ji** = -**ij**, **kj** = -**jk**, **ik** = -**ki** (9.15), we have for **xy**,

$$\begin{aligned}(x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k})(y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}) &= x_1y_1\mathbf{i}^2 + x_2y_2\mathbf{j}^2 + x_3y_3\mathbf{k}^2 \\ &\quad + (x_1y_2 - x_2y_1)\mathbf{ij} + (x_2y_3 - x_3y_2)\mathbf{jk} + (x_3y_1 - x_1y_3)\mathbf{ki}.\end{aligned}$$

Notice that the simultaneous cyclic interchange 1,2,3 and **i**,**j**,**k** moves us cyclically round the last three expressions. The last expression may be written $-(x_1y_3 - x_3y_1)\mathbf{j}$ and, applying the multiplication rules (9.15) for **i**,**j**,**k**, we obtain the result -**x**.**y** + **x** x **y**. For the second equality in (9.24) we interchange **x** and **y** to get **yx** = -**y**.**x** + **y** x **x**, whose conjugate equals **xy** since **y** x **x** = -**x** x **y**. Now for two theorems which contain probably the deepest results of this section.

THEOREM 9.24 Quaternion multiplication satisfies $\overline{\mathbf{ab}} = \overline{\mathbf{b}}\,\overline{\mathbf{a}}$

Proof With **á** denoting the pure part of quaternion **a**, we have

$$\begin{aligned}\mathbf{ab} &= (a_0 + \mathbf{á})(b_0 + \mathbf{b}') \\ &= a_0b_0 + a_0\mathbf{b}' + b_0\mathbf{á} + \mathbf{á}\mathbf{b}'.\end{aligned}$$

Hence
$$\begin{aligned}\overline{\mathbf{ab}} &= a_0b_0 - a_0\mathbf{b}' - b_0\mathbf{á} + \mathbf{b}'\mathbf{á}, && \text{by (9.24)}, \\ &= (b_0 - \mathbf{b}')(a_0 - \mathbf{á}), && \text{by inspection}, \\ &= \overline{\mathbf{b}}\,\overline{\mathbf{a}}.\end{aligned}$$

THEOREM 9.25 The norm of a quaternion product is the product of the norms:

$$|\mathbf{ab}| = |\mathbf{a}||\mathbf{b}|. \tag{9.25}$$

Proof
$$\begin{aligned}|\mathbf{ab}|^2 &= (\mathbf{ab})(\overline{\mathbf{ab}}) && \text{by definition of norm}, \\ &= (\mathbf{ab})(\overline{\mathbf{b}}\,\overline{\mathbf{a}}) && \text{by Theorem 9.24}, \\ &= \mathbf{a}(\mathbf{b}\overline{\mathbf{b}})\overline{\mathbf{a}} && \text{as } \mathbf{H} \text{ is associative, Theorem 9.22}, \\ &= \mathbf{a}\,|\mathbf{b}|^2\overline{\mathbf{a}} && \text{by definition of norm}, \\ &= \mathbf{a}\overline{\mathbf{a}}\,|\mathbf{b}|^2 && \text{as } |\mathbf{b}|^2 \text{ is real}, \\ &= |\mathbf{a}|^2\,|\mathbf{b}|^2 && \text{by definition of norm}, \\ &= (|\mathbf{a}|\,|\mathbf{b}|)^2.\end{aligned}$$

Hence the result, since norms are non-negative.

EXAMPLE 9.26 The set of all unit quaternions may be viewed as the set of all unit vectors in 4-space, forming the 3-sphere S^3 (the '3' refers to the number of independent coordinates). This set satisfies the axioms for a group, albeit an infinite one. Firstly it is closed under multiplication for if **a**, **b** are unit quaternions then $|\mathbf{ab}| = |\mathbf{a}||\mathbf{b}| = 1 \times 1 = 1$, by Theorem 9.25. Multiplication is associative by Theorem 9.22. The real number 1 acts as identity, and every unit quaternion, being nonzero (by (9.17A)) has an inverse, by (9.21), namely $\bar{\mathbf{a}}$. Of special interest are finite subgroups related to symmetry groups of polyhedra (see Table 9.4 above Figure 9.13, and Coxeter, 1974). A subgroup with eight elements is the *quaternion group* $Q = \{\pm 1, \pm\mathbf{i}, \pm\mathbf{j}, \pm\mathbf{k}\}$.

EXERCISE Prove that $|axa^{-1}| = |x|$ $(a \neq 0)$.

EXERCISE Calculate $(\mathbf{i}+3\mathbf{j}+\mathbf{k})(2\mathbf{i}-\mathbf{j}+5\mathbf{k})$ using (9.24). Check the answer satisfies $|xy| = |x||y|$.

9.2.3 The polar form of a quaternion

COPIES OF C WITHIN H The quaternions of form $a_0 + a_1\mathbf{j}$ form a copy of the complex numbers with **j** in place of **i**. Indeed, if we write suggestively $\mathbf{I} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$, of unit norm, then by Theorem 9.20 $\mathbf{I}^2 = -1$, and so

the quaternions $\lambda + \mu\mathbf{I}$ *(λ,μ real) form a copy of the complex numbers.* (9.26)

In particular the set of such numbers is commutative, though we must stick to the same **I**. In fact every polynomial in **I** with real coefficients reduces to the form $\lambda + \mu\mathbf{I}$ because of $\mathbf{I}^2 = -1$. This contrasts with the noncommutativity of the quaternions in general, which we observed as early as Example 9.17.

POLAR FORM Suppose $\mathbf{b} = b_0 + \mathbf{b'}$ is a quaternion of modulus 1. Then $-1 \le b_0 \le 1$ and from Figure 9.8 $b_0 = \cos\theta$ for a unique angle θ with $0 \le \theta \le \pi$ $(\theta = \cos^{-1}(b_0))$, and hence $|\mathbf{b'}|^2 = 1 - b_0^2 = \sin^2\theta$. But from $\sin\theta \ge 0$ for $0 \le \theta \le \pi$, we may infer that $|\mathbf{b'}| = \sin\theta$. We call θ the *argument* of **b**. Now unless **b** is real we have $\sin\theta \neq 0$, so we may write $\mathbf{I} = (1/\sin\theta)\mathbf{b'}$, giving $\mathbf{b} = \cos\theta + \mathbf{I}\sin\theta$. Hence:

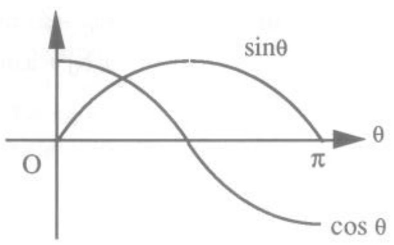


Figure 9.8 sin θ and cos θ for $0 \le \theta \le \pi$.

A non-real unit quaternion has a unique expression in the form $\cos\theta + \mathbf{I}\sin\theta$, $0 \le \theta \le \pi$, where $\mathbf{I}^2 = -1$. (9.27)

Notice that, analogously to the case of complex numbers, the conjugate $\cos\theta - \mathbf{I}\sin\theta$ equals $\cos(-\theta) + \mathbf{I}\sin(-\theta)$. Also, in the special case **b** real, we have the same expression as (9.27) but **I** is arbitrary because its coefficient $\sin\theta$ is zero. Then $b = 1$, $\theta = 0$ or $b = -1$, $\theta = \pi$. Indeed (9.27) to (9.32) go through as they did in case $\mathbf{I} = \mathbf{i}$.

DEFINITION $e^{\mathbf{I}\theta} = \cos\theta + \mathbf{I}\sin\theta$ $(\mathbf{I}^2 = -1)$. (9.28)

CONSEQUENCE $e^{\mathbf{I}\alpha}e^{\mathbf{I}\beta} = e^{\mathbf{I}(\alpha+\beta)} = e^{\mathbf{I}\beta}e^{\mathbf{I}\alpha}$ (9.29)

CONSEQUENCE The inverse of $e^{\mathbf{I}\theta}$ is its conjugate $e^{-\mathbf{I}\theta}$. (9.30)

For a general quaternion **a** of norm $r > 0$ we have $\mathbf{a} = r\mathbf{b}$, with $\mathbf{b} = (1/r)\mathbf{a}$ of unit norm, hence the unique polar form

$$\mathbf{a} = re^{\mathbf{I}\theta} = r(\cos\theta + \mathbf{I}\sin\theta). \quad (9.31)$$

Given that there is a unique θ between 0 and π we allow θ to be changed by multiples of 2π as we do with the complex number polar form, itself a special case of the quaternionic, since this leaves $\cos\theta$, $\sin\theta$ and hence the quaternion unchanged. Notice that all pure quaternions have $\theta = \cos^{-1}(0) = \pi/2$, when we insist that $0 \le \theta \le \pi$. We emphasise that, for a fixed quaternion square root **I** of -1, expressions of the form $e^{\mathbf{I}\alpha}$, $e^{\mathbf{I}\beta}$, $e^{\mathbf{I}\gamma}$ may be multiplied in any order, using (9.29). In particular we have *De Moivre's Theorem for quaternions*: if $\mathbf{I}^2 = -1$ and $n = 0, \pm 1, \pm 2, \ldots$, then

$$(\cos\theta + \mathbf{I}\sin\theta)^n = \cos n\theta + \mathbf{I}\sin n\theta \quad (9.32)$$

EXAMPLES 9.27 (Some polar forms.) As indicated in Table 9.2 below, we determine **I** by dividing the pure part by its norm, then find the angle, using an Argand diagram if necessary. We emphasise that all pure quaternions can be expressed with $\theta = \pi/2$.

(1) $\mathbf{i} = e^{\mathbf{i}\pi/2}$, $\mathbf{j} = e^{\mathbf{j}\pi/2}$, $\mathbf{k} = e^{\mathbf{k}\pi/2}$. This gives a simple illustration of the need to keep **I** fixed if we are to ensure commutativity, for $e^{\mathbf{i}\pi/2}e^{\mathbf{j}\pi/2} = \mathbf{ij} = \mathbf{k}$, whereas the other order gives $-\mathbf{k}$. Observe that $-\mathbf{k} = e^{\mathbf{I}\pi/2}$ with $\mathbf{I} = -\mathbf{k}$, but still $\theta = \pi/2$.

(2) The polar form of $\mathbf{i}+\mathbf{j}$. Another pure case, so it has $\theta = \pi/2$. We may highlight I by writing $\mathbf{i}+\mathbf{j} = \sqrt{2}.\frac{\mathbf{i}+\mathbf{j}}{\sqrt{2}} = \sqrt{2}e^{\mathbf{I}\pi/2} = \sqrt{2}e^{(\mathbf{i}+\mathbf{j})\pi/2\sqrt{2}}$.

TABLE 9.2 To find the polar form of a quaternion **a**:

(1) Write a as $\mathbf{a} = a_0 + |\mathbf{a'}|(\mathbf{a'}/|\mathbf{a'}|)$,

(2) Plot $(a_0, |\mathbf{a'}|)$ in the Argand diagram to find r,θ.

or

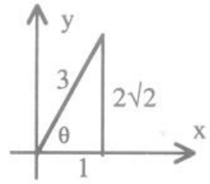
Calculate $r = |\mathbf{a}|$ from the components and identify a_0/r with a known cosine (cf. Example 9.28(2)).

Note: it may be notationally or otherwise convenient to find r,θ for the conjugate then correct the sign.

(3) The polar form of $1 + \mathbf{k}\sqrt{3}$. Not pure, so we may plot $(1,\sqrt{3})$ on an Argand diagram (i.e. in the xy-plane) and obtain $r = 2$, $\theta = \pi/3$. Thus $1 + \mathbf{k}\sqrt{3} = 2e^{\mathbf{k}\pi/3}$. By De Moivre's Theorem $(1/2)(1 + \mathbf{k}\sqrt{3})$ is a cube root of -1 (its cube has angle π).

(4) The polar form of $\sqrt{2} + \mathbf{j} - \mathbf{k} = \sqrt{2} + \sqrt{2}\,\frac{\mathbf{j}-\mathbf{k}}{\sqrt{2}}$. The technique of plotting $(\sqrt{2}, \sqrt{2})$ clarifies that $r = 2$, $\theta = \pi/4$, so $\sqrt{2} + \mathbf{j} - \mathbf{k} = 2e^{(\mathbf{j}-\mathbf{k})\pi/4\sqrt{2}}$. Now De Moivre's Theorem predicts that $(\sqrt{2} + \mathbf{j} - \mathbf{k})^8 = 256$, which is not exactly obvious on inspection.

(5) The polar form of $1+\mathbf{i}+\sqrt{3}\mathbf{j}+2\mathbf{k}$. The pure part has squared norm $1+3+4 = 8$, so $\mathbf{I} = (\mathbf{i}+\sqrt{3}\mathbf{j}+2\mathbf{k})/2\sqrt{2}$, and by the Argand diagram $r = 3$, $\theta = \cos^{-1}(1/3)$, a first example in which θ is not a rational multiple of π. Later we will use this as an example for rotation in 3-space.



EXAMPLES 9.28 (When quaternions give simpler numbers.) (1) In using quaternions we can sometimes avoid the $\sqrt{}$ signs required in the complex case. A prime example is the cube root of 1, $w = (1/2)(-1+i\sqrt{3})$. By De Moivre's Theorem, any square root of -3 can replace $i\sqrt{3}$, and a convenient choice in quaternions is $\mathbf{i}+\mathbf{j}+\mathbf{k}$, giving $w = (1/2)(-1+\mathbf{i}+\mathbf{j}+\mathbf{k})$, an especially simple form which we probably would not guess, starting from the basic multiplication laws (9.15).

(2) We have $\cos \pi/5 + i \sin \pi/5 = (1 + \sqrt{5})/4 + i\sqrt{(5 - \sqrt{5})/8}$ as complex tenth root of unity, but having one square root within another. Using quaternions we can take $\cos \pi/5 + \mathbf{I} \sin \pi/5 = (1 + \sqrt{5})/4 + \mathbf{i}/2 - \mathbf{j}(1 - \sqrt{5})/4$, involving only single level square roots. More compactly written this is $(1/2)(\tau + \mathbf{i} - \sigma\mathbf{j})$, where (see Table 9.1 in Section 9.1.3) σ,τ are the solutions of $x^2 - x - 1 = 0$. Then $\mathbf{I} = (\mathbf{i} - \sigma\mathbf{j}) / \sqrt{(1 + \sigma^2)}$.

EXERCISE Find the polar forms of (i) $\sqrt{3}+\mathbf{i}-\mathbf{j}+\mathbf{k}$ and (ii) $-\sigma + \mathbf{j} - \tau\mathbf{k}$ [Harder - see (2) above and use the properties of σ,τ given in Table 9.1]. Deduce a fourth root of -1 from question (i).

9.3 Quaternions and rotation

In this section we bring out three ways in which quaternion multiplication provides isometries, the third being the important representation of rotations in 3-space.

9.3.1 Left and right multiplication

This provides a nice application of results on matrices from Chapters 7 and 8. We identify a quaternion $\mathbf{x} = x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$ with the point/vector (x_0, x_1, x_2, x_3) in 4-space $\mathbf{R}^4$ and the standard basis $\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ with respective quaternions $1, \mathbf{i}, \mathbf{j}, \mathbf{k}$, where $\mathbf{e}_0 = (1,0,0,0)$ and so on. Given a quaternion $\mathbf{a}$, the corresponding *left multiplication* transformation $L_\mathbf{a}$ of $\mathbf{R}^4$ sends $\mathbf{x}$ to $\mathbf{ax}$, the *right multiplication* $R_\mathbf{a}$ sends $\mathbf{x}$ to $\mathbf{xa}$. They are linear, for example $L_\mathbf{a}(\mathbf{x}+\alpha\mathbf{y}) = \mathbf{a}(\mathbf{x}+\alpha\mathbf{y}) = \mathbf{ax}+\alpha\mathbf{ay} = L_\mathbf{a}(\mathbf{x}) + \alpha L_\mathbf{a}(\mathbf{y})$ $(\alpha \in \mathbf{R})$, and hence are representable by matrices, where $\mathbf{x} \to \mathbf{x}M_L$ and $\mathbf{x} \to \mathbf{x}M_R$ respectively. The matrix elements may be obtained by a bare-handed approach or by applying the 4-dimensional version of Theorem 8.5 (which systematises the calculation). Thus the rows of M_L correspond to $1, \mathbf{i}, \mathbf{j}, \mathbf{k}$, and the $\mathbf{i}$ row for instance consists of the components of $L_\mathbf{a}(\mathbf{i}) = \mathbf{ai}$. Thus we calculate from (9.15)

$$\mathbf{a}1 = a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}, \quad \mathbf{ai} = a_0\mathbf{i} - a_1 - a_2\mathbf{k} + a_3\mathbf{j},$$
$$\mathbf{aj} = a_0\mathbf{j} + a_1\mathbf{k} - a_2 - a_3\mathbf{i}, \quad \mathbf{ak} = a_0\mathbf{k} - a_1\mathbf{j} + a_2\mathbf{i} - a_3,$$

and this gives M_L below. For M_R we calculate $1\mathbf{a}, \mathbf{ia}, \mathbf{ja}, \mathbf{ka}$.

$$M_L(\mathbf{a}) = \begin{bmatrix} a_0 & a_1 & a_2 & a_3 \\ -a_1 & a_0 & a_3 & -a_2 \\ -a_2 & -a_3 & a_0 & a_1 \\ -a_3 & a_2 & -a_1 & a_0 \end{bmatrix}, \qquad M_R(\mathbf{a}) = \begin{bmatrix} a_0 & a_1 & a_2 & a_3 \\ -a_1 & a_0 & -a_3 & a_2 \\ -a_2 & a_3 & a_0 & -a_1 \\ -a_3 & -a_2 & a_1 & a_0 \end{bmatrix}$$

Considering $M = M_L$, we find that the components a_i are so distributed and signed that the rows of the matrix are mutually orthogonal, and each has length $|\mathbf{a}|$. Equivalently, $MM^T = |\mathbf{a}|^2 I$, where I is the identity matrix. Modifying the argument of Remark 8.9 we have $|\mathbf{x}M|^2 = (\mathbf{x}M)(\mathbf{x}M)^T = \mathbf{x}MM^T\mathbf{x}^T = \mathbf{x}|\mathbf{a}|^2 I\mathbf{x}^T = |\mathbf{a}|^2|\mathbf{x}|^2$. Thus $L_\mathbf{a}$ scales all vectors by a factor $|\mathbf{a}|$, in agreement with $|\mathbf{ax}| = |\mathbf{a}||\mathbf{x}|$ of Theorem 9.25. Now let $\mathbf{a}$ be a unit quaternion, so that $L_\mathbf{a}$ is an isometry and M is orthogonal. What else can we say about M, apart from $M^{-1} = M^T$? Without going into details we define the isometry in $\mathbf{R}^4$ to be direct if $|M|$ is 1 and indirect if it is -1. Is $|M|$ always the same, independently of $\mathbf{a}$?

Calculating $|M|$ directly looks rather formidable, but because the determinant is the product of the eigenvalues, it can be done simply, on the observation that M is skew-symmetric apart from the diagonal terms a_0. That is, $M = a_0 I + S$, where $S^T = -S$. It follows from the eigenvalue equation $|M - \lambda I| = 0$ that the eigenvalues of M are those of S increased by a_0. Now $MM^T = I$ gives $S^2 = (a_0^2-1)I$. This implies that each eigenvalue λ of S satisfies $\lambda^2 = (a_0^2 -1)$, for $\mathbf{x}S = \lambda\mathbf{x} \Rightarrow \mathbf{x}S^2 = \lambda\mathbf{x}S = \lambda^2\mathbf{x}$. Thus S has eigenvalues $\pm i\sqrt{(1 - a_0^2)}$ and those of M are $a_0 \pm i\sqrt{(1 - a_0^2)}$ (assuming $a_0^2 < 1$). Since their sum equals the trace $4a_0$ of M, the + and - signs occur twice each, and $|M| = (a_0^2 - i^2(1-a_0^2))^2 = 1$. In case $a_0^2 = 1$, we have $M = a_0 I$, $|M| = 1$. (Notice that a_0^2 cannot exceed 1. Why?) A similar analysis holds for $M = M_R$.

EXERCISE Verify that M_R is the matrix given, by determining it in the manner described.

Now, why should M have the essentially skew-symmetric form we observed? According to the component formula of Theorem 8.5 we have discovered that, for $\mathbf{a} \in \mathbf{H}$, and $\mathbf{e}_0,\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3$ standing for $1,\mathbf{i},\mathbf{j},\mathbf{k}$ respectively,

$$(\mathbf{a}\mathbf{e}_s).\mathbf{e}_t = \begin{cases} -(\mathbf{a}\mathbf{e}_t).\mathbf{e}_s, & \text{if } s \neq t \\ a_0, & \text{if } s = t \end{cases} \tag{9.33}$$

This is interesting because for $s \neq t$ it is a unified result, not distinguishing between 1 and any of $\mathbf{i}, \mathbf{j}, \mathbf{k}$. It must be a consequence of such a result about multiplication of the basis elements; and here is that result, in particular parts (iii) and (iv).

THEOREM 9.29 Let $\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ *denote* $1,\mathbf{i},\mathbf{j},\mathbf{k}$ *respectively. Then for r,s,t taking values 0,1,2,3 we have*

(i) $\mathbf{e}_r^{-1} = \pm\mathbf{e}_r$, (ii) $\mathbf{e}_r\mathbf{e}_s = \pm\mathbf{e}_t$, for some t,
(iii) $\mathbf{e}_s\mathbf{e}_t^{-1} = -\mathbf{e}_t\mathbf{e}_s^{-1}$ $(s \neq t)$, (iv) $(\mathbf{e}_r\mathbf{e}_s).\mathbf{e}_t = -(\mathbf{e}_r\mathbf{e}_t).\mathbf{e}_s$ $(s \neq t)$.

Proof Everything comes from the multiplication table (9.15), with $1\mathbf{e}_r = \mathbf{e}_r$ assumed. Parts (i) and (ii) may be read straight off that table. For example $\mathbf{j}(-\mathbf{j}) = 1$ so $\mathbf{j}^{-1} = -\mathbf{j}$ (inverses are unique by (9.21)), whereas $1^{-1} = 1$. For (iii), suppose firstly that $\mathbf{e}_s = 1$. Then $\mathbf{e}_t = \mathbf{i}, \mathbf{j}$, or $\mathbf{k}$ so the assertion is $\mathbf{e}_t^{-1} = -\mathbf{e}_t$, which is true by (9.15). Similarly for case $\mathbf{e}_t = 1$. If $s, t \geq 1$ $(s \neq t)$ the result is $\mathbf{ji} = -\mathbf{ij}$ or a cyclic shift of it, hence true by (9.15). Part *(iv)* : From the fact that, as vectors the $\mathbf{e}_i$ satisfy $\mathbf{e}_r.\mathbf{e}_s = \delta_{rs}$ (1 if r = s, else 0) we have the following simple table which establishes (iv).

	Condition	$(\mathbf{e}_r\mathbf{e}_s).\mathbf{e}_t$	$(\mathbf{e}_r\mathbf{e}_t).\mathbf{e}_s$
$\mathbf{e}_r\mathbf{e}_s = \mathbf{e}_t$	i.e. $\mathbf{e}_r = \mathbf{e}_t\mathbf{e}_s^{-1} = -\mathbf{e}_s\mathbf{e}_t^{-1}$	1	-1
$\mathbf{e}_r\mathbf{e}_s = -\mathbf{e}_t$	i.e. $\mathbf{e}_r = -\mathbf{e}_t\mathbf{e}_s^{-1} = \mathbf{e}_s\mathbf{e}_t^{-1}$	-1	1
	$\mathbf{e}_r\mathbf{e}_s \neq \pm\mathbf{e}_t$	0	0

9.3.2 Quaternions and rotation in 3-space

We come now to an important application of quaternions: as an alternative to rotation matrices. The case seems most clear cut in their use for smoothing animation (Section 9.4). Here we explore how quaternions work out in representing rotations in 3-space. The first step is to spot the non-obvious fact that if **x** is a pure quaternion, alias a point in 3-space, then so is $\mathbf{axa}^{-1}$ for *every* quaternion **a**, pure or otherwise. This is not the case for left or right multiplication, so we do not quite have a correspondence with the similar looking isometry combination of Theorem 2.12. For example left multiplication by **i** sends the point **i** = (0,1,0,0) to (-1,0,0,0).

*THEOREM 9.30 (Rotation about unit axis vector **I** through angle 2θ.)*

(a) *If* **x** *is a pure quaternion then so is* $\mathbf{axa}^{-1}$ *for every quaternion* **a**.

(b) *The transformation of 3-space* T: $\mathbf{R}^3 \to \mathbf{R}^3$ *given by*

$$T(\mathbf{x}) = \mathbf{axa}^{-1}, \quad \textit{with } \mathbf{a} = re^{\mathbf{I}\theta}, \quad (r > 0,\ \mathbf{I}^2 = -1),$$

is rotation through angle 2θ about axis vector I based at the origin.

(c) *With* **a** = t + u**i** + v**j** + w*k*, |**a**| = 1, *the matrix for* T(**x**) = **x**M *is*

$$M(\mathbf{a}) = \begin{bmatrix} 1-2v^2-2w^2 & 2tw+2uv & 2uw-2tv \\ 2vu-2tw & 1-2w^2-2u^2 & 2tu+2vw \\ 2tv+2wu & 2wv-2tu & 1-2u^2-2v^2 \end{bmatrix}.$$

(d) $M(\mathbf{a}^{-1}) = M(\mathbf{a})^{-1} = M(\mathbf{a})^T, \qquad M(\mathbf{a})M(\mathbf{b}) = M(\mathbf{ba}) \quad (\mathbf{a},\mathbf{b} \text{ in } \mathbb{H}).$

REMARKS 9.31 (1) In the matrix (c) we follow the notation of Shoemake (1985), though it may be convenient to replace the diagonal elements by their equivalents $\cos 2\theta + 2u^2$, $\cos 2\theta + 2v^2$, $\cos 2\theta + 2w^2$. The matrix is given in a rather different form in Theorem 8.49. To make the transition from one to the other, we replace 2θ here by φ. However, we shall derive the present formula in terms of quaternions, for completeness. Either formula may be preferable, depending on the precise purpose.

(2) Transposing (d) we have $M(\mathbf{b})^T M(\mathbf{a})^T = M(\mathbf{ba})^T$. These transposes are the matrices required if we work with the matrix formulation $T(\mathbf{x}) = M^T\mathbf{x}^T$, as is often done. We emphasise again that if we know the matrices for one formulation then we know them for the other: simply transpose.

Notice that if quaternion multiplication were commutative the map T would be the identity.

EXERCISE Demonstrate the equivalence of the two versions of the rotation matrix, in Theorems 8.49 and 9.30.

Proof of Theorem 9.30 Note first that (a) can be proved by multiplying the appropriate matrices $M_L(\mathbf{a})M_R(\mathbf{a})^T$ from the previous section to obtain the first row and column all zeros except for leading entry 1, which shows that this matrix maps a vector of form (0,a,b,c) into another vector with zero first coordinate, i.e. pure quaternions to pure quaternions. We shall carry out a coordinate free proof, obtaining a most useful formula which exhibits the rotation produced. Without loss of generality we may assume that **a** is a unit quaternion, since $(\lambda\mathbf{a})\mathbf{x}(\lambda\mathbf{a})^{-1} = \mathbf{axa}^{-1}$ for a nonzero real number λ, by (9.21). It will be very helpful to have the following three results before us for pure quaternions (hence vectors) **u**, **v**, **w**.

(i) $\mathbf{uv} = -\mathbf{u}.\mathbf{v} + \mathbf{u} \times \mathbf{v}$, (9.24)

(ii) $\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{w} \times \mathbf{v}) \times \mathbf{u} = (\mathbf{u}.\mathbf{w})\mathbf{v} - (\mathbf{u}.\mathbf{v})\mathbf{w}$ (Theorem 7.35)

(iii) The product $[\mathbf{u},\mathbf{v},\mathbf{w}] = \mathbf{u}.(\mathbf{v} \times \mathbf{w})$ is unchanged if **u**,**v**,**w** shift cyclically, and is zero if any two are equal (Theorem 7.31).

Suppose the polar form is $\mathbf{a} = c + \mathbf{I}s$, where $c = \cos\theta$, $s = \sin\theta$ for some angle θ, and $\mathbf{I}^2 = -1$. In anticipation of (b) let us write T(**y**) for $\mathbf{aya}^{-1}$. Then

$$\begin{aligned} T(\mathbf{y}) &= (c + \mathbf{I}s)\,\mathbf{y}\,(c - \mathbf{I}s) && \text{by (9.30) for } \mathbf{a}^{-1}, \\ &= c^2\mathbf{y} + sc(\mathbf{Iy} - \mathbf{yI}) - s^2\,\mathbf{I}\,\mathbf{y}\,\mathbf{I}, \\ &= c^2\mathbf{y} + 2sc(\mathbf{I} \times \mathbf{y}) - s^2\,\mathbf{I}\,\mathbf{y}\,\mathbf{I}, && \text{by (i) for } \mathbf{Iy},\ \mathbf{yI}. \end{aligned}$$

$$\begin{aligned} \mathbf{I}\,\mathbf{y}\,\mathbf{I} &= (\mathbf{Iy})\mathbf{I} \\ &= -(\mathbf{I}.\mathbf{y})\mathbf{I} + (\mathbf{I} \times \mathbf{y})\,\mathbf{I} && \text{by (i) for } \mathbf{Iy}, \\ &= -(\mathbf{I}.\mathbf{y})\mathbf{I} - (\mathbf{I} \times \mathbf{y}).\mathbf{I} + (\mathbf{I} \times \mathbf{y}) \times \mathbf{I}, && \text{by (i) for } (\mathbf{I} \times \mathbf{y})\mathbf{I} \\ &= -(\mathbf{I}.\mathbf{y})\mathbf{I} + (\mathbf{I}.\mathbf{I})\,\mathbf{y} - (\mathbf{I}.\mathbf{y})\,\mathbf{I}, && \text{by (ii)}, \end{aligned}$$

since $(\mathbf{I} \times \mathbf{y}).\mathbf{I} = 0$ by (iii). Finally, $\mathbf{I}.\mathbf{I} = |\mathbf{I}|^2 = 1$ simplifies the expression to

$$\mathbf{I}\,\mathbf{y}\,\mathbf{I} = \mathbf{y} - 2(\mathbf{I}.\mathbf{y})\mathbf{I}, \tag{9.34}$$

Now we substitute (9.34) in the expression so far for T(**y**), noting that $c^2 - s^2 = \cos 2\theta$ and $2sc = \sin 2\theta$, to obtain a formula that will be useful again,

$$T(\mathbf{y}) = \mathbf{y}\cos 2\theta + (\mathbf{I} \times \mathbf{y})\sin 2\theta + (1 - \cos 2\theta)(\mathbf{I}.\mathbf{y})\,\mathbf{I} \tag{9.35}$$

where **y**, **I** are pure and $\mathbf{I}^2 = -1$. Since $\mathbf{I} \times \mathbf{y}$ is also pure, so is T(**y**), as we wished to show.

(b) There are various ways to show that T is rotation through 2θ about axis vector **I**. We make maximum use of the classification of isometries. *T is an isometry* because the definition (8.1) is satisfied as follows

$$\begin{aligned} |T(\mathbf{x}) - T(\mathbf{y})| &= |\mathbf{axa}^{-1} - \mathbf{aya}^{-1}| && (\mathbf{x}, \mathbf{y} \text{ pure}) \\ &= |\mathbf{a}(\mathbf{x}-\mathbf{y})\mathbf{a}^{-1}| \\ &= |\mathbf{a}|\,|\mathbf{x}-\mathbf{y}|\,|\mathbf{a}^{-1}| && \text{by Theorem 9.25,} \\ &= |\mathbf{x}-\mathbf{y}| && \text{as } |\mathbf{a}^{-1}| = 1/|\mathbf{a}|. \end{aligned}$$

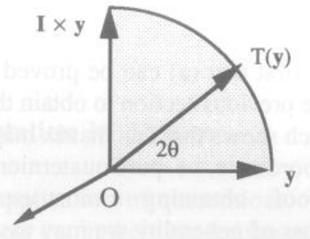


Figure 9.9 Effect of the transformation $T(\mathbf{y}) = \mathbf{aya}^{-1}$, with $\mathbf{a} = e^{\mathbf{I}\theta}$, $\mathbf{I}^2 = -1$.

Let $\mathbf{y}$ be a vector perpendicular to I, so that $\mathbf{I.y} = 0$, and from (9.35) we have as illustrated in Figure 9.9

$$T(\mathbf{y}) = \mathbf{y}\cos 2\theta + (\mathbf{I}\times\mathbf{y})\sin 2\theta. \qquad (9.36)$$

This shows that T has the same effect as rotation 2θ about axis **I**, on the plane Π through O normal to **I**. By the Classification Theorem 8.42, T is either the rotation claimed or a rotary reflection in Π. But $T(\mathbf{I}) = \mathbf{a\,I\,a}^{-1} = e^{\mathbf{I}\theta}\,\mathbf{I}\,e^{-\mathbf{I}\theta} = e^{\mathbf{I}\theta}\,e^{-\mathbf{I}\theta}\,\mathbf{I} = \mathbf{I}$, so T must be the rotation. Thus the key result is established. (Alternatively, T acts as the rotation on four non-collinear points: the point with position vector **I** and any three non-collinear points in Π. Hence result (b) by the earlier Theorem 8.22.)

(c) We are computing the 3 x 3 submatrix $M = [m_{rs}]$ of $N = M_L(\mathbf{a})M_R(\mathbf{a})^T$ corresponding to the last three rows and columns, numbered 1,2 ,3, from

$$m_{rs} = (\text{row } r \text{ of } M_L(\mathbf{a}))\,.\,(\text{row } s \text{ of } M_R(\mathbf{a})) \qquad (*).$$

What is more obvious from $m_{rs} = (\mathbf{ae}_r\mathbf{a}^{-1})\,.\,\mathbf{e}_s = f(a_0, a_1, a_2, a_3)$, say, is that $m_{r+1,s+1} = f(a_0, a_2, a_3, a_1)$, where $1\to 2\to 3\to 1$ (0 fixed), and that therefore we may obtain the second and third rows of M from the first (this being found from (*)). For example $m_{12} = 2a_1a_2 + 2a_0a_3$ implies that $m_{23} = 2a_2a_3 + 2a_0a_1$. Representing the diagonal we have $m_{11} = a_1^2 + a_0^2 - a_3^2 - a_2^2 = t^2 + u^2 - v^2 - w^2$, which equals $1 - 2v^2 - 2w^2$ (the stated expression) because of $1 = |\mathbf{a}|^2 = t^2 + u^2 + v^2 + w^2$. [The fact that N is orthogonal and sends vectors (0,a,b,c) to (0,d,e,f) implies that its first row and column have the form (1,0,0,0), though we do not need this.]

(d) The matrix result $M(\mathbf{ab}) = M(\mathbf{b})M(\mathbf{a})$ holds, since $(\mathbf{ab})\mathbf{x}(\mathbf{ab})^{-1} = \mathbf{a}(\mathbf{bxb}^{-1})\mathbf{a}^{-1}$ implies that $M(\mathbf{b})$ is applied first. The first statement of (d) now follows from the equalities $M(\mathbf{a}^{-1})M(\mathbf{a}) = M(\mathbf{aa}^{-1}) = M(1) = I$ (the identity matrix), because a 1-sided inverse is 2-sided by (7.17).

EXERCISE Prove that $\mathbf{y}, T(\mathbf{y}), \mathbf{I}$ is a right-handed triple in part (b) above, if $0 < 2\theta < \pi$.

EXERCISE Complete the calculation of $M(\mathbf{a})$ from $M_L(\mathbf{a})M_R(\mathbf{a})^T$.

REMARKS 9.32 The reader may wish to postpone these comments, logically placed here, until after seeing the examples that follow. (1) Sometimes in the literature one finds $T(\mathbf{x}) = \mathbf{a}^{-1}\mathbf{xa}$, $\mathbf{a} = e^{\mathbf{I}\theta}$. Since $\mathbf{a}^{-1} = e^{-\mathbf{I}\theta}$ the result is rotation, still about axis **I**, but through -2θ rather than 2θ. There may be applications for which this sign reversal perhaps does not matter, but for computer graphics it seems risky to let in a gratuitous minus sign. Therefore we stick to $T(\mathbf{x}) = \mathbf{axa}^{-1}$.

(2) We emphasise that in contrast with (i), replacing **a** by r**a** for any $r \neq 0$ leaves T unchanged, since $(r\mathbf{a})\mathbf{x}(r\mathbf{a})^{-1} = r\mathbf{ax}r^{-1}\mathbf{a}^{-1} = \mathbf{axa}^{-1}$. This is especially useful if obtaining **a** of unit modulus requires division by a square root.

(3) Again, replacing θ, **I** by −θ, −**I** leaves the rotation unchanged, since $e^{(-\theta)(-\mathbf{I})} = e^{\mathbf{I}\theta}$. This is illustrated in Section 8.4.3. On the other hand, replacing **a** by its conjugate or **I** by its negative reverses the turn.

(4) The unit quaternion **a** gives a 1/n'th turn if and only if **a** has order 2n (is a 2n'th root of unity and nothing less). Reason: rotation $2\pi/n$ requires argument π/n.

To encourage our faith that quaternions accomplish rotations, we start with several examples in which the answer is easy to check.

EXAMPLE 9.33 Calculate the effect on the x-axis, of rotation by π/2 about the z-axis, using quaternions directly. We use $T(\mathbf{x}) = \mathbf{axa}^{-1}$. In the usual notation $2\theta = \pi/2$, $\mathbf{I} = \mathbf{k}$, and $\mathbf{a} = \cos\pi/4 + \mathbf{k}\sin\pi/4 = (1+\mathbf{k})/\sqrt{2}$. Since $|\mathbf{a}| = 1$ we have $\mathbf{a}^{-1} = \bar{\mathbf{a}} = (1-\mathbf{k})/\sqrt{2}$, and so

$$T(\mathbf{i}) = \tfrac{1}{2}(1+\mathbf{k})\,\mathbf{i}\,(1-\mathbf{k}) = \tfrac{1}{2}(\mathbf{i}+\mathbf{k})\,(\mathbf{i}+\mathbf{j}) = \tfrac{1}{2}(\mathbf{i}-\mathbf{ik}+\mathbf{j}-\mathbf{jk}) = \mathbf{j}.$$

Therefore T maps the x-axis into the y-axis, vindicating Remark (1) above. See Figure 9.10 on the right. Also $T(\mathbf{k}) = \mathbf{k}$ (no calculation required - why?) and $T(\mathbf{j}) = -\mathbf{i}$, which confirms by the classification of isometries that we have the correct rotation, since it is correct on four noncoplanar points

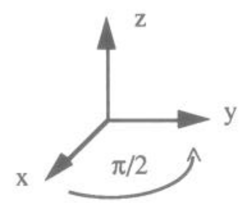


Figure 9.10

EXAMPLE 9.34 We calculate the matrix for the Example above by the formula of Theorem 9.30(c), with $\mathbf{a} = (1+\mathbf{k})/\sqrt{2} = t + u\mathbf{i} + v\mathbf{j} + w\mathbf{k}$, giving $t = 1/\sqrt{2} = w$, $u = v = 0$. Hence, in agreement with the rotation formula of (7.26), the matrix is M_1 below.

$$M_1 = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad M_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

EXAMPLE 9.35 We use quaternions to find the matrix for rotation about axis vector (1,1,1) through a 1/3 turn. For this we have $2\theta = 2\pi/3$, $\theta = \pi/3$, $\mathbf{I} = (\mathbf{i}+\mathbf{j}+\mathbf{k})/\sqrt{3}$, which gives simply $\mathbf{a} = (1+\mathbf{i}+\mathbf{j}+\mathbf{k})/2$, $t = u = v = w = 1/2$. By Theorem 9.30(c), the matrix is M_2 above. This sends x-axis → y-axis → z-axis cyclically, as we knew it ought. Indeed, given this, we could write down the matrix straightaway by Theorem 8.5. Notice that **a** is a cube root of 1.

TABLE 9.3 A compendium of quaternions **a** for rotation $T(x) = axa^{-1}$
For axis **I** replace **k** by **I** except in case (*)

Rotation	θ	axis	quaternion a	|a|
1/2 turn	π/2	z-axis	**k**	1
1/3 turn	π/3	z-axis	1 + √3**k**	2
1/4 turn	π/4	z-axis	1 + **k**	√2
1/5 turn	π/5	τ**j** + **k** (*)	τ + **j** - σ**k**	2
1/6 turn	π/6	z-axis	√3 + **k**	2

(*) Refers to a 1/5 turn about OA with A(0,1,-σ) a vertex of an icosahedron - the regular solid bounded by twenty equilateral triangular faces, five at each of the 12 vertices. Here (see Table 9.1) $\tau = (1+\sqrt{5})/2 = 2\cos\pi/5$ and $\sigma = (1-\sqrt{5})/2$ are the roots of $x^2-x-1 = 0$. Hence $\sigma+\tau = 1$, $\sigma\tau = -1$, $\sigma^2+\tau^2 = 3$, $\sigma^2 = 2-\tau$ and $\tau^2 = 2-\sigma$. The points $(0, \pm\tau, \pm 1)$ and their cyclic shifts serve as vertices. $\sin\pi/5$ cannot be expressed conveniently in the way $\cos\pi/5$ can.

9.3.3 Composition of rotations, by quaternions

EXAMPLE 9.36 We find the composition of a 60° rotation about the y-axis followed by a 60° rotation about the x-axis, without using matrices. Reversing order as prescribed, the appropriate product of quaternions is

$$\frac{\sqrt{3}+\mathbf{i}}{2}\cdot\frac{\sqrt{3}+\mathbf{j}}{2} \;=\; \frac{3}{4} + \frac{1}{4}(\mathbf{i}\sqrt{3} + \mathbf{j}\sqrt{3} + \mathbf{k}) \;=\; \cos\theta + \mathbf{I}\sin\theta,$$

hence the result is a rotation of $2\cos^{-1}(3/4) = 83^o$, about an axis vector $(\sqrt{3}, \sqrt{3}, 1)$.

EXAMPLE 9.37 We shall combine rotation symmetries,

$$R_{OE}(1/2)R_{OA}(1/3),$$

of the regular tetrahedron, shown in Figure 9.11. This solid is bounded by four equilateral triangles, three at each of the 4 vertices. Inspection shows that the rotation symmetries are as listed below, with origin O at the centre:

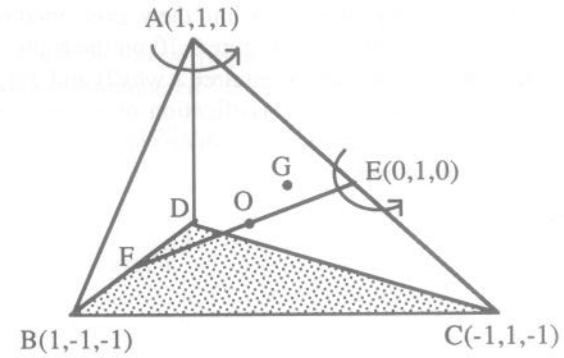


Figure 9.11 The regular tetrahedron. Its vertices are (±1, ±1, ±1) with none or two minus signs.

(1)	a 1/2 turn about each line EF joining the midpoints of opposite edges	3
(2)	a 1/3 and a 2/3 turn about OA for each vertex A	8
(3)	The identity rotation (do nothing)	1
		12

The product of any two of these symmetries must be a third. We require the result of a 1/2 turn about OE followed by a 1/3 turn about OA. Using Table 9.3 as an aid, we find that suitable rotation quaternions are for $R_{OE}(1/2)$: **a** = **j**, for $R_{OA}(1/3)$: **b** = 1 + (√3)<u>OA</u> /|<u>OA</u>| = 1+**i**+**j**+**k**. By Theorem 9.30(d), a rotation quaternion for $R_{OE}(1/2)R_{OA}(1/3)$ is **ba** = (1+**i**+**j**+**k**)**j** = -1-**i**+**j**+**k** = -1 - √3(<u>OB</u>/|<u>OB</u>|). From Table 9.3 (note that **a** and -**a** give the same rotation), the result is a 1/3 turn about OB. Equivalently a 2/3 turn about OG, where G is the centroid of triangle ADC.

EXERCISE Find $R_{OE}(1/2)R_{OB}(1/3)$ in the tetrahedron of Example 9.37. Check your answer by its effect on three suitable points besides the origin, or by using matrices.

EXAMPLE 9.38 The cube of Figure 9.12 contains a symmetrically placed copy of the tetrahedron ABCD of Example 9.37, whose symmetries are thus also symmetries of the cube. We find the product of the cube rotation symmetries indicated:

$$R_{OZ}(1/4)\,R_{OH}(-1/3).$$

An appropriate quaternion product is, by Table 9.3, (1-√3(<u>OH</u> / |<u>OH</u>|) x (1+**k**) = (1-**i**-**j**+**k**)(1+**k**) = 2(-**i**+**k**), giving a 1/2 turn about OL.

For reflections, cf. Example 8.25.

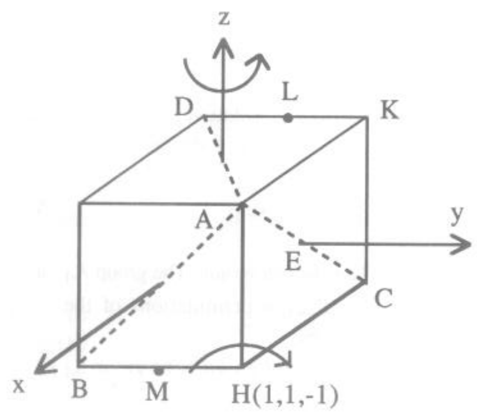


Figure 9.12 The cube, centred at the origin, with vertices (±1, ±1, ±1), and inscribed tetrahedron ABCD from Figure 9.11.

EXERCISE Compute $R_{OY}(3/4)R_{OK}(1/3)$, above. Is this a symmetry of the tetrahedron?

A fund of rotation examples

By rotation *about a point* A of a polyhedron (solid bounded by plane faces) we will mean rotation about axis OA, where as always O is the origin. The *rotation group* of a figure is the group of all its rotational symmetries. For the *cube* of Example 9.38, symbolised by {4,3} because its faces are squares, three at a vertex, the rotation group consists of the identity and all powers of: a 1/2 turn about the midpoint of each edge, a 1/3 turn about each vertex, a 1/4 turn about each *face centre* (centre of face). Total 24. The face centres of a cube form the vertices of its 'dual', the *octahedron* {3,4} with four regular triangular faces at each vertex. Its symmetries are therefore those of a cube, but with face and vertex switched in the description.

The *icosahedron* {3,5} is a solid bounded by twenty regular triangular faces, five at each of the twelve vertices. Rotation symmetries come from: a 1/2 turn about the mid point of each edge, a 1/3 turn about each face centre, and a 1/5 turn about each vertex. A total of sixty in the group. As before we may take the face centres as vertices of a 'dual', this time the *dodecahedron* {5,3}, with regular pentagonal faces, three at a vertex. The five solids enumerated are called the *Platonic solids*, and exhaust the possibilities for a regular polyhedron {p,q} that is *convex*, meaning that any line segment joining two points in the solid is also within the solid.

Table 9.4 gives the unit quaternions for each of the three distinct rotation groups, and a geometrical description of them in terms of *permutations* of certain subfigures. A permutation of a list of objects, say denoted by 1,2, .., n, is a reordering of them, and is called *even* or *odd* according as it requires an even or odd number of *transpositions* (i.e.

Then for example 1,2,3 may be cyclically shifted to 3,1,2 by transposing 2,3 then 1,3, so is even. The group of all permutations of n objects is called the *symmetric group* S_n, and its subgroup of all even permutations is the *alternating group* A_n.

For further information, see Coxeter (1973) who shows that such a polyhedral group, the rotation group of a polyhedron, has order twice the number of edges in the polyhedron's boundary. The reader may know the Euler polyhedron formula $V-E+F = 2$ (cf. Table 9.4), holding for a wide variety of surfaces bounded by polygonal faces. For its application to Solid modelling in computer graphics, see Baumgart (1974) and Mäntylä (1988). A useful introduction is found in Foley et al (1990).

TABLE 9.4 Rotation groups of the Platonic solids, and associated quaternions
V, E, F denote the number of vertices, edges, and faces

Solid	Rotation group	V	E	F	Corresponding group of unit quaternions
Tetrahedron	*Tetrahedral group.* The group A_4 of all even permutations of the vertices.	4	6	4	**1, i, j, k**, $1 \pm \mathbf{i} \pm \mathbf{j} \pm \mathbf{k}$, and their negatives.
Cube Octahedron	*Octahedral group.* The group S_4 of all permutations of the main diagonals of the cube.	8 6	12 12	6 8	The above, & $(1\pm\mathbf{i})/\sqrt{2}$, $(1\pm\mathbf{j})/\sqrt{2}$, $(1\pm\mathbf{k})/\sqrt{2}$, $(\mathbf{i}\pm\mathbf{j})/\sqrt{2}$, $(\mathbf{j}\pm\mathbf{k})/\sqrt{2}$, $(\mathbf{k}\pm\mathbf{i})/\sqrt{2}$, & negatives.
Icosahedron, Dodecahedron	*Icosahedral group.* The group A_5 of all even permuations of five regular tetrahedra on the dodecahedron's 5 x 4 vertices.	12 20	30 30	20 12	The tetrahedral ones, & $(\pm t \pm u\mathbf{i} \pm v\mathbf{j} \pm w\mathbf{k})/2$, with tuvw an even permutation of $-1\ 0\ \sigma\ \tau$.

EXERCISE What group of rotations comes from quaternion group $Q = \{\pm 1, \pm\mathbf{i}, \pm\mathbf{j}, \pm\mathbf{k}\}$?

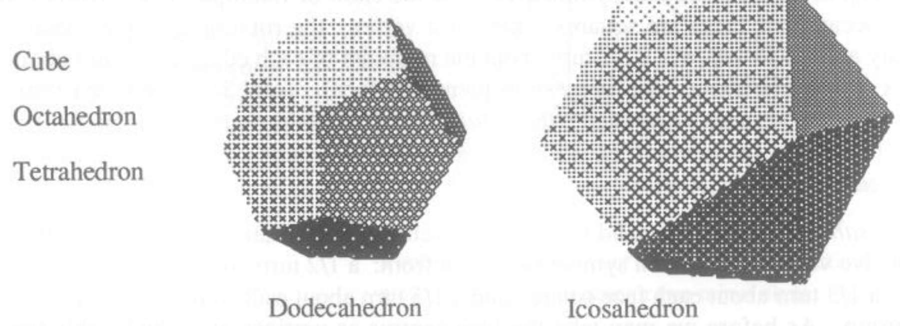


Figure 9.13 The five Platonic solids

A connection with groups We give now a small amount of information concerning a wider context. For further reading on quaternions, see Coxeter (1974) and references therein. Just as the unit pure quaternions can be regarded as the points of the unit 2-sphere S^2, in 3-space, so can the unit quaternions be viewed as forming the 3-dimensional sphere in 4-space:

$$S^3 = \{(a_0, a_1, a_2, a_3): \text{ each } a_i \text{ is real}, \ \Sigma a_i^2 = 1\},$$

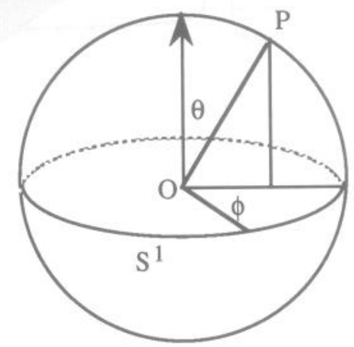
in which the pure quaternion subset S^2 is the "equatorial" subsphere defined by $a_0 = 0$. (In Section 9.4 we consider how to take a smooth path through a sequence of points on S^3 to achieve animation 'in-betweening'). But there is more. Since the product of two unit quaternions is a third, multiplication of quaternions gives S^3 the structure of an infinite group, with identity from the real number 1 (Example 9.26). We have a map

$$F: S^3 \to SO(3),$$

which sends $\pm\mathbf{a}$ to the same matrix $M(\mathbf{a})^T$. And F is a *group homomorphism,* meaning that $F(\mathbf{ab}) = F(\mathbf{a})F(\mathbf{b})$. This is simply $M(\mathbf{ab})^T = M(\mathbf{a})^T M(\mathbf{b})^T$. It is described as *2:1 and onto* because every matrix in SO(3) is the image of exactly two members of S^3. We have just introduced the Platonic solids, whose rotation groups G are finite subgroups of SO(3). The corresponding sets of unit quaternions in Table 9.4 are finite subgroups of S^3, namely the inverse images $F^{-1}(G)$; collectively they are called the *binary* polyhedral groups $2A_4$, $2S_4$, $2A_5$, of orders 24, 48, 120.

Coordinates on the n-sphere We build up from the circle S^1 to S^2 to S^3, and similarly beyond, so that S^n is coordinatised by n angles. For the equatorial circle S^1 in S^2 we have a typical point $\mathbf{u} = (\cos\phi, \sin\phi)$, $0 \le \phi \le 2\pi$. Now take basis vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \ldots$ each viewed as belonging to a higher dimensional space as required. Then we extend S^1 to S^2, adding a further coordinate on basis vector $\mathbf{e}_3$, obtaining the points of the ordinary sphere S^2 as $\mathbf{v} = \mathbf{u}\sin\theta + \mathbf{e}_3\cos\theta$, $(0 \le \theta \le \pi)$. In coordinates,

$$\mathbf{v} = (\cos\phi \sin\theta, \sin\phi \sin\theta, \cos\theta).$$



The sphere S^2 with equator S^1.

To coordinatise S^3 we simply apply the same idea again with new basis vector $\mathbf{e}_4$ and angle say ψ, where $0 \le \psi \le \pi$ (we might have chosen $\theta_1, \theta_2, \theta_3$). Then S^3 has points

$$\mathbf{w} = \mathbf{v}\sin\psi + \mathbf{e}_4\cos\psi = (\cos\phi \sin\theta \sin\psi, \sin\phi \sin\theta \sin\psi, \cos\theta \sin\psi, \cos\psi).$$

9.4 Quaternion in-betweening

Background references for this Section are given at the end of the Chapter.

9.4.1 Why in-betweening and why quaternions?

A prime case in which interpolation or 'in-betweening' is required, is that of computer animation. The aspect we address here is that of animation in 3-space with no change in the actual shape of a moving object. A series of *key frames* are established for giving the position and orientation of an object at certain time intervals, and we wish to generate a suitable sequence of intermediate states so that the object will appear to the eye to move smoothly between and past key frames. A key frame may be specified by a 4 by 4 matrix as in Section 8.3.3 describing the position and orientation of an object in 3-space, relative to a chosen origin and three mutually perpendicular coordinate axes. Since we are dealing here with strictly rigid motion, only direct isometries are allowed, which, according to the classification of Section 8.5.4 are

Translation,
Rotation, or
Screw,

where a screw consists of rotation combined with translation parallel to the axis.

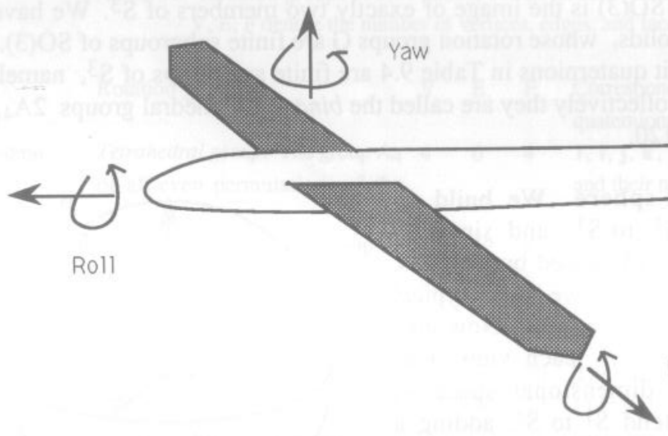


Figure 9.14 Yaw, pitch and roll for an aircraft.

But why quaternions? We cannot interpolate directly beween transformation matrices because their entries are partially dependent, for example a 3 x 3 rotation matrix is subject to six quadratic relations. The traditional set of independent coordinates describing rotation in three dimensions has been Euler angles, introduced successfully by Euler (1758) to solve differential equations, and still useful in this regard (see e.g. Miller (1972)). Euler angles are an extension of the idea of latitude and longitude on the ordinary 2-sphere. They specify rotation as the composition of three independent rotations about given axes through an origin, in a given order. A typical system used in aircraft dynamics (Figure 9.14) is to specify in order:

1. Yaw, or heading, around a vertical axis,
2. Pitch, around a horizontal axis through the wings,
3. Roll, around an axis along the fuselage.

This is not the only system in use. The reader may care to list 12 = 3 x 2 x 2 viable alternate systems with mutually perpendicular axis in the object. See Hughes (1986). On the other hand a quaternion represents orientation as a single rotation, enabling a much simpler approach to interpolation in particular. This is probably the most important gain in using quaternions. The complications of Euler angles for animation are further discussed in Shoemake (1985).

REMARK Robotics is another area where in-betweening is required. See for example the tutorial of Heise and Macdonald (1989).

9.4.2 Interpolating between two orientations

Clearly there is no problem in applying linear interpolation to the translation parts of successive placings of an object in space. Here we address the question of a suitable way to interpolate between two orientations as unit quaternions, that is, as points on the unit sphere S^3 in 4-space.

As we will soon show, any two distinct points **a**, **b** on S^3 lie on a unique circle in this sphere, called the *great circle* through **a**, **b**, divided by these points into two *great arcs* (if one is the shorter we call it *the* great arc between **a**, **b**). Since the points of this circle may be parametrised by angle in the usual way, it is natural to use *spherical linear interpolation* between **a** and **b** as indicated in Figure 9.15: if the angle between vectors **a**, **b** is θ then the interpolated point **q**(t) for time t, where $0 \le t \le 1$, is given by the vector at angle tθ to **a**.

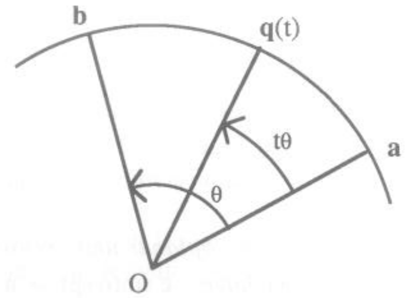


Figure 9.15 Spherical interpolation along a great arc **a b**.

A formula is simple to obtain once we appreciate that **a**, **b** are on an honest circle of radius 1 in an ordinary 2-dimensional plane Π, which happens to be sitting in 4-space. This works as follows. The 4-vectors **a**, **b** define a unique plane Π through the origin consisting by definition of all linear combinations of **a**, **b**,

$$\Pi = \{\lambda\mathbf{a} + \mu\mathbf{b} : \lambda, \mu \text{ are arbitrary real numbers}\}.$$

Angles in Π are defined by innerproducts of its vectors calculated from their coordinates as 4-vectors. But how do we know this is like the usual x-y plane?

APPROACH 1 By similar arguments to the 3-dimensional case, Theorem 8.55, there exists a 4 by 4 orthogonal matrix M sending unit vectors **a**,**b** to unit vectors **a'**,**b'** with last two coordinates zero and with **a'**.**b'** = **a**.**b**. Also M sends $\lambda\mathbf{a} + \mu\mathbf{b}$ to $\lambda\mathbf{a}' + \mu\mathbf{b}'$. Thus geometry in Π is the same as that in

$$\Pi' = \lambda(a_1', a_2', 0\ 0) + \mu(b_1', b_2', 0\ 0): \quad \lambda, \mu \text{ real},$$

and hence the same as a standard plane Σ based on 2 coordinates with, say (a'_1, a'_2) defining one axis and a perpendicular vector (c'_1, c'_2) the other, as in Figure 9.16.

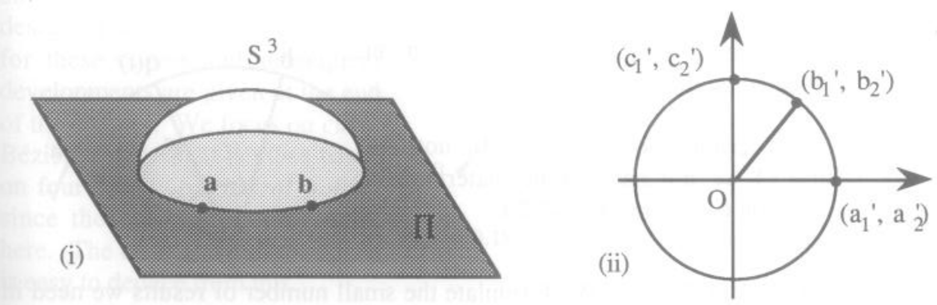


Figure 9.16 (i) Plane Π intersects sphere S^3 in a great circle, (ii) the plane Σ.

APPROACH 2 Π itself contains a unit vector **c** perpendicular to **a**, so that **a**,**c** define axes for x,y coordinates in Π. To obtain **c** we subtract from **b** its component along **a** and make the result into a unit vector. Thus $\mathbf{c} = \mathbf{v}/|\mathbf{v}|$, where $\mathbf{v} = \mathbf{b} - (\mathbf{b.a})\mathbf{a}$. As a check, $\mathbf{v.a} = \mathbf{b.a} - (\mathbf{b.a})\mathbf{a.a} = 0$, since $\mathbf{a.a} = 1$. Using $\mathbf{a.a} = 1 = \mathbf{c.c}$ and $\mathbf{a.c} = 0$ we calculate the inner product in Π (inherited from $\mathbf{R}^4$) as $(x_1\mathbf{a} + x_2\mathbf{c}).(y_1\mathbf{a} + y_2\mathbf{c}) = x_1y_1 + x_2y_2$. Thus Π is a standard plane.

The great circle C through **a**, **b** is by definition the intersection of Π and the sphere S^3, namely the points of Π constituting its unit circle about the origin. We have now justified Figure 9.15, and laid the foundation for later arguments here and in Section 9.4.4. The formula quoted in Shoemake (1985) and Heise and Macdonald (1989) may now be derived simply by an argument in the plane where, if **a**, **b**, **c** are unit vectors at given angles α, β as shown in Figure 9.17 then **c** is a linear combination $\lambda\mathbf{a}+\mu\mathbf{b}$, determined by α, β alone, and this relation continues to hold if the plane consists of 4-vectors as does Π above.

LEMMA 9.39 For coplanar unit vectors ***a, b, c****, forming nonzero angles α, β as shown in Figure 9.17, we have* $\mathbf{c}\sin(\alpha+\beta) = \mathbf{a}\sin\beta + \mathbf{b}\sin\alpha$.

Proof We find the linear relationship $\mathbf{c} = \lambda\mathbf{a} + \mu\mathbf{b}$ by viewing **a**, **b**, **c** as complex numbers (as we may, by Approach 2), when $\mathbf{a} = \mathbf{c}e^{-i\alpha}$, $\mathbf{b} = \mathbf{c}e^{i\beta}$, and so $\mathbf{c} = \lambda\mathbf{c}e^{-i\alpha} + \mu\mathbf{c}e^{i\beta}$ or, dividing through by **c** as a complex number :

$$1 = \lambda(\cos\alpha - \mathbf{i}\sin\alpha) + \mu(\cos\beta + \mathbf{i}\sin\beta).$$

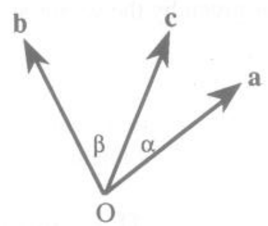


Figure 9.17 Coplanar unit vectors ($\alpha, \beta \neq 0$).

Equating real and imaginary parts (NB: the real part of the left hand side must equal the real part of the right hand side and similarly for the imaginary parts), we obtain two equations for the unknowns λ, μ:

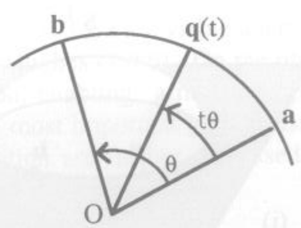
$$\lambda\cos\alpha + \mu\cos\beta = 1, \qquad -\lambda\sin\alpha + \mu\sin\beta = 0.$$

To complete the proof it remains to solve these easy equations and to apply the relation $\sin(\alpha+\beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta$.

COROLLARY 9.40 The quaternion obtained by spherical linear interpolation from ***a*** *to* ***b*** *on the unit sphere S^3 in 4-space, with parameter t $(0 \le t \le 1)$, where* $\mathbf{a.b} = \cos\theta$, *is given by*

$$\mathbf{q}(t) = \frac{\sin(1-t)\theta}{\sin\theta}\mathbf{a} + \frac{\sin t\theta}{\sin\theta}\mathbf{b} \qquad (\sin\theta \neq 0) \qquad (9.37)$$

Proof We apply Lemma 9.39 with $\alpha = \theta t$, $\beta = \theta - \theta t$. (Figure 9.15 is repeated on the right).



REMARK 9.41 Concerning the case $\sin\theta = 0$, note that $\sin\theta = 0 \Leftrightarrow \cos\theta = \pm 1 \Leftrightarrow \mathbf{b} = \pm\mathbf{a} \Leftrightarrow$ unit quaternions **a**, **b** give the same rotation (see Remark 9.32(2)).

DIFFERENTIATING VECTORS We formulate the small number of results we need in terms of plane vectors, the extension to higher dimensions being straightforward. Suppose the position vector $\mathbf{r}(t) = (x(t), y(t))$ of a point P varies with a parameter t, possibly time. With the usual convention of a dot to denote differentiation with respect to t, we define the derivative of **r** as $\dot{\mathbf{r}} = (\dot{x},\dot{y})$ (we say **r** is *differentiable* when $\dot{x}$ and $\dot{y}$ exist). Then $\dot{\mathbf{r}}$ is a tangent to the curve traced out by P since it represents the direction in which **r** is changing as t varies. Further, the length $|\dot{\mathbf{r}}|$ represents the rate at which **r** is changing - how fast P moves along the curve for a given rate of change of t. The following Lemma will suffice.

LEMMA 9.42 If vectors **u**, **v**, *and scalar a(t) are differentiable then*
(a) $(d/dt)\,\mathbf{u.v} = \mathbf{u.\dot{v}} + \mathbf{\dot{u}.v}$ (b) $(d/dt)\,a\mathbf{v} = a\dot{\mathbf{v}} + \dot{a}\mathbf{v}$

Proof (a) $(d/dt)\,\mathbf{u.v} = (d/dt)(u_1v_1 + u_2v_2) = u_1\dot{v}_1 + \dot{u}_1v_1 + u_2\dot{v}_2 + \dot{u}_2v_2 = \mathbf{u.\dot{v}} + \mathbf{\dot{u}.v}$ (on regrouping the terms). The second part is slighter shorter.

The classic example is the unit circle with $1 = \mathbf{r.r}$, hence $2\mathbf{r.\dot{r}} = 0$ and the tangent is at right angles to the radius, as confirmed by $\mathbf{r.\dot{r}} = (\cos t, \sin t).(-\sin t, \cos t) = 0$. Now, from (9.37), with **a**,**b** constant, so that $\dot{\mathbf{a}} = \dot{\mathbf{b}} = 0$, we have the derivative of **q**(t) as

$$\dot{\mathbf{q}}(t) = \frac{\theta\cos t\theta}{\sin\theta}\mathbf{b} - \frac{\theta\cos(1-t)\theta}{\sin\theta}\mathbf{a} \qquad (\sin\theta \neq 0) \qquad (9.38)$$

EXAMPLE 9.43 Show that $|\dot{\mathbf{q}}(t)|^2 = \theta^2$ (which is what we would expect).

Solution We have $|\dot{\mathbf{q}}(t)|^2 = \dot{\mathbf{q}}(t).\dot{\mathbf{q}}(t) =$

$$= \left(\frac{\theta\cos t\theta}{\sin\theta}\right)^2\mathbf{b.b} - 2\left(\frac{\theta\cos t\theta}{\sin\theta}\right)\left(\frac{\theta\cos(1-t)\theta}{\sin\theta}\right)\mathbf{a.b} + \left(\frac{\theta\cos(1-t)\theta}{\sin\theta}\right)^2\mathbf{a.a}.$$

Setting $\mathbf{a.a} = \mathbf{b.b} = 1$, $\mathbf{a.b} = \cos\theta$, expanding $\cos(\theta - t\theta)$, and collecting terms, we obtain $\theta^2[\cos^2 t\theta - \cos^2\theta\cos^2 t\theta + \sin^2\theta\sin^2 t\theta]/\sin^2\theta$. The bracketed expression simplifies to $\sin^2\theta$, and we are done.

EXERCISE Verify that $|\mathbf{q}(t)| = 1$ by direct calculation as in Example 9.43.

9.4.3 Bézier curves

Paul Bézier produced his famous curves in response to a need in the design of car bodies. References for these curves and subsequent developments are given at the end of the chapter. We focus on *cubic* Bézier curves, that is those based on four *knots*, or *control points*, since these are what we require here. The case of n control points is easy to deduce from this.

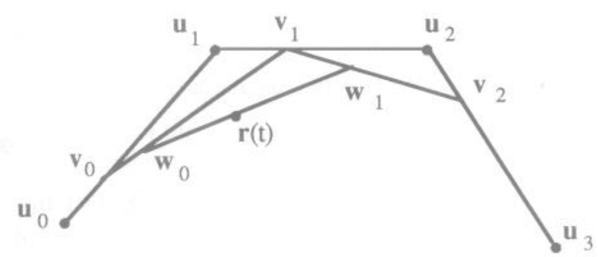


Figure 9.18 Geometrical construction of Bezier curve with four knots.