Quaternions and Rotations in 3-Space:
How it Works

Vernon Chi
Microelectronic Systems Laboratory
Department of Computer Science
University of North Carolina at Chapel Hill
25 September 1998

Informal summary

Think of a quaternion Q as a vector augmented by a real number to make a four element entity. It has a *real* part $Q\rfloor_{re}$ and a *vector* part $Q\rfloor_{ve}$. If $Q\rfloor_{re}$ is zero, Q represents an ordinary vector; if $Q\rfloor_{ve}$ is zero, it represents an ordinary real number. In any case, the ratio between the real part and the *magnitude* of the vector part $|Q\rfloor_{ve}|$ plays an important role in rotations, and is conveniently represented by the parameter $\phi = tan^{-1}(|Q\rfloor_{ve}|/Q\rfloor_{re})$. A unit magnitude quaternion Q_u has a Pythagorean distance of 1 over its four elements, and its product with any vector S_v gives another vector having the same magnitude as S_v but rotated in space. If $S_v \perp Q_u\rfloor_{ve}$, the rotation is by an angle ϕ about the vector $Q_u\rfloor_{ve}$ (or simply about Q_u). If S_v is arbitrary, however, certain cross-terms of the product spoil this convenient relationship. Even in this general case however, these cross-terms cancel in the triple product $U_v = Q_u S_v Q_u^{-1}$, where $Q_u^{-1} \equiv 1/Q_u$. The rotations of the two successive products are in the same direction, so U_v represents a rotation of S_v about $Q_u\rfloor_{ve}$ by an angle 2ϕ, which depends only on Q_u. Thus, the operation $Q_u S_v Q_u^{-1}$ performs a rotation of S_v which is entirely characterized by the unit quaternion Q_u. The rotation occurs about an axis parallel to Q_u by an amount $2\,tan^{-1}(|Q\rfloor_{ve}|/Q\rfloor_{re})$. Quaternion notation conveniently handles composition of any number of successive rotations into one equivalent rotation: $Q_u = Q_1 Q_2 \cdots Q_n$ where each unit quaternion Q_i represents one of the succession of rotations. Other operations useful in inertial navigation problems are also presented.

1   Historical background

Quaternions were devised by Sir William Hamilton in his extensions of vector algebras to satisfy the properties of division rings (roughly, quotients exist in the same domain as the operands). In [1], Art.112, Hamilton notes, "...that *for the complete determination, of what we have called the geometrical* QUOTIENT *of two Co-initial Vectors, a System of Four Elements, admitting each separately of numerical expression, is generally required.* ... we have already a *motive* for saying, that 'the Quotient of two Vectors is generally a Quaternion.' "

Quaternions can also be considered to be an extension of classical algebra into the hypercomplex number domain D, satisfying a property that $|p|^2 \cdot |q|^2 = |p \cdot q|^2$ for $(p, q) \in D$ [2]. This domain consists of symbolic expressions of n terms with real coefficients where n may be 1 (real numbers), 2 (complex numbers), 4 (quaternions), 8 (Cayley numbers), but no other possible values (proved by Hurwitz in 1898). Thus, quaternions also share many properties with complex numbers.

While Hamilton provides geometrical interpretations of various proved properties throughout [1], the development itself is fundamentally algebraic, that is, based on the properties of a particular axiomatic set of symbolic operations. The geometric properties of quaternions are nevertheless sweeping, the composition of successive rotations through successive multiplications being just one, albeit an important one.

2   Axiomatic properties of quaternions

Quaternions are defined as sums of 4 terms of the form $Q = 1 \cdot q_1 + i \cdot q_2 + j \cdot q_3 + k \cdot q_4$ where q_1, q_2, q_3, q_4 are reals, 1 is the multiplicative identity element, and i, j, k are symbolic elements having the properties:

$$i^2 = -1, \; j^2 = -1, \; k^2 = -1,$$
$$ij = k, \; ji = -k,$$
$$jk = i, \; kj = -i,$$
$$ki = j, \; ik = -j.$$

Customarily, the extension of an algebra should attempt to preserve the properties of the operators defined in the original algebra. Generalizing from the classical algebra of real and complex numbers to quaternions motivates the following operator rules.

2.1   Addition of quaternions

The addition rule for quaternions is component-wise addition:

$$P+Q = (p_1+ip_2+jp_3+kp_4)+(q_1+iq_2+jq_3+kq_4) = (p_1+q_1)+i(p_2+q_2)+j(p_3+q_3)+k(p_4+q_4).$$

This rule preserves the associativity and commutivity properties of addition, and provides a consistent behavior for the subset of quaternions corresponding to real numbers, i.e.,

$$P_r + Q_r = (p + 0i + 0j + 0k) + (q + 0i + 0j + 0k) = p + q.$$

2.2   Multiplication of quaternions

The multiplication rule for quaternions is the same as for polynomials, extended by the multiplicative properties of the elements i, j, k given above. Written out for close inspection, we have:

$$
\begin{aligned}
PQ &= (p_1 + ip_2 + jp_3 + kp_4)(q_1 + iq_2 + jq_3 + kq_4) \\
&= (p_1 q_1 - p_2 q_2 - p_3 q_3 - p_4 q_4) + i(p_1 q_2 + p_2 q_1 + p_3 q_4 - p_4 q_3) \\
&\quad + j(p_1 q_3 + p_3 q_1 + p_4 q_2 - p_2 q_4) + k(p_1 q_4 + p_4 q_1 + p_2 q_3 - p_3 q_2).
\end{aligned}
$$

A term-wise inspection reveals that commutivity is not preserved. Associativity is preserved, however, the proof being left to the reader. And as desired for the subset of reals, $P_r Q_r = pq$.

2.3   Conjugates of quaternions

Consistent with complex numbers, let us define the *conjugate* operation on a given quaternion Q to be,

$$
\overline{Q} = \overline{(q_1 + iq_2 + jq_3 + kq_4)} \equiv (q_1 - iq_2 - jq_3 - kq_4).
$$

As with complex numbers, note that both $(Q + \overline{Q})$ and $(Q\overline{Q})$ are real. Moreover, if we define the absolute value or *norm* of Q to be,

$$
|Q| = \sqrt{q_1^2 + q_2^2 + q_3^2 + q_4^2}, \quad \text{where the positive root is assumed,}
$$

then apparently $Q\overline{Q} = \overline{Q}Q = |Q|^2$. The conjugate operation is distributive over addition, that is, $\overline{P + Q} = \overline{P} + \overline{Q}$. With respect to multiplication however, $\overline{PQ} = \overline{Q}\,\overline{P}$, the proof of which is left as an exercise to the reader.

3   Other properties of quaternions

The axioms in the previous section completely *define* quaternions in terms of the desired properties under three basic operations. Many other properties may be proved.

3.1   General properties

Mathematically, the most important property is that the quaternions form a division ring (i.e., quaternion quotients exist).

3.1.1   Division of quaternions

Since multiplication is not commutative, let us derive both a *left quotient* Q_L^{-1} and a *right quotient* Q_R^{-1} by defining the symbolic expression P/Q to be solutions of the following two identities,

$$
QQ_L^{-1} = P, \qquad\qquad Q_R^{-1}Q = P.
$$

Multiplying both sides of these identities respectively on the left and right by $\overline{Q}/|Q|^2$ we have immediately,

$$Q_L^{-1} = \frac{\overline{Q}P}{|Q|^2}, \qquad Q_R^{-1} = \frac{P\overline{Q}}{|Q|^2}.$$

Thus in general two distinct quotients will occur, however in the special case where $P = 1$, we have by definition the multiplicative *inverse* of a quaternion,

$$Q_L^{-1} = Q_R^{-1} = Q^{-1} = \frac{\overline{Q}}{|Q|^2}$$

3.1.2   Quaternion multiplication is distributive over addition

A term-wise expansion of $P(Q + R) = PQ + PR$ proves this property and is left as an exercise for the reader.

3.1.3   Unit quaternions

The subspace Q_u of *unit quaternions* which satisfy the condition $|Q_u| = 1$ have some important properties. A trivially apparent one is,

$$Q_u^{-1} = \overline{Q}_u.$$

A less obvious, but very useful one is,

$$Q_u = R_u cos\,\phi + P_u sin\,\phi = cos\,\phi + P_u sin\,\phi,$$

where $R_u = (1, 0, 0, 0)$ is a real unit quaternion, $P_u = (0, ip_2, jp_3, kp_4)$ is a vector unit quaternion parallel to the vector part of Q_u, and ϕ is an arbitrary real number. The proof is straightforward:

$$|Q_u|^2 = Q_u\overline{Q}_u = (R_u cos\,\phi + P_u sin\,\phi)\overline{(R_u cos\,\phi + P_u sin\,\phi)}$$
$$= R_u\overline{R}_u cos^2\phi + (R_u\overline{P}_u + P_u\overline{R}_u)sin\,\phi\,cos\,\phi + P_u\overline{P}_u sin^2\phi$$
$$= cos^2\phi + sin^2\phi = 1.$$

At this time, let's interpret ϕ as simply quantifying the ratio of the real part to the magnitude of the vector part of a quaternion. Its geometrical representation as specifying an angle of rotation will be presented later.

3.2   Vector properties of quaternions

The quaternion $Q = (q_1 + iq_2 + jq_3 + kq_4)$ can be interpreted as having a real part q_1, and a vector part $(iq_2 + jq_3 + kq_4)$, where the elements $\{i, j, k\}$ are given an added *geometric* interpretation as unit vectors along the x, y, z axes, respectively. Accordingly, the subspace $Q_r = (q_1 + 0i + 0j + 0k)$ of *real quaternions* may be regarded as being equivalent to the real numbers, $Q_r = q$. Similarly, the subspace $Q_v = (0 + iq_2 + jq_3 + kq_4)$ of *vector quaternions* may be regarded as being equivalent to the ordinary vectors, $Q_v = \boldsymbol{q} \equiv (iq_x + jq_y + kq_z)$.

3.2.1 Products of real quaternions

The product of real quaternions is real, and the operation is commutative:

$$P_r Q_r = pq = qp = Q_r P_r.$$

Moreover, the operation is associative:

$$(P_r Q_r)R_r = (pq)r = p(qr) = P_r(Q_r R_r).$$

3.2.2 Product of a real quaternion with a vector quaternion

The product of a real and a vector quaternion is a vector, and the operation is commutative:

$$P_r Q_v = (0 + p_1 q_2 i + p_1 q_3 j + p_1 q_4 k) = (0 + q_2 p_1 i + q_3 p_1 j + q_4 p_1 k) = Q_v P_r.$$

3.2.3 Parallel and perpendicular quaternions

Let's call quaternions P and Q *parallel* $(P \parallel Q)$ if their *vector parts* $P{\rfloor}_{ve} = (P - \overline{P})/2$ and $Q{\rfloor}_{ve} = (Q - \overline{Q})/2$ are parallel; i.e., if $(R - \overline{R}) = 0$, where $R = P{\rfloor}_{ve} Q{\rfloor}_{ve}$. Similarly, let's call them *perpendicular* $(P \perp Q)$ if $P{\rfloor}_{ve}$ and $Q{\rfloor}_{ve}$ are perpendicular; i.e. if $(R + \overline{R}) = 0$.

3.2.4 Products of vector quaternions

The product of two vector quaternions has the remarkable property,

$$P_v Q_v = -(p_2 q_2 + p_3 q_3 + p_4 q_4) + (p_3 q_4 - p_4 q_3)i + (p_4 q_2 - p_2 q_4)j + (p_2 q_3 - p_3 q_2)k$$
$$= -\boldsymbol{p} \cdot \boldsymbol{q} + \boldsymbol{p} \times \boldsymbol{q},$$

where the "·" and "×" operators are respectively the "dot" and "cross" products of classical vector algebra. This is clearly a general quaternion except in two special cases: if $P_v \parallel Q_v$ the product is a real quaternion equal to $-\boldsymbol{p} \cdot \boldsymbol{q}$ and if $P_v \perp Q_v$ the product is a vector quaternion equal to $\boldsymbol{p} \times \boldsymbol{q}$.

3.2.5 Product of a unit quaternion and a perpendicular vector quaternion

Properties of this curiously specialized case are useful in understanding how quaternions can be used to rotate vectors in 3-space. Let S_v be a vector quaternion, Q_u be a unit quaternion, and $S_v \perp Q_u$. Then according to section 3.1.3, we can write,

$$T = Q_u S_v = (cos\,\phi + sin\,\phi\,P_u)S_v = cos\,\phi\,S_v + sin\,\phi\,P_u S_v,$$

where $P_u \parallel Q_u$. The first term is a vector $T_{v(1)} \parallel S_v$. Since $S_v \perp P_u$, the second term must also be a vector $T_{v(2)}$; moreover $T_{v(2)} \perp S_v$ and $T_{v(2)} \perp Q_u \parallel P_u$. Since the product T is a sum of vectors it must also be a vector, $T = T_v$. Both $T_{v(1)}$ and $T_{v(2)}$ lie in a

plane perpendicular to Q_u. Thus $T_v = T_{v(1)} + T_{v(2)}$ can be geometrically interpreted as a rotation of S_v by an angle ϕ in this plane, i.e., about an axis parallel to Q_u.

Now consider the product,

$$U_v = T_v Q_u^{-1} = T_v \overline{Q}_u = cos\,\phi\,T_v + sin\,\phi\,T_v \overline{P}_u = cos\,\phi\,T_v - sin\,\phi\,T_v P_u.$$

The vector identity $T_v P_u = -P_u T_v$ can be used to rewrite this as,

$$U_v = cos\,\phi\,T_v + sin\,\phi\,P_u T_v,$$

which is another rotation of angle ϕ about Q_u. The rotation ϕ is in the same sense for these two products, so the operation

$$U_v = Q_u S_v Q_u^{-1}$$

performs a rotation of S_v about Q_u by an angle 2ϕ.

3.3   General rotations in 3-space; Reference frames

In section 3.2.5 we saw how the operation $Q_u S_v Q_u^{-1}$ rotated a *perpendicular* vector S_v about a unit quaternion Q_u. Now let's consider how this operation behaves with an *arbitrary* vector V_v. We can decompose $V_v = W_v + S_v$ where $W_v \parallel Q_u$ and $S_v \perp Q_u$. Then,

$$Q_u V_v Q_u^{-1} = Q_u(W_v + S_v)Q_u^{-1} = Q_u W_v Q_u^{-1} + Q_u S_v Q_u^{-1} = Q_u W_v Q_u^{-1} + U_v,$$

where U_v is S_v rotated about Q_u by an angle 2ϕ. To evaluate the first term, note that since $W_v \parallel Q_u$ we can write $W_v = zP_u$, where z is a real number and unit vector $P_u \parallel Q_u$. Thus,
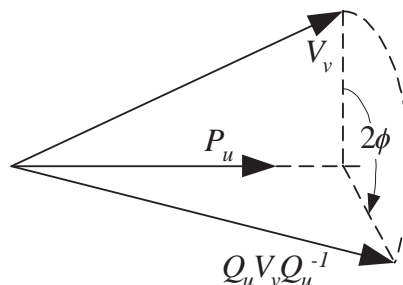
$$Q_u W_v Q_u^{-1} = Q_u z P_u Q_u^{-1} = z Q_u P_u Q_u^{-1} = z P_u Q_u Q_u^{-1} = z P_u = W_v.$$

That $Q_u P_u = P_u Q_u$ is left as an exercise to the reader. Finally then, we have:

$$Q_u V_v Q_u^{-1} = W_v + U_v.$$

Geometrically, we interpret this as a rotation of V_v about Q_u by an angle of 2ϕ.

Figure 1:

Arbitrary vector V_v is rotated by unit quaternion Q_u about a unit vector $P_u \parallel Q_u$, through angle 2ϕ.

This operation performs the same rotation on *all* vectors including the unit vectors of a coordinate system. Therefore, it can be used to rigidly transform the coordinates of any *reference frame* into a new frame of different orientation. This is a very useful property.

3.4　Composition of successive rotations

Let Q_1 and Q_2 be two unit quaternions representing arbitrary rotations in 3-space as described in section 3.3. Applying them in succession to a vector V_v,

$$Q_2(Q_1 V_v Q_1^{-1})Q_2^{-1} = (Q_2 Q_1)V_v(Q_1^{-1} Q_2^{-1}) = (Q_2 Q_1)V_v(Q_2 Q_1)^{-1} = Q_i V_v Q_i^{-1},$$

where the unit quaternion $Q_i = Q_2 Q_1$ is the successive composition of two rotations. This property generalizes to the composition of any number of rotations. In this reverse order composition, each successive rotation is relative to the *initial* reference frame as is illustrated in Figure 2a.
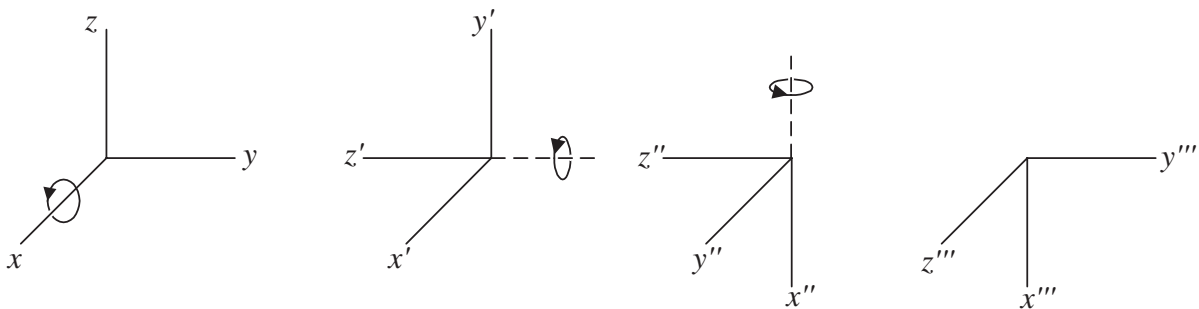

Figure 2a: 90° rotations of a reference frame about the initial x, y, z axes, respectively

Composing a rotation in the forward order, $Q_c = Q_1 Q_2 \ldots$, has the effect of performing each successive rotation relative to its *current* reference frame, illustrated in Figure 2b.
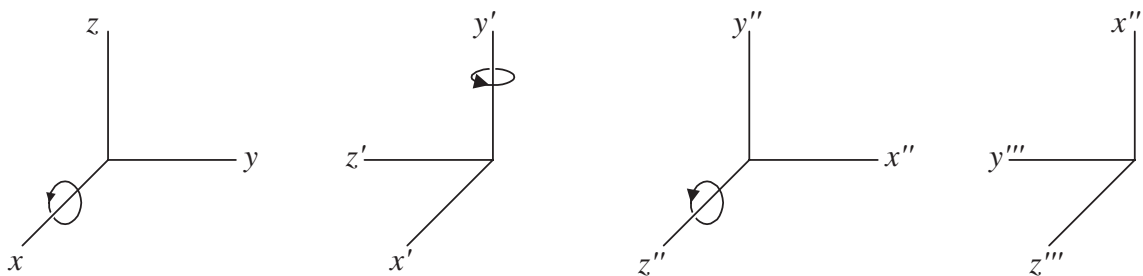

Figure 2b: 90° rotations of a reference frame about its current x, y, z axes, respectively.

4　Strapdown inertial navigation system (INS) applications

Usage of quaternions by this branch of engineering is common, but the notation often differs in some respects from the above, and a more detailed annotation is provided to relate variables to reference frames. Specifically in this section, I'll follow the notation used in Titterton and Weston [3]. Next I'll introduce this notation, then derive expressions for some of the commonly used operations for INS engineering.

4.1   Frames and coordinates

It is often convenient to represent the same physical situation in a number of different frames of reference which may differ by displacement, rotation, and system of coordinates. Each frame comprises a complete definition of these parameters. A privileged, *inertial* family of frames are those in which physical objects experience no inertial forces.

Cartesian coordinate systems, while not necessary, are generally used as coordiante systems of the frames discussed in [3]. The non-scalar data types used are vectors, matrices, and quaternions. Distinct from the data types, are the kinds of variables treated, i.e., positions, linear velocities, and angular rates.

4.2   Superscripts and subscripts

Superscripts and subscripts are used to associate certain attributes of a variable with coordinate frames. On a gross level, the notation is consistent, but there are fine nuances, depending on the kind of the variable but not its type.

Superscripts are used consistently for all kinds of variables. S^i indicates that the variable S is expressed in the coordinates of the ith frame.

The subtle nuances in the interpretation of subscripts is illustrated next for each of the three kinds of variable.

4.2.1   The position variable X_j^i

X_j^i represents the position of a point relative to the origin of the jth frame, expressed in the coordinates of the ith frame. In most cases $i = j$, and it is common to use implicit notations. X_j and X^j both represent X_j^j, where the choice of super- or subscript depends on what is being emphasized.

4.2.2   The velocity variable V_j^i

The variable V_j^i represents a velocity taken relative to the jth frame, expressed in coordinates of the ith frame. The velocity in any frame is not dependent on the location of the origin of the frame; rather it may be taken relative to the velocity of *any fixed point* in that frame. Just as for position variables, $V_j = V_j^j$ is implied.

4.2.3   The angular rate variable Ω_{jk}^i

The variable Ω_{jk}^i represents an angular rate of rotation of the kth entity relative to the jth frame, expressed in coordinates of the ith frame. Just as for velocities, the location of origin of reference frame j is not relevant; rather the angular rate is taken relative to the angular rate of any fixed point in the jth frame. Often, the kth entity is another frame, so this notation conveniently expresses the angular rate of rotation of the kth frame relative to the jth.

4.3   A vector representation of a rotation

While representing a rotation as a unit quaternion provides useful mathmatical properties, representing it as a vector is more compact, as well as being intuitively straightforward. Let the vector $\boldsymbol{a} = a\hat{\boldsymbol{a}}$ represent a rotation where its unit vector $\hat{\boldsymbol{a}}$ is aligned with the axis of rotation and its magnitude a is equal to the angle of rotation. From this we can uniquely construct a unit quaternion,

$$A = cos(a/2) + sin(a/2)\hat{\boldsymbol{a}},$$

such that $AR_v\overline{A}$ performs a rotation of R_v about $\hat{\boldsymbol{a}}$ by an angle equal to a.

4.4   Time derivative of a rotation quaternion

Assume a b-frame that is rotating with respect to a reference n-frame. At any instant, let the unit quaternion Q represent a rotation of a constant vector C^b in the b-frame into a vector $C^n = QC^b\overline{Q}$ in the n-frame. Since this rotation progresses continuously in time, $Q = Q(t)$ has a time derivative $\dot{Q}$ which we now derive.

Applying the derivative of products rule to C^n, we have

$$\dot{C}^n = \dot{Q}\,C^b\,\overline{Q} + Q\,C^b\,\dot{\overline{Q}} = \dot{Q}\,C^b\,\overline{Q} + \overline{\dot{Q}\,\overline{C}^b\,\overline{Q}} = \dot{Q}\,C^b\,\overline{Q} - \overline{\dot{Q}\,C^b\,\overline{Q}}.$$

In the vector formulation of classical mechanics [4], a vector $\boldsymbol{p}$ is used to represent an instantaneous rate of rotation, $\dot{\boldsymbol{c}} = \boldsymbol{p} \times \boldsymbol{c}$, where $\boldsymbol{c}$ is an arbitrary vector, and $\dot{\boldsymbol{c}}$ is its variation with time. In the n-frame, a quaternion formulation of this equation is,

$$\dot{C}^n = (P^n C^n - \overline{P^n C^n})/2.$$

Since $\boldsymbol{c}$ is arbitrary, this equation can be applied to an entire coordinate system, and we can represent the rate of rotation of the b-frame in the n-frame as $P^n = P_{nb}^n$.

Equating the expressions for $\dot{C}^n$, we have, $\dot{Q}C^b\overline{Q} = P_{nb}^n C^n/2 = P_{nb}^n(QC^b\overline{Q})/2$, or $\dot{Q} = P_{nb}^n Q/2$. It is often the case that the rotational rate is measured in the rotating b-frame, so we can substitute the identity $P_{nb}^n = QP_{nb}^b\overline{Q}$, to obtain

$$\dot{Q} = QP_{nb}^b/2.$$

4.5   Interpolation between rotations

Given two arbitrary rotations Q_{10}, Q_{20} from the 0-frame to the 1 and 2-frames respectively, geometric intuition would suggest an interpolation between them would be along the single rotation Q_{21} taking the 1-frame into the 2-frame. In fact, this can be visualized as a great circle on a unit 4-sphere which connects the images of Q_{10} and Q_{20}. This great circle lies in a plane normal to $Q_{21}\rfloor_{ve}$. The locus of points lying between Q_{10} and Q_{20} on the great circle corresponds to a rotational angle of between 0 and $cos^{-1}(Q_{21}\rfloor_{re})$.

Now $Q_{20} = Q_{10}Q_{21} \Rightarrow Q_{21} = \overline{Q_{10}}Q_{20}$. Let $Q_{21} = cos(\phi_{21}) + \hat{\boldsymbol{u}}_{21}sin(\phi_{21})$, whence we can calculate $\phi_{21} = cos^{-1}(Q_{21}\rfloor_{re})$ and $\hat{\boldsymbol{u}}_{21} = Q_{21}\rfloor_{ve}/sin(\phi_{21})$.

Given $\phi_{x1} \ni (0 \leq \phi_{x1} \leq \phi_{21})$ we construct $Q_{x1} = cos(\phi_{x1}) + \hat{\boldsymbol{u}}_{21}sin(\phi_{x1})$, from which we calculate the interpolated rotation,

$$Q_{x0} = Q_{10}Q_{x1}.$$

5   References

[1] Sir William Rowan Hamilton, "Elements of Quaternions," Third Edition, Chelsea Publishing Co., New York, 1963.

[2] I. L. Kantor and A. S. Solodovnikov, "Hypercomplex Numbers," English translation, Springer-Verlag, New York, 1989.

[3] D. H. Titterton and J. L. Weston, "Strapdown inertial navigation technology," Peter Peregrinus, Ltd., IEE, Stevenage, UK, 1997.

[4] Herbert Goldstein, "Classical Mechanics," Addison-Wesley, Reading MA, 1950, pp. 132–134.