Angle between principal axis triples

Walter Tape¹ and Carl Tape²

¹Department of Mathematics, University of Alaska, Fairbanks, Alaska 99775, USA
²Geophysical Institute and Department of Geology & Geophysics, University of Alaska, Fairbanks, Alaska 99775, USA. E-mail: carltape@gi.alaska.edu

Accepted 2012 August 21. Received 2012 August 19; in original form 2012 May 14

SUMMARY

The principal axis angle $\xi_0$, or Kagan angle, is a measure of the difference between the orientations of two seismic moment tensors. It is the smallest angle needed to rotate the principal axes of one moment tensor to the corresponding principal axes of the other. This paper is a conceptual review of the main features of $\xi_0$. We give a concise formula for calculating $\xi_0$, but our main goal is to illustrate the behaviour of $\xi_0$ geometrically. When the first of two moment tensors is fixed, the angle $\xi_0$ between them becomes a function on the unit ball. The level surfaces of $\xi_0$ can then be depicted in the unit ball, and they give insights into $\xi_0$ that are not obvious from calculations alone. We also include a derivation of the known probability density $p_{\xi_0}$ of $\xi_0$. The density $p_{\xi_0}(t)$ is proportional to the area of a certain surface $L_w(t)$. The easily seen variation of $L_w(t)$ with $t$ then explains the rather peculiar shape of $p_{\xi_0}$. Because the curve $p_{\xi_0}$ is highly non-uniform, its shape needs to be considered when analysing distributions of empirical $\xi_0$ values. We recall an example of Willemann which shows that $\xi_0$ may not always be the most appropriate measure of separation for moment tensor orientations, and we offer an alternative measure.

Key words: Theoretical seismology.

1 INTRODUCTION

The seismic moment tensor is a mathematical expression of an earthquake source. In matrix terms, it is a $3 \times 3$ symmetric matrix; its eigenvalues express the size and type of the earthquake source, and its eigenvectors express the orientation. When comparing moment tensors, some quantitative measure of separation between them is often desirable. A natural measure is the ordinary euclidean distance in the space of $3 \times 3$ matrices. But while the euclidean distance is the best overall measure, it leaves something to be desired conceptually, since it does not reveal the separate contributions of the size, type, and orientation. Some supplementary measures are therefore desirable as well. Differences in size and type, that is, differences in eigenvalue triples, are relatively easy to measure. It is less obvious what should constitute a reasonable measure of difference in moment tensor orientation.

In this paper we review one commonly used measure, which we call the principal axis angle and denote by $\xi_0$. The principal axis angle $\xi_0$ between two moment tensors is the smallest angle required to rotate the principal axes of one of them into the corresponding principal axes of the other. In Section 3 we give a more precise definition of $\xi_0$ as well as a concise formula for its calculation (eq. 34). In Sections 4 and 7 we give some intuition for $\xi_0$ by depicting it as a function on the unit ball. In Section 5 and Appendix A we give a derivation of the known probability density $p_{\xi_0}$ for $\xi_0$, the underlying assumption being that the moment tensor orientations are random. We give a geometric characterization of $p_{\xi_0}$ that helps to explain its distinctive and highly non-uniform nature. We remind readers that the distinctive shape of $p_{\xi_0}$ should be kept in mind when interpreting $\xi_0$ values for real data sets. In Section 8 we suggest that $\xi_0$ may not always be the best measure of separation for moment tensor orientations, and we offer an alternative measure.

Although we have just described $\xi_0$ as a measure of difference in moment tensor orientation, the only relevant feature of the moment tensor is its orthorhombic symmetry, which is embodied in its principal axes. Orthorhombic symmetry is the symmetry of a brick, or of a three-coloured cube whose opposite faces are coloured the same. The change in orientation of any object having orthorhombic symmetry can be measured by $\xi_0$.

The angle $\xi_0$ has probably seen its widest application and most thorough study in connection with crystalline textures, where it is just one of many ‘misorientation angles’ each associated with a symmetry group (e.g. Sutton & Balluffi 1995; Morawiec 2004). Mackenzie & Thomson (1957) introduced the concept of misorientation angle in connection with cubic symmetry and used Monte Carlo methods to approximate its probability density function $p_{\xi_0}$. Mackenzie (1958) and Handscomb (1958) each gave an analytic expression for $p_{\xi_0}$, again for cubic symmetry. Grimmer (1979a,b) found $p_{\xi_0}$ for other symmetry groups. Grimmer gave no derivation but said that he had used the methods of Handscomb (1958). Later Morawiec (1995), using ideas of Frank (1988) as well as
Handscomb (1958), published a derivation of $p_0$ and gave examples for many symmetry groups.

In seismology the angle $\xi_0$ has been studied by Kagan (e.g. 1991, 2007), and in fact in some literature the angle is referred to as the Kagan angle. The angle $\xi_0$ has been used in the statistical analysis of real and synthetic sets of moment tensors, faults, and stress (e.g. Kagan 1992; Hardebeck & Shearer 2002; Hardebeck 2006). Hardebeck, for example, was able to infer unexpected homogeneity in earthquake faulting in California. The angle has also been used to compare moment tensors from different earthquakes (e.g. Okal et al. 2011). And it has been used in comparing moment tensors computed for the same events using different techniques (e.g. Liu et al. 2004), the principal axis angle $\xi_0$ is apt to find increased use in measuring moment tensor differences. In spite of its simple description as a minimum rotation angle, the angle $\xi_0$ has some subtle features that users should be aware of. The purpose of this paper is to provide a conceptual exposition of $\xi_0$ that will elucidate those features.

## 2 Rotation Matrices

This section reviews some prerequisites involving rotation matrices. Most of the results are stated without proof. Wikipedia is a sufficient reference.

### 2.1 Rotation axis and rotation angle

A rotation matrix is a $3 \times 3$ orthogonal matrix $U$ with determinant equal to 1. That is,

$$UU^T = I,$$

$$\det U = 1,$$

where $U^T$ is the transpose of $U$ and where $I$ is the identity matrix. If $U$ is considered as a linear transformation of $\mathbb{R}^3$, then eq. (1a) says that $U$ preserves distances, and eq. (1b) says that $U$ preserves handedness.

A typical rotation matrix $U$ has two normalized rotation axes and two rotation angles. If $\xi$ is the rotation axis associated with the rotation matrix $U$, then the effect of $U$ is to rotate each point in $\mathbb{R}^3$ through angle $\xi$ about the axis $\mathbf{u}$, with positive rotation being clockwise. If $\xi$ is the rotation axis associated with the rotation matrix $U$, then $-\xi$ is the rotation axis associated with the rotation matrix $-U$.

We will always choose the rotation axis $\mathbf{u}$ and associated rotation angle $\xi$ so that $0 \leq \xi \leq \pi$. Then, $\mathbf{u}$ (normalized) and $\xi$ satisfy

$$\cos \xi = \frac{1 + \text{trace} U}{2}, \quad 0 \leq \xi \leq \pi,$$

$$\mathbf{u} \cdot \mathbf{u} = 1, \quad (\mathbf{u} \cdot \mathbf{e}) \geq 0,$$

where $\mathbf{e}$ is a non-zero vector orthogonal to $\mathbf{u}$. Eq. (2a) uniquely determines $\xi$. If $U$ is neither the identity nor a $180^\circ$ rotation, then eqs (2b) and (2c) uniquely determine $\mathbf{u}$, eq. (2b) determines $\mathbf{u}$ up to sign, and then eq. (2c) chooses between the two sign possibilities, by requiring that the triple $\mathbf{u}$, $\mathbf{e}$, $U\mathbf{e}$ be right-handed. If $U$ is a $180^\circ$ rotation, then it has two normalized rotation axes $\pm \mathbf{u}$, both satisfying eqs (2). (The left-hand side of eq. (2c) is then zero.) And if $U$ is the identity, then any unit vector $\mathbf{u}$ satisfies eqs (2) with $\xi = 0$.

Eqs (2b) and (2c) nicely express the conceptual meaning of the rotation axis, but they have the disadvantage that they require $U$ to be truly orthogonal, not just approximately so. Appendix F has another approach to calculating the rotation axis.

### 2.2 Quaternion summary

A unit quaternion has the form

$$q = (w, x, y, z), \quad w^2 + x^2 + y^2 + z^2 = 1,$$

where $w, x, y, z$ are real numbers. We also use the notations

$$q = (w, v) = (w, (x, y, z)),$$

$$q = w + xi + yj + zk.$$

To multiply quaternions, one writes them in the latter form and then proceeds as one might hope, except that the multiplication is not commutative, and

$$i^2 = j^2 = k^2 = ijk = -1.$$

For unit quaternions $q$, the inverse is the conjugate. That is,

$$q^{-1} = \bar{q},$$

where $(w, v) = (w, -v)$.

There is a two-to-one correspondence between unit quaternions and rotation matrices. The rotation matrix associated with the unit quaternion $q = (w, x, y, z)$ is

$$U(q) = \begin{pmatrix}
  w^2 + x^2 - y^2 - z^2 & 2(xy - wz) & 2(wy + xz) \\
  2(xy + wz) & w^2 - x^2 + y^2 - z^2 & 2(yz - wx) \\
  2(xz - wy) & 2(wx + yz) & w^2 - x^2 - y^2 + z^2
\end{pmatrix}. $$

From eq. (7),

$$U(q_1)U(q_2) = U(q_2q_1).$$

The unit quaternions for a rotation $U$ are $\pm q$, where

$$q = \left(\frac{\cos \xi}{2}, \left(\begin{pmatrix}
  \sin \frac{\xi}{2}
\end{pmatrix}
\right) \mathbf{u}\right),$$

and where $\xi$ and $\mathbf{u}$ are the rotation angle and normalized rotation axis for $U$. The rotation angle for a rotation $U$ is therefore given by

$$\frac{\xi}{2} = |w|,$$

where $(w, x, y, z)$ is a unit quaternion for $U$.

The $180^\circ$ rotations $X_\pi$, $Y_\pi$, $Z_\pi$ about the $x, y, z$ coordinate axes play an important role in what follows (e.g. Fig. 3). From eq. (7), their respective unit quaternions are $\pm i$, $\pm j$, and $\pm k$. If $q = w + xi + yj + zk$ is a unit quaternion for $U$, then unit quaternions for $U$, $UX_\pi$, $UY_\pi$, $UZ_\pi$ are, respectively,

$$q = w + xi + yj + zk,$$

$$q_i = (w + xi + yj + zk)i = -x + wi + zj - yk.$$

© 2012 The Authors, GJI, 191, 813–831

Geophysical Journal International © 2012 RAS
\[\begin{align*}
q_j &= (w + xi + yj + zk)j = -y - zi + wj + xk, \\
q_k &= (w + xi + yj + zk)k = -z + yi - xj + wk.
\end{align*}\]

From eq. (10) the rotation angles \(\xi, \xi, \xi, \xi\) for \(U, UX_x, UY_y, UZ_z\), are therefore given by
\[
\begin{align*}
\cos \frac{\xi}{2} &= |w|, \\
\cos \frac{\xi}{2} &= |x|, \\
\cos \frac{\xi}{2} &= |y|, \\
\cos \frac{\xi}{2} &= |z|.
\end{align*}
\]
Thus the entries of the unit quaternion for \(U\) are the rotation angles, thinly disguised, of \(U, UX_x, UY_y, UZ_z\).

### 2.3 Depicting rotations in the unit ball

Eq. (9) suggests that rotation matrices be represented as points in the unit ball, that is, in the set
\[
B = \{ v \in \mathbb{R}^3 : \|v\| \leq 1 \}.
\]
The correspondence between rotations \(U\) and points \(v\) of \(B\) is
\[
U \leftrightarrow v = \left( \frac{\xi}{\sin \frac{\xi}{2}} \right) u,
\]
where \(\xi\) and \(u\) are the rotation angle and normalized rotation axis of \(U\). Thus each point \(v\) in \(B\) determines a rotation matrix; the direction of \(v\) gives the rotation axis, and the norm of \(v\) determines the rotation angle \(\xi\) by
\[
\|v\| = \sin \frac{\xi}{2}.
\]
Stated otherwise, the rotation matrix at the point \(v\) is \(U(q)\), where \(U(q)\) is as in eq. (7) and where \(q\) is the unit quaternion \((\sqrt{1 - \|v\|^2}, \ v)\). See Fig. 1.

Thus the unit ball \(\mathbb{B}\) represents the group \(U\) of all rotations. All rotations can in principle be visualized at once, as suggested in Fig. 2. First some standard object is chosen, such as the cube in the figure; this is the ‘initial object’. Then, for each point \(v\) in the unit ball, the corresponding rotation \(U\) is applied to the initial object, and the result is displayed at \(v\). Since the rotation corresponding to the zero vector is the identity, the object seen at the origin is necessarily the initial object itself. By comparing the object at \(v\) with the object at the origin, one sees the effect of the rotation \(U\).

Points \(v\) on the unit sphere—the boundary of \(\mathbb{B}\)—have \(\|v\| = 1\) and hence have \(\xi = \pi\) (eq. 15); they represent 180° rotations. Antipodal points on the unit sphere therefore represent the same rotation matrix. The points need to be considered the same when \(B\) represents rotations.

Usually we will not distinguish carefully between the unit ball \(\mathbb{B}\) and the group \(U\) of rotations. Likewise, we will not distinguish between a point \(v\) of \(\mathbb{B}\) and its corresponding rotation \(U\) in \(U\).

### 3 PRINCIPAL AXIS ANGLE \(\xi_0\)

A ‘frame’ is a right-handed orthonormal basis for \(\mathbb{R}^3\). Initially the principal axis angle \(\xi_0\) will be defined for pairs of frames. Eventually \(\xi_0\) will be defined for pairs of moment tensors, by applying it to the frames associated with the principal axes of the moment tensors.

\[\begin{align*}
U_1U_1^{-1}(U_1e_i) &= U_2e_i, \\
U_2U_2^{-1} : U_1 \rightarrow U_2.
\end{align*}\]
is the result of applying $U$ of the unit ball, and $180^\circ$ balls. It is therefore reasonable to define the distance between frames of cubes shown, much is left to the imagination.

It is therefore reasonable to define the distance between frames $U_1$ and $U_2$ by

$$d(U_1, U_2) = \xi(U_2 U_1^{-1}),$$

where $\xi(V)$ denotes the rotation angle of $V$. Then the function $d$ is indeed a metric in the usual sense. That is,

$$d(U_1, U_2) \geq 0,$$

$$d(U_1, U_2) = 0 \text{ iff } U_1 = U_2,$$

$$d(U_1, U_2) = d(U_2, U_1) \quad \text{(symmetry)},$$

$$d(U_1, U_2) + d(U_2, U_3) \geq d(U_1, U_3) \quad \text{(triangle inequality)}. $$

Eq. (19c) follows from eq. (2a) and from the fact that trace $U = \text{trace } (U^{-1})$ if $U$ is a rotation. Eq. (19d) is proved in Appendix C. The metric $d$ is both right and left invariant:

$$d(U_1 V, U_2 V) = d(U_1, U_2),$$

$$d(V U_1, V U_2) = d(U_1, U_2).$$

the latter equation following from trace $(V U V^{-1}) = \text{trace } U$. From eq. (18),

$$d(I, U) = \xi(U).$$

Since we usually consider frames to be points of the unit ball $\mathbb{B}$, we sometimes need to distinguish between the distance $d$ and the ordinary euclidean distance in $\mathbb{B}$. In that case we will refer to $d$ as the rotation metric, or the $d$-metric, or something similar.

Huynh (2009) has much more on metrics for rotations.

3.2 Definition of $\xi_0$

Let $X_\pi, Y_\pi, Z_\pi$ be the rotations through angle $\pi$ about the $x, y, z$ coordinate axes. In particular, $X_\pi, Y_\pi, Z_\pi$ are the $180^\circ$ rotations about the coordinate axes.

For any frame $U$, the frame $UX_\pi$ can be written $(UX, U^{-1})U$. The frame $UX_\pi$ is therefore the result of rotating the frame $U$ through an angle of $180^\circ$ about its own first frame vector. Equivalently, $UX_\pi$ results from $U$ by reversing the directions of its second and third frame vectors. Similarly for $UY_\pi$ and $UZ_\pi$. See Fig. 3.

Thus, for any frame $U$, the frames $U, UX_\pi, UY_\pi, UZ_\pi$ all correspond to the same ordered triple of principal axes. Technically, in fact, the principal axis triple determined by the frame $U$ is the set $\mathcal{U} = U D_2 = \{U, UX_\pi, UY_\pi, UZ_\pi\}$, (22)

where $D_2$ is the symmetry group

$$D_2 = \{I, X_\pi, Y_\pi, Z_\pi\}. $$

To define a reasonable measure of separation $\xi_0$ between principal axis triples that are associated with frames $U_1$ and $U_2$, we therefore need to consider not just the rotation

$$U_1 \rightarrow U_2,$$

but rather all of the rotations

$$U_1 R_1 \rightarrow U_2 R_2, \quad R_1, R_2 \in D_2.$$ (25)

These rotations are the matrices $U_2 R_1 (U_1 R_1)^{-1}$ (eq. 17). The principal axis angle $\xi_0(U_1, U_2)$ is defined to be the smallest of their rotation angles.
Since \( U_1 R_2(U_1 R_1)^{-1} = U_2 R_2 R_1 U_1^{-1} \) and \( R_2 R_1 \in \mathbb{D}_2 \), there are only four rotations to consider (not 16). Their rotation angles can be found from eq. (2a); the smallest of them gives \( \xi_0 \).

This brute force calculation of \( \xi_0 \) is workable, and it requires no calculation of rotation axes, but it does not give much insight. Eq. (34) below gives a more concise expression for \( \xi_0 \) that can be used to visualize \( \xi_0 \) as a function.

### 3.3 Some properties of \( \xi_0 \)

In terms of the rotation distance \( d \) (eq. 18), the principal axis angle is

\[
\xi_0(U_1, U_2) = \min_{R_1, R_2 \in \mathbb{D}_2} d(U_1 R_1, U_2 R_2).
\]

(26)

For frames, the function \( \xi_0 \) fails to be a metric, because distinct frames can have zero separation \( \xi_0 \) between them. For example, \( \xi_0(X_\infty, I) = 0 \) but \( X_\infty \neq I \). The remaining requirements for a metric, however, are satisfied. For \( U_1, U_2, U_3 \in \mathbb{U} \),

\[
\xi_0(U_1, U_2) \geq 0,
\]

(27a)

\[
\xi_0(U_1, U_2) = \xi_0(U_2, U_1) \quad \text{(symmetry)},
\]

(27b)

\[
\xi_0(U_1, U_2) + \xi_0(U_2, U_3) \geq \xi_0(U_1, U_3) \quad \text{(triangle inequality)}.
\]

(27c)

Eqs (27a) and (27b) follow from the corresponding properties for \( d \) and from eq. (26). The triangle inequality for \( \xi_0 \) is proved in Appendix D.

From eqs (20b) and (26), the angle \( \xi_0 \) is left-invariant:

\[
\xi_0(V U_1, V U_2) = \xi_0(U_1, U_2).
\]

(28)

However, right invariance fails more often than not. As a reminder, we write

\[
\xi_0(U_1 V, U_2 V) \neq \xi_0(U_1, U_2).
\]

(29)

### 3.4 Distance between principal axis triples

From eq. (26),

\[
\xi_0(U_1 R_1, U_2 R_2) = \xi_0(U_1, U_2), \quad R_1, R_2 \in \mathbb{D}_2.
\]

(30)

This is the fundamental property that allows \( \xi_0 \) to be defined not just for frames but for principal axis triples, and hence for moment tensors (Section 8). For principal axis triples \( U_1 \) and \( U_2 \) (eq. 22), the angle \( \xi_0 \) is defined by

\[
\xi_0(U_1, U_2) = \xi_0(U_1, U_2).
\]

(31)

Whether \( U_1 \) is written as \( U_1, U_1 X_\infty, U_1 Y_\infty \), or \( U_1 Z_\infty \), the result on the right-hand side of eq. (31) is the same, by eq. (30), and similarly for \( U_2 \). But note that the analogue of eq. (31) for \( \xi \) would not work. This is the insight underlying the definition of \( \xi_0 \).

Eqs (27) and (28), which applied to frames, carry over automatically to principal axis triples, via eq. (31). Moreover,

\[
\xi_0(U_1, U_2) = 0 \Rightarrow \xi_0(U_1, U_2) = 0
\]

\[
\Rightarrow U_1 = U_2 R \quad \text{for some} \ R \in \mathbb{D}_2
\]

\[
\Rightarrow U_1 = U_2.
\]

(32)

On the set of principal axis triples, the function \( \xi_0 \) is therefore a metric. It is left-invariant. The function \( d \) is a natural measure of distance between frames. The function \( \xi_0 \) is the associated distance between principal axis triples.

© 2012 The Authors, GJI, 191, 813–831

Geophysical Journal International © 2012 RAS

### 3.5 A more concise formula for \( \xi_0 \)

Continuing from eq. (26),

\[
\xi_0(U_1, U_2) = \min_{R_1, R_2 \in \mathbb{D}_2} d(U_1 R_1, U_2 R_2)
\]

\[
= \min d(I, U_1^{-1} U_2 R)
\]

\[
= \min(\xi, \xi, \xi, \xi, \xi, \xi, \xi).
\]

(33)

where \( U = U_1^{-1} U_2 \) and where \( \xi, \xi, \xi, \xi, \xi, \xi, \xi \) are the rotation angles for \( U, U_1 X_\infty, U_1 Y_\infty, U_1 Z_\infty \). Then, from eqs (12), the principal axis angle \( \xi_0(U_1, U_2) \) is given by

\[
\cos \frac{\xi_0}{2} = \max(|w|, |x|, |y|, |z|),
\]

(34)

where \( w, x, y, z \) is a unit quaternion for \( U_1^{-1} U_2 \). Such a quaternion is, from eq. (9),

\[
w = \cos(\xi/2),
\]

(35)

\[(x, y, z) = (\sin(\xi/2)) u,
\]

where \( \xi \) and \( u \) are the rotation angle and normalized rotation axis of \( U_1^{-1} U_2 \). The quaternion can also be calculated directly, as explained in Appendix F.

Appendix E has a sample calculation of \( \xi_0 \).

### 3.6 The case \( U_1 = I \)

From eq. (28) it follows immediately that

\[
\xi_0(U_1, U_2) = \xi_0(I, U_1^{-1} U_2).
\]

(36)

Thus the general case \( \xi_0(U_1, U_2) \) can always be reduced to the case \( \xi_0(I, U) \) in which the first frame is the identity. The matrix \( U_1^{-1} U_2 \) that appears in connection with eq. (34) is then just \( U \). Stated less formally: one’s own frame of reference can always be adjusted so that the first of \( U_1 \) and \( U_2 \) appears to be the identity frame.

In most of what follows, we will be interested in \( \xi_0(I, U) \), which will then be abbreviated to \( \xi_0(U) \).

### 3.7 The frame coordinates \( \xi, \xi, \xi, \xi \)

The frame coordinates \( \xi, \xi, \xi, \xi \) of a rotation \( U \) are defined to be the rotation angles of \( U, U X_\infty, U Y_\infty, U Z_\infty \), respectively:

\[
\xi = \xi(U), \quad \xi = \xi(U X_\infty),
\]

(37)

\[
\xi = \xi(U Y_\infty), \quad \xi = \xi(U Z_\infty).
\]

They are also the respective (rotation) distances from \( U \) to \( I, X_\infty, Y_\infty, Z_\infty \). That is,

\[
\xi = \xi(U) = d(I, U),
\]

(38)

\[
\xi = \xi(U X_\infty) = d(I, U X_\infty) = d(X_\infty, U),
\]

\[
\xi = \xi(U Y_\infty) = d(I, U Y_\infty) = d(Y_\infty, U).
\]

\[
\xi = \xi(U Z_\infty) = d(I, U Z_\infty) = d(Z_\infty, U).
\]

(39)

The frame coordinates are related to the \( xyz \) cartesian coordinates in the unit ball \( B \) by eqs (12) and by

\[
w = (1 - \|v\|^2)^{1/2}, \quad v = (x, y, z).
\]

The four frame coordinates are not independent, however, since \( w^2 + x^2 + y^2 + z^2 = 1 \). Moreover, the eight symmetrically positioned points \((\pm x, \pm y, \pm z)\) all have the same frame coordinates, due to the absolute values in eqs (12), so the frame coordinates do not uniquely determine a frame unless an octant of \( B \) is specified.
which we abbreviate to $\xi = t$ for $t = 70^\circ$. The frame coordinates $\xi, \eta, \zeta$ for a frame $U$ are the rotation angles of $U, UX_x, UX_y, UX_z$, respectively. Hence the set $\xi = t$ consists of the frames $U$ for which the rotation angle of $UX_x$ is $t$. As a subset of the unit ball $B$, the set $\xi = t$ is the pair of discs $|x| = \cos(t/2)$. (Right) The set $\xi \geq t$ for $t = 140^\circ$. It consists of the frames $U$ for which the rotation angle of $UX_x$ is at least $t$. In $B$ it is the slab bounded by the horizontal discs $|z| = \cos(t/2)$.

In $B$ the level surfaces for $\xi$ are spheres, and the level surfaces for $\xi, \eta, \zeta$ are planes or, rather, discs. More precisely, the level surface $\xi = t$ is the sphere of euclidean radius $\sin(t/2)$, and the level surface $\xi = t$ is the pair of discs $|x| = \cos(t/2)$ perpendicular to the $x$-axis. The level surfaces $\xi = t$ and $\xi = t$ are analogous.

Many sets relevant to $\xi_0$ can be usefully described in terms of their frame coordinates. Consider, for example, the set

$$\{ U \in U : \xi(U) \geq t \},$$

which we abbreviate to

$$\{ \xi \geq t \} \quad \text{or} \quad \xi_0 \geq t.$$

It consists of the rotations $U$ such that the rotation angle of $UX_x$ is at least $t$. Equivalently, it is the set of rotations whose rotation distance (eq. 18) to $Z_x$ is at least $t$. It is the set $|z| \leq \cos(t/2)$, which is the horizontal slab between the two discs $|z| \leq \cos(t/2)$. Fig. 4 shows the slab for $t = 140^\circ$. Table 1 gives several other sets of rotations, both in frame coordinates and in cartesian coordinates.

From eq. (33), the parameter $\xi(t)$ is $\xi(0(I, U)$ is expressed in terms of the frame coordinates as

$$\xi_0 = \min(\xi, \eta, \zeta).$$

Hence $\xi(t)$ is the (rotation) distance from $U$ to the closest of $I, X_x, Y_y, Z_z$.

### 3.8 A partition of the unit ball

The group $U$ of rotations, or frames, is not an efficient representation of principal axis triples, since each such triple corresponds to four frames (Fig. 3). There are many ways to partition $U$, or the unit ball $B$, into four subsets in such a way that each subset gives an efficient representation of principal axis triples.

Eq. (40) suggests that $B$ be partitioned into subsets $B_w, B_x, B_y, B_z$, according to whether the principal axis angle $\xi_0 = \xi, \eta, \zeta$, or $\zeta$. In frame coordinates,

$$B_w = \{ \xi_0 = \xi \},$$

or

$$B_w = \{ \min(\xi, \eta, \zeta) = \xi \}.$$ (41)

$$B_x = \{ \xi_0 = \xi \}, B_y = \text{etc.},$$

$$B_w = \{ \min(\xi, \eta, \zeta) = \xi \}, B_z = \text{etc.}.$$ (42)

The sets $B_w, B_x, B_y, B_z$ are separated from each other by the surfaces $\xi = \xi, \eta = \xi, \xi = \xi, \xi = \xi, \zeta = \xi$, and $\zeta = \xi$. The latter three are pairs of planes. The surface $\xi = \xi_0$ is (eqs 12)

$$\cos \xi = \cos \xi_0$$

$$1 - x^2 - y^2 - z^2 = x^2$$

$$2x^2 + y^2 + z^2 = 1,$$

which is an ellipsoid of revolution. It has vertices $(\pm 1/\sqrt{2}, 0, 0)$ on the $x$-axis, and its equatorial section is the unit circle $y^2 + z^2 = 1$ in the $yz$-plane. Points $v = (x, y, z)$ that are on or inside the ellipsoid have $w \geq |x|$, that is, $\xi \leq \xi_0$. The ellipsoids $\xi = \xi_0$ and $\zeta = \xi_0$ are analogous. Points $v$ that are on or within all three ellipsoids have $\xi = \min(\xi, \xi, \zeta)$; they are the points in $B_w$. See Figs 5 and 6.

The definitions of $B_w, B_x, B_y, B_z$ can be paraphrased as follows. Each point $v \in B$ represents a frame $U$. Associated with $U$ are the four frames $U, UX_x, UX_y, UX_z$. If the smallest of the rotation angles for $U, UX_x, UX_y, UX_z$ is the rotation angle for $U$ itself, then the point $v$ is in $B_w$ (green in Fig. 5). If the smallest of the four rotation angles is the rotation angle for $UX_x$, then $v$ is in $B_x$ (red), and so forth.

The frames $U, UX_x, UX_y, UX_z$ all correspond to the same triple of principal axes (Fig. 3). The set $B_w$ consists of the frames $U$ for which $U$ is the ‘best’ (i.e. closest to the identity frame, that is, smallest rotation angle) of the frames $U, UX_x, UX_y, UX_z$. The set $B_z$ consists of the frames $U$ for which $UX_z$ is the best of $U, UX_x, UX_y, UX_z$. And so forth. Any one of the sets $B_w, B_x, B_y, B_z$ can be regarded as the set of all principal axis triples. Hence any one of them can be thought of as the space of double couple moment tensors.

### Table 1: Some sets of frames, described both in frame coordinates $\xi, \eta, \zeta$, and in cartesian coordinates $x, y, z$. In the second column, $v = (x, y, z)$ and $w = \sqrt{1 - ||v||^2}$.

<table>
<thead>
<tr>
<th>$\xi$, $\eta$, $\zeta$, $\zeta$</th>
<th>$x$, $y$, $z$</th>
<th>in $B$</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\xi = t$</td>
<td>$</td>
<td>v</td>
<td>= \sin(t/2)$</td>
</tr>
<tr>
<td>$\xi = t$</td>
<td>$</td>
<td>x</td>
<td>= \cos(t/2),</td>
</tr>
<tr>
<td>$\xi \geq t$</td>
<td>$</td>
<td>z</td>
<td>\leq \cos(t/2),</td>
</tr>
<tr>
<td>$\min(\xi, \xi, \zeta) = t$</td>
<td>$\max(</td>
<td>x</td>
<td>,</td>
</tr>
<tr>
<td>$\xi = \xi_0$</td>
<td>$2x^2 + y^2 + z^2 = 1$</td>
<td>ellipsoid, Fig. 5</td>
<td></td>
</tr>
<tr>
<td>$\xi_0 = \xi$</td>
<td>$w \geq \max(</td>
<td>x</td>
<td>,</td>
</tr>
<tr>
<td>$\xi_0 = \zeta$</td>
<td>$</td>
<td>z</td>
<td>\geq \max(</td>
</tr>
<tr>
<td>$\xi_0 = t$</td>
<td>$\max(</td>
<td>x</td>
<td>,</td>
</tr>
<tr>
<td>$\xi_0 = \xi_0$</td>
<td>$\max(</td>
<td>x</td>
<td>,</td>
</tr>
</tbody>
</table>
Finally, using eqs (38) and (40), we can characterize the frames in \( B_u \) as those frames that are at least as close to \( I \) as to \( X, Y, \) or \( Z \). The frames in \( B_v \) are those that are at least as close to \( X \), as to \( Y, Z \). And so forth.

The above ‘partition’ of \( B \) is not quite a partition in the technical set-theoretic sense, since the pairwise intersections of the four sets \( B_u, B_v, B_w, B_z \) are not empty. But they are nearly empty, being contained in the boundaries of the four sets. The partition can be turned into a true partition by carefully redefining the four sets on their boundaries, but we do not do so here. (For another partition, the redefinition is done in Appendix B of Tape & Tape (2012).)

3.9 Four symmetries of \( U \)

Each of the symmetry mappings \( U \rightarrow UX, U \rightarrow UY, U \rightarrow UZ \) permutes the subsets \( B_u, B_v, B_w, B_z \) among each other. Eqs (37) and (41) show, for example, that the mapping \( U \rightarrow UX \) swaps \( B_u \) with \( B_w \), and swaps \( B_v \) with \( B_z \). Since multiplication by any rotation preserves the natural (i.e. Haar) measure \( \mu \) on the group of rotations, then each of \( B_u, B_v, B_w, B_z \) has the same measure. And since any two of the sets overlap only on their boundaries, which have measure zero, then the probability that a random rotation \( U \) lies in any one of \( B_u, B_v, B_w, B_z \) is the same for each, namely \( \mu = 1/4 \). And since each of the symmetry mappings preserves the metric \( d \) (eqs 20a), then the sets \( B_u, B_v, B_w, B_z \) are all congruent with respect to \( d \).

4 Level Surfaces of \( \xi_0 \)

We let \( \mathbb{L}(t) \) be the level surface of \( \xi_0 \) for the value \( t \). That is,

\[
\mathbb{L}(t) = \{ \xi_0 = t \}.
\]

In words, \( \mathbb{L}(t) \) consists of all frames whose principal axis angle \( \xi_0 \) is \( t \). When considered as subsets of the unit ball \( B \), the level surfaces make up all of \( B \):

\[
B = \bigcup_t \mathbb{L}(t).
\]

Knowing all the level surfaces of \( \xi_0 \) is equivalent to knowing the function \( \xi_0 \) itself.

In principle the level surfaces of \( \xi_0 \) are simple. Since \( \xi_0 = \min(\xi, \xi_x, \xi_y, \xi_z) \) (eq. 40), then in \( B_u \), the level surfaces of \( \xi_0 \) must coincide with the level surfaces of \( \xi \), and in \( B_v \) they must coincide with the level surfaces of \( \xi_x \), and so forth. The level surfaces of \( \xi \) are spheres centred at the origin, and the level surfaces of \( \xi_x, \xi_y, \) and \( \xi_z \) are planes perpendicular to the \( x \), \( y \), and \( z \)-axes, respectively (Section 3.7). The details can nevertheless stand some elaboration.

We define subsets \( \mathbb{L}_w(t), \mathbb{L}_v(t), \mathbb{L}_u(t), \mathbb{L}_z(t) \) of \( \mathbb{L}(t) \) by

\[
\mathbb{L}_w(t) = B_w \cap \mathbb{L}(t) = \{ \xi_0 = \xi = t \},
\]

\[
\mathbb{L}_v(t) = B_v \cap \mathbb{L}(t) = \{ \xi_0 = \xi_x = t \},
\]

\[
\mathbb{L}_u(t) = B_u \cap \mathbb{L}(t) = \{ \xi_0 = \xi_y = t \},
\]

\[
\mathbb{L}_z(t) = B_z \cap \mathbb{L}(t) = \{ \xi_0 = \xi_z = t \}.
\]

In Section 3.8 we said that any one of \( B_u, B_v, B_w, B_z \) could be thought of as the set of all principal axis triples (or double couple moment tensors). Similarly, any one of \( \mathbb{L}_w(t), \mathbb{L}_v(t), \mathbb{L}_u(t), \mathbb{L}_z(t) \) can be thought of as the set of principal axis triples having principal axis angle \( \xi_0 = t \).
4.1 In two dimensions

Fig. 7(a) shows how to construct the set \( L_w(t) \) in the \( z = 0 \) section of the unit ball. Figs 7(b) and (c) do the same for \( L_s(t) \) and \( L_r(t) \).

For the \( t \)-value under consideration, namely \( t = 100^\circ \), the level set \( L(t) = L_w(t) \cup L_s(t) \cup L_r(t) \) consists of four little isosceles right ‘triangles’ having curved hypotenuses.

4.2 In three dimensions

The nature of \( L(t) \), whether in two or three dimensions, is determined by the relative disposition of a certain sphere \( S(t) \) and cube \( C(t) \). They, as well as the related solids \( \tilde{S}(t) \) and \( \tilde{C}(t) \), are defined in cartesian coordinates by

\[
\begin{align*}
S(t) &= \{ \|v\| = \sin \frac{t}{2} \}, \\
\tilde{S}(t) &= \{ \|v\| \geq \sin \frac{t}{2} \}, \\
C(t) &= \{ \max(|x|, |y|, |z|) = \cos \frac{t}{2} \}, \\
\tilde{C}(t) &= \{ \max(|x|, |y|, |z|) \leq \cos \frac{t}{2} \}.
\end{align*}
\]

Thus the sphere \( S(t) \) has radius \( \sin(t/2) \), and the cube \( C(t) \) has vertices \(( \pm s, \pm s, \pm s) \), where \( s = \cos(t/2) \). The set \( \tilde{S}(t) \) consists of the points that are on or outside \( S(t) \), and \( \tilde{C}(t) \) consists of the points that are on or inside \( C(t) \).

In frame coordinates, from eqs (12),

\[
\begin{align*}
S(t) &= \{ \xi = t \}, \\
\tilde{S}(t) \cap B &= \{ \xi \geq t \}, \\
C(t) \cap B &= \{ \min(\xi_x, \xi_y, \xi_z) = t \}, \\
\tilde{C}(t) \cap B &= \{ \min(\xi_x, \xi_y, \xi_z) \geq t \}.
\end{align*}
\]

In connection with \( C(t) \), we define points \( D, E, F \) by

\[
\begin{align*}
D(t) &= (s, 0, 0) \quad \text{(centre of a face of } C(t)), \\
E(t) &= (s, s, 0) \quad \text{(midpoint of an edge of } C(t)), \\
F(t) &= (s, s, s) \quad \text{(a vertex of } C(t)),
\end{align*}
\]

\( s = \cos \frac{t}{2} \).

From eqs (46) and (48),

\[
\begin{align*}
L_w(t) &= \{ \xi_0 = \xi = t \} \\
&= \{ \xi = t \} \cap \{ \min(\xi_x, \xi_y, \xi_z) \geq t \} \\
&= S(t) \cap \tilde{C}(t) \cap B \\
&= S(t) \cap \tilde{C}(t).
\end{align*}
\]

Fig. 8 shows the evolution of \( L_w(t) \) with increasing \( t \). According to eq. (50), the set \( L_w(t) \) is the portion of the sphere \( S(t) \) that is on or inside the cube \( C(t) \). As \( t \) increases, \( S(t) \) grows and \( C(t) \) shrinks. The set \( L_w(t) \) changes qualitatively when any of the points \( D, E, F \) is on \( S(t) \). From eqs (49),

\[
\begin{align*}
\|D(t)\| &= \sin(t/2) \text{ iff } t = t_1, \\
\|E(t)\| &= \sin(t/2) \text{ iff } t = t_2, \\
\|F(t)\| &= \sin(t/2) \text{ iff } t = t_3,
\end{align*}
\]

where

\[
\begin{align*}
t_1 &= \pi/2, \\
t_2 &= \cos^{-1}(-1/3) = 109.5^\circ, \\
t_3 &= 2\pi/3.
\end{align*}
\]
The set \( L_w(t) \) is the sphere \( S(t) \) minus six open spherical caps. The caps are empty when \( t \leq t_1 \), and they overlap when \( t > t_2 \). They completely cover the sphere when \( t > t_1 \), in which case \( L_w(t) \) is empty.

In particular, regardless of the direction of the rotation axis,

\[
\xi \leq \pi/2 \Rightarrow \xi_0 = \xi,
\]

\[
\xi > 2\pi/3 \Rightarrow \xi_0 < \xi.
\]  

(53)

(These also follow directly from eq. 34.) Also see Figs 13(a) and (d).

To describe \( L_w(t) \cup L_y(t) \cup L_z(t) \), we first find from eq. (46) that

\[
L_w(t) = \{ \xi_0 = \xi_1 = \xi \} = [\min(\xi_1, \xi_1, \xi_1) = t] \cap [\min(\xi_1, \xi_1, \xi_1) = t] = [\{ \xi_1 = t \} \cap (\xi \geq t)] \cap [\{ \min(\xi_1, \xi_1, \xi_1) = t \}].
\]  

(54)

Then, from eqs (48) and (54),

\[
L_y(t) \cup L_z(t) \cup L_w(t) = [\{ \min(\xi_1, \xi_1, \xi_1) = t \} \cap (\xi \geq t)] = C(t) \cap \mathbb{B} \cap \bar{S}(t).
\]  

(55)

Thus \( L_w(t) \cup L_y(t) \cup L_z(t) \) is the portion of \( C(t) \cap \mathbb{B} \) that is on or outside \( S(t) \); see Fig. 9.

Eqs (50) and (55) together describe the level surface \( L(t) \):

\[
L(t) = (S(t) \cap C(t)) \cup (C(t) \cap \mathbb{B} \cap \bar{S}(t)).
\]  

(56)

The evolution of the level surface \( L(t) \) with increasing \( t \) is illustrated in Fig. 10. Like \( L_w(t) \), the set \( L(t) \) changes qualitatively when \( t = t_1, t_2, \) or \( t_3 \). Just as the symmetry mappings \( U \rightarrow UX \), \( U \rightarrow UY \), \( U \rightarrow UZ \) permute the sets \( B_x, B_y, B_z, \) among each other (Section 3.9), so too they permute \( L_w(t) \), \( L_y(t) \), \( L_z(t) \), \( L(t) \). In spite of appearances, those latter four sets are therefore all congruent, but congruent with respect to the rotation distance \( d \) rather than with respect to the ordinary euclidean distance in \( \mathbb{B} \). In Fig. 10(a), for example, since \( t \leq \pi/2 \), each of \( L_w(t) \), \( L_y(t) \), \( L_z(t) \), \( L(t) \) is a \( d \)-sphere with radius equal to \( t \); the sphere \( L_w(t) \) is centred at \( I \), the sphere \( L_y(t) \) is centred at \( X_y \), etc. In \( \mathbb{B} \) the sphere \( L_z(t) \) appears to be two disconnected red discs, but in fact the discs are connected to each other along their boundaries, since antipodal points on the unit sphere are the same.
4.3 The transition values $t_1, t_2, t_3$

There are many equivalent ways to describe the ‘times’ $t_1, t_2, t_3$ when the level set $L_t$ changes qualitatively. As in eqs (51), $t_1, t_2, t_3$ are the times when the respective centres, edge midpoints, and vertices of the cube $C(t)$ are on the sphere $S(t)$. They are also the times when the points $A, B, C$ of Fig. 5 are on $S(t)$. The time $t = t_1$ is also when any two of the $d$-spheres $\xi = t, \xi_x = t, \xi_y = t, \xi_z = t$ are tangent to each other. The time $t = t_2$ is when any one of the $d$-spheres meets the intersection of two others tangentially. Note, however, that tangencies are not depicted correctly on the boundary of $B$.

5 PROBABILITY OF $\xi_0$

In Appendix A we show that the probability density for $\xi_0$ is

$$p_{\xi_0}(t) = \frac{2}{\pi^2} \left( \text{area of } L_{\xi_0}(t) \right).$$

That is, the probability that a random rotation have $\xi_0$ between 0 and $t$ is

$$P(0 \leq \xi_0 \leq t) = \frac{2}{\pi^2} \int_0^t \left( \text{area of } L_{\xi_0}(\xi) \right) d\xi.$$

The behaviour of $p_{\xi_0}(t)$ can therefore be seen at a glance in Fig. 8.
Because small $\xi_0$ values are so unlikely for random orientations, the leftward skewed histograms are even more telling than they might first appear.

6 DEPICTING ROTATIONS IN RODRIGUES SPACE

In this paper we have depicted rotations as points in the unit ball $B$. Some authors replace $B$ with all of $\mathbb{R}^3$, which they then refer to as Rodrigues space (e.g. Frank 1988). Underlying each approach is a parameterization of rotations. The parameter points $v = (x, y, z)$ and $v' = (x', y', z')$ for a rotation $U$ are

$$v = \left(\sin\frac{\xi}{2} \right) u \quad \text{(for unit ball)},$$

$$v' = \left(\tan\frac{\xi}{2} \right) u \quad \text{(for Rodrigues)},$$

where $\xi$ and $u$ are the rotation angle and normalized rotation axis of $U$. The points $v$ and $v'$ are then related by

$$v' = \left(\sec\frac{\xi}{2} \right) v = \frac{v}{(1 - \|v\|^2)^{1/2}},$$

$$v' = \left(\cos\frac{\xi}{2} \right) v' = \frac{v'}{(1 + \|v\|^2)^{1/2}}.$$  

From eq. (61a) the Rodrigues coordinates $x', y', z'$ are related to our usual $x, y, z$ coordinates by

$$x' = x/w, \quad y' = y/w, \quad z' = z/w.$$  

The Rodrigues setting has some advantages. In Rodrigues space, with distances between rotations measured by the rotation metric $d$ (eq. 18), the locus of rotations equidistant from two given rotations turns out to consist of two planes. In Rodrigues space the set $B_\xi$, whose boundary surfaces are formed by $\xi = \xi_0, \xi = \xi_1$, and $\xi = \xi_2$, must therefore be a solid polyhedron, with truly planar faces. In fact it is the solid unit cube; the boundary surface $\xi = \xi_1$, for example, is $|x'| = 1$, since

$$|x'| = \frac{|x|}{w} = \cos(\xi_1/2).$$

But Rodrigues space also has its drawbacks. It does not contain 180° rotations, and such rotations therefore have to be treated as limiting cases. And although some sets, like $B_\xi$, look better in Rodrigues space than in the unit ball, other sets look worse. The level surfaces $\xi_0 = \xi_1, \xi_1 = \xi_0$, and $\xi_0 = \xi_2, \xi_1 = \xi_2$, which were planar in the unit ball (e.g. Fig. 4), become hyperboloids in Rodrigues space. Thus the sets $L_0(t), L_1(t), L_2(t)$, which were parts of cubic faces in Fig. 10, look extinct in Rodrigues space (Fig. 12). And eq. (57) is an example of a result that is better expressed in the context of the unit ball than in Rodrigues space. In any event, one can easily go back and forth between the (interior of the) unit ball and Rodrigues space, using eqs (61).

Frank (1988) has a readable exposition of Rodrigues space, and details can be found in Morawiec (2004). In Frank's terminology, our set $B_\xi$ is the fundamental zone for the group $\mathbb{D}_2$. Our angle $\xi_0$ is the misorientation angle for $\mathbb{D}_2$.

7 THE ANGLE $\xi_0$ FOR FIXED $\xi$

For purposes of comparing $\xi_0$ with other measures of orientation separation, it is convenient to plot contours of $\xi_0$ on the spheres $S(t)$. For each $t$, $0 \leq t \leq \pi$, the contour plot on $S(t)$ displays $\xi_0$ as a function of rotation axis, with rotation angle $\xi = t$ fixed. Due to the
the differences in appearance ought to be due to differences in orientation. Thus \( \xi_0 \) is not capturing our informal sense of difference in moment tensor orientations. Willemann (1993) makes the same point when, referring to \( \xi_0 \), he says

\[
\ldots \text{the rotation required for alignment is unreasonable in the sense that a rotation of only 90° is required to align opposite double couples.}
\]

### 8.1 An alternative to \( \xi_0 \)

In trying to define a reasonable measure of difference in moment tensor orientations, we are motivated by the fact that the eigenvalues of a moment tensor give its size and source type (its pattern, in beachball language), and that the eigenvectors give the orientation. For measuring differences in orientation, the eigenvalues should therefore be ignored (but see below). In the definition of \( \xi_0 \) for moment tensors (eq. 64), the only role of the eigenvalues is to order the principal axes, so that the principal axes of the one moment tensor can be associated with the principal axes of the other. Once the association is made, it does not matter which of the three axes is the \( T \), \( B \), or \( P \) axis.

That may be going too far, as was suggested in Fig. 14. The parameter \( \omega \), defined next, gives an alternative measure of orientation differences that does capture some of the intuition for the differences seen in Fig. 14. For orientations (i.e. frames) \( U_1 \) and \( U_2 \), the function \( \omega \) is defined by

\[
\omega(U_1, U_2) = \angle([A_0]_{U_1}, [A_0]_{U_2}), \quad A_0 = (1, 0, -1),
\]

(66)

where \( \angle \) is the angle in the space of \( 3 \times 3 \) matrices. That is, \( 3 \times 3 \) matrices are treated as elements of \( \mathbb{R}^9 \), and the angle between them is computed using the standard inner product in \( \mathbb{R}^9 \). Thus, for \( 3 \times 3 \) matrices \( M \) and \( N \),

\[
\cos \angle(M, N) = \frac{M \cdot N}{\|M\| \|N\|} = \frac{\sum m_{ij} n_{ij}}{(\sum m^2_{ij})^{1/2} (\sum n^2_{ij})^{1/2}}.
\]

(67)

As was done earlier for \( \xi_0 \), we write \( \omega(U) \) for \( \omega(I, U) \).

For the 90° rotations \( U_1 \) and \( U_2 \) in Fig. 14, we find from eq. (66) that \( \omega(U_1) = \pi/3 \) and \( \omega(U_2) = \pi \), which is more in keeping with our informal sense of the orientation differences than the values \( \xi_0(U_1) = \xi_0(U_2) = \pi/2 \). More generally, Fig. 15 shows how \( \xi_0 \) and \( \omega \) compare for 90° rotations.

Since \( [A]_{U_1} = [A]_{U_2} \) for \( R \in \mathbb{D}_2 \), then

\[
\omega(U_1, R_1, U_2, R_2) = \omega(U_1, U_2), \quad R_1, R_2 \in \mathbb{D}_2
\]

(68)

This means that \( \omega \) can be defined for moment tensors, not just for frames. The definition is analogous to eq. (64):

\[
\omega([A]_{U_1}, [A]_{U_2}) = \omega(U_1, U_2), \quad A_1, A_2 \in \mathbb{W}_0,
\]

(69)

or, equivalently,

\[
\omega([A]_{U_1}, [A]_{U_2}) = \angle([A]_{U_1}, [A]_{U_2}), \quad A_1, A_2 \in \mathbb{W}_0.
\]

(70)

For any \( A \in \mathbb{W}_0 \) the closest double couple to the moment tensor \( [A]_{U_1} \) is \( [A_{DC}]_{U_1} \), where \( A_{DC} \) is the vector projection of \( A \) onto \( A_0 \). That is,

\[
A_{DC} = \frac{A \cdot A_0}{\|A\|^2} A_0.
\]

(71)

Since \( A_{DC} \) and \( A_0 \) are scalar multiples of each other, then eq. (70) can be rewritten as

\[
\omega([A]_{U_1}, [A]_{U_2}) = \angle([A_{DC}]_{U_1}, [A_{DC}]_{U_2}), \quad A_1, A_2 \in \mathbb{W}_0.
\]

(72)

In words, \( \omega(A_1, A_2) \) is the angle, in matrix space, between the closest double couples to \( A_1 \) and \( A_2 \). The intuition behind \( \omega \) is that,
Figure 13. Contour plots of the rotation angle $\xi$ and the principal axis angle $\xi_0$ on one octant of the sphere $S(t)$. On $B_x \cap S(t)$ the angle $\xi$ coincides with $t$, and the contour plot for $\xi$ is a solid colour in each case (left column). On the subset $L_{\omega}(t)$ the angle $\xi_0$ coincides with $\xi$, and so the contour plot for $\xi_0$ (middle column) on $L_{\omega}(t)$ is a solid colour matching that for $\xi$. On $B_y \cap S(t)$ the angle $\xi_0$ coincides with $\xi_y$, whose contour surfaces in $B$ are planes perpendicular to the $x$-axis and whose contour curves on $S(t)$ are therefore circles with poles on the $x$-axis. Similarly on $B_y \cap S(t)$ and $B_z \cap S(t)$. The set $L_{\omega}(t)$ varies with $t$ as shown in Fig. 8 and as described in eq. (50). The right-hand diagrams show the rotations $X_t$ and $Y_t$ at their correct locations in $B$.

First, by using angles (in matrix space) rather than distances, we ignore all information about size. Second, by replacing $\Lambda_1$ with $\Lambda_0$ (or $\Lambda_{DC}$), we ignore the information about source type. The remaining information should be orientation. However, the distinction between the $T$, $B$, and $P$ axes is not lost, due to the presence of $\Lambda_0$ in eq. (70), and the three axes get treated differently in calculating $\omega$. In this regard, $\omega$ differs from $\xi_0$.

It is clear from eqs (64) and (69) that $\xi_0([\Lambda_1]_{01}, [\Lambda_2]_{02})$ and $\omega([\Lambda_1]_{01}, [\Lambda_2]_{02})$ are both independent of the eigenvalue triples $\Lambda_i \in \mathcal{W}_0$. That fact can be disconcerting at first, since changing
The principal axis angle $\xi_0$ measures differences in moment tensor orientations. Conceptually, however, $\xi_0$ can stand alone as a measure of differences between principal axis triples, by which we mean ordered triples of mutually perpendicular lines, not necessarily associated with any moment tensor. To elucidate the logic of $\xi_0$, we have left moment tensors out of the discussion until the very end.

Informally, the angle $\xi$ between two principal axis triples is the smallest angle required to rotate one triple into the other. Formally, $\xi(U_1, U_2)$ is defined in Section 3.2 as a function of frames $U_1$ and $U_2$. The general case $\xi(U_1, U_2)$ is easily reduced to the case $\xi(U, \bar{U})$ where the first frame is the identity. Eq. (34) gives a concise formula for $\xi_0$.

9 OTHER PARTITIONS OF FRAMES

The partition $\mathcal{B}_x, \mathcal{B}_y, \mathcal{B}_z$ of the unit ball $\mathcal{B}$ in Fig. 5 was constructed so that any one of the four subsets could represent the set of principal axis triples or, equivalently, the set of double couple moment tensors. Many other partitions can do the same.

Fig. 16 shows a partition of $\mathcal{B}$ into unnamed green, red, blue, and yellow subsets. Eq. (11b) shows that the mapping $U \mapsto UX_y$ swaps the green with the red subset (and blue with yellow). Similarly, $U \mapsto UY_x$ swaps green with blue, and $U \mapsto UZ_x$ swaps green with yellow. Those properties by themselves are enough to ensure that the green subset can serve to represent all principal axis triples, since, for any frame $U$, one of $UX_y, UY_x, UZ_x$ will be in the green subset. So the two partitions have much in common. Many other such partitions of rotations are also possible; Tape & Tape (2012) have one that is based on slip and dip angles. What distinguishes the partition $\mathcal{B}_x, \mathcal{B}_y, \mathcal{B}_z$ is the definition condition $\xi_0 = \xi$ for $\mathcal{B}_w$.
The behaviour of $\xi_0$ is not at all transparent, especially if one relies on calculations alone. In Section 2.3 we describe how rotation matrices, or frames, can be depicted as points in the unit ball $\mathcal{B}$. We do not always distinguish between a point $v$ of $\mathcal{B}$ and its corresponding rotation matrix $U$, and we do not always distinguish between $\mathcal{B}$ and the group $\mathcal{U}$ of rotations. The angle $\xi_0 = \xi_0(I, U)$ then becomes a function on $\mathcal{B}$ and can be visualized. Fig. 13 shows how $\xi_0$ varies for frames $U$ having fixed rotation angle. Fig. 10 exhibits the level surfaces of $\xi_0$ in $\mathcal{B}$. It shows, for example, that it is easy to construct $U$ with $\xi_0 = 40^\circ$ but difficult to construct $U$ with $\xi_0 = 112^\circ$; in the latter case, both the rotation axis and the rotation angle of $U$ must be chosen judiciously.

Each triple of principal axes corresponds to four frames (Fig. 3). The group $\mathcal{U}$ of frames is therefore too big by a factor of 4 for representing principal axis triples efficiently, or for representing double couple moment tensors. There are many ways to partition $\mathcal{U}$, e.g., exhibiting principal axis triples efficiently, or for representing double couple moment tensors such that any one of the subsets represents all principal axes in $\mathcal{B}$. One such partition, stated for the unit ball $\mathcal{B}$, is $\mathcal{U} = \{I, U_1, U_2, U_3\}$, where $I$ is the identity matrix and $U_1, U_2, U_3$ are the rotation matrices for the three principal axes. Another such partition is shown in Fig. 16. What distinguishes the first partition is its close connection with the function $\xi_0$.

The set $\mathcal{L}_0(t)$ consists of the rotations whose principal axis angle $\xi_0$ and rotation angle $\xi$ both equal $t$. When those rotations are considered as points in the unit ball, the set $\mathcal{L}_0(t)$ is a subset of the sphere of radius $\sin(t/2)$ (Fig. 8). In Appendix A we show that for random orientations the probability density $p_{\xi_0}(t)$ for $\xi_0$ is proportional to the area of $\mathcal{L}_0(t)$. We then calculate the area of $\mathcal{L}_0(t)$, which gives an analytical expression for $p_{\xi_0}$.

The curve $p_{\xi_0}$ shows what to expect for random orientations (Fig. 11). It serves as a reference curve against which to compare distributions of empirical $\xi_0$ values. It shows that, for inferring similarities of orientations, small values of $\xi_0$ are more significant than they might appear.

The mathematical setting for discussing $\xi_0$ is the group $\mathcal{U}$ of frames. The distance $d$ between two frames is defined to be the rotation angle of the rotation that takes the one frame to the other. Then $d$ is indeed a metric for frames, and $\xi_0$ is the metric for principal axes triples that is the counterpart of $d$.

Fig. 14 shows that the angle $\xi_0$ sometimes conflicts with our intuitive feeling as to what constitutes a reasonable measure of separation for moment tensor orientations. We introduce an alternate measure $\omega$ and compare it with $\xi_0$. The angle $\omega$ between moment tensors $M_1$ and $M_2$ is the angle, in matrix space, between the closest double couples to $M_1$ and $M_2$.

ACKNOWLEDGMENTS

We thank two reviewers for helpful suggestions. One reviewer was Yan Kagan, and the other was anonymous. Both Hans Grimmer and Adam Morawiec kindly responded to our e-mail inquiries regarding the historical evolution of concepts related to the misorientation angle.

REFERENCES


Grimmer, H., 1979a. The distribution of misorientation angles if all relative orientations of neighbouring grains are equally probable, Scr. Metall., 13, 161–164.


APPENDIX A: PROBABILITY DENSITY \( p_{\xi}(t) \) IS PROPORTIONAL TO THE AREA OF \( \mathbb{L}_\omega(t) \)

In this appendix we verify eq. (57) of Section 5.

Since the sets \( \mathbb{L}_\omega(t), \mathbb{L}_x(t), \mathbb{L}_y(t), \mathbb{L}_z(t) \) are permuted among each other by the symmetry mappings \( U \to UX_x, U \to UX_y, U \to UX_z \), so too are the four sets \( \mathbb{L}_\omega(t), \mathbb{L}_x(t), \mathbb{L}_y(t), \mathbb{L}_z(t) \) defined by
\[
\mathbb{L}_\omega(t) = \{ \xi_0 = \xi \leq t \} = \bigcup_{t/2 \leq \xi \leq t} \mathbb{L}_\omega(t'),
\]
\[
\mathbb{L}_x(t) = \{ \xi_0 = \xi, \xi \leq t \} \cap \mathbb{L}_\omega(t'), \quad \text{etc.}
\]
(A1)

The four sets have the same (Haar) measure \( \mu \), since the symmetry mappings preserve the measure. The pairwise intersections of the sets have measure zero, and so the probability that a random rotation has \( \xi_0 \leq t \) is
\[
P(\xi_0 \leq t) = 4\mu(\mathbb{L}_\omega(t)).
\]
(A2)

where \( \mu \) is assumed to be normalized so that \( \mu(\mathbb{B}) = 1 \).

Since \( \mathbb{L}_\omega(t') \subset \mathbb{S}(t') \) (e.g. eq. 50), then
\[
\mathbb{L}_\omega(t') \cap \mathbb{S}(t') = \{ \mathbb{L}_\omega(t'), t' \leq t, \theta, t' > t \}.
\]
(A3)

The probability that a random rotation has its rotation angle between \( \xi \) and \( \xi + d\xi \) and has the longitude and colatitude of its rotation axis between \( \phi \) and \( \phi + d\phi \) and between \( \theta \) and \( \theta + d\theta \) is (e.g. Miles 1965)
\[
\frac{1}{2\pi^2} \sin \frac{\xi}{2} \sin \theta \sin \phi \sin \theta \sin \phi \ d\xi.
\]
(A4)

If \( F_1 \) is the usual parameterization of the sphere \( \mathbb{S}(\xi) \) of radius \( \sin(\xi/2) \), that is, if
\[
F_1(\theta, \phi) = \sin \frac{\xi}{2} \left( \cos \phi \sin \theta, \sin \phi \sin \theta, \cos \theta \right),
\]
\( 0 \leq \theta \leq \pi, \ 0 \leq \phi \leq 2\pi \),
(A5)

then the parameterization of the unit ball in terms of \( (\xi, \theta, \phi) \) is
\[
(\xi, \theta, \phi) \rightarrow F_1(\theta, \phi).
\]
(A6)

From eq. (A4), the measure of a set \( \mathbb{A} \subset \mathbb{B} \) is
\[
\mu(\mathbb{A}) = \frac{1}{2\pi^2} \int_{\mathbb{S}(\xi)} \sin \frac{\xi}{2} \sin \theta \sin \phi \sin \theta \sin \phi \ d\xi
\]
\[
= \frac{1}{2\pi^2} \int_{0}^{\pi/2} \int_{0}^{2\pi} \sin \frac{\xi}{2} \sin \theta \sin \phi \sin \theta \sin \phi \ d\phi \ d\theta
\]
\[
= \frac{1}{2\pi^2} \int_{0}^{\pi} \text{area of } (\mathbb{A} \cap \mathbb{S}(\xi)) d\xi.
\]
(A7)

If we take \( \mathbb{A} = \mathbb{L}_\omega(t) \) in eq. (A7), then eqs (A2) and (A3) give
\[
P(\xi_0 \leq t) = 2 \mu(\mathbb{L}_\omega(t)) = \frac{1}{\pi^2} \int_{0}^{t} \text{area of } \mathbb{L}_\omega(\xi) d\xi.
\]
(A8)

The derivative of eq. (A8) gives the probability density for \( \xi_0 \):
\[
p_{\xi_0}(t) = \frac{2}{\pi^2} \left( \text{area of } \mathbb{L}_\omega(t) \right).
\]
(A9)

which is eq. (57).

To get a sense of how the invariant measure \( \mu \) differs from the ordinary volume in \( \mathbb{B} \), we make the substitution \( \rho = \sin(\xi/2) \) in eq. (A7) and find
\[
\mu(\mathbb{A}) = \frac{1}{\pi^2} \int_{0}^{1} \left( \frac{\text{area of } (\mathbb{A} \cap \mathbb{S}(\rho))}{\sqrt{1 - \rho^2}} \right) d\rho,
\]
(A10)

where \( \mathbb{S}(\rho) \) is the sphere of euclidean radius \( \rho \). Without the factor \( (1 - \rho^2)^{-1/2} \), the integral in eq. (A10) would give the ordinary volume of \( \mathbb{A} \).

This is the natural place to calculate the distribution of the ‘mis-orientation axes’—the rotation axes of the orientations in \( \mathbb{B}_\omega \). We instead refer to Morawiec (1996).

APPENDIX B: AREA OF \( \mathbb{L}_\omega(t) \) WHEN \( t \geq t_2 \)

In this appendix we verify eq. (59) of Section 5.

The set \( \mathbb{H}' \) shown in Fig. B1 is \( 1/48^{th} \) of the set \( \mathbb{L}_\omega(t) \). It is the radial projection of the red planar set \( H \) to the sphere \( \mathbb{S}(t) \) of radius \( \sin(t/2) \). The set \( H \) is on the cube face \( x = \cos(t/2) \), and the radial projection \( f: H \to \mathbb{H}' \)
\[
f(y, z) = \left( \sin \frac{t}{2}, \frac{\cos(t/2)}{\sqrt{\cos(t/2)}}, \frac{\cos(t/2)}{\sqrt{\cos(t/2)}} \right).
\]
(B1)

The integral can be evaluated using polar coordinates, that is,
\[
y = r \cos \theta, \ z = r \sin \theta.
\]
As in Fig. B1(c), the set \( H \) is then
\[
H : r_0 \leq r \leq r(\theta), \ t_0 \leq \theta \leq \frac{\pi}{4},
\]
(B3)

where
\[
r_0 = \sqrt{-\cos t},
\]
\[
\frac{r(\theta)}{\cos t} = \frac{\sec \theta}{2},
\]
(B4)

\[
r_0 \cos \theta = \frac{t}{2}.
\]
(B5)

Eq. (B2) becomes
\[
\text{Area of } \mathbb{H}' = \sqrt{2} \int_{t_0}^{\pi/4} \int_{0}^{r(\theta)} \frac{r \sin(t/2)}{(1 + 2r^2 + \cos(t))^3/2} dr \ d\theta
\]
\[
= \frac{\sin t}{2} \left( \theta - \tan \frac{t}{2} \sin^{-1} \left( \frac{\sin \theta}{\sqrt{2}} \right) \right)^{\theta=\pi/4}_{\theta=t_0}.
\]
(B5)

Each of the eight curved triangles in \( \mathbb{L}_\omega(t) \) consists of six regions congruent to \( \mathbb{H}' \), so the area of \( \mathbb{L}_\omega(t) \) (Fig. 8e) is 48 times the area in eq. (B5).

APPENDIX C: TRIANGLE INEQUALITY FOR \( d \)

In this appendix we verify eq. (19d) of Section 3.1, the triangle inequality for \( d \).
For quaternions $q_1$ and $q_2$, we let $\theta(q_1, q_2)$ be the angle between them in $\mathbb{R}^4$. Thus if both $q_1$ and $q_2$ are unit quaternions, then $\cos(\theta(q_1, q_2)) = q_1 \cdot q_2$. (The notation $\cdot$ is the standard inner product in $\mathbb{R}^4$, not quaternion multiplication.)

If $q_1$ and $q_2$ are unit quaternions for rotations $U_1$ and $U_2$, and if $q_1 q_2^{-1} = (w, x, y, z)$, then

$$\cos \frac{d(U_1, U_2)}{2} = |w|$$

$$= |q_1 q_2^{-1} \cdot (1, 0, 0, 0)|$$

$$= |q_1 q_2|$$

$$= |\cos(\theta(q_2, q_1))|,$$

(C1)

where the third equation uses the invariance of the inner product under multiplication by a unit quaternion. Hence

$$d(U_1, U_2) = 2 \theta(q_1, q_2),$$

(C2)

where $q_1 \cdot q_2 \geq 0$—as can always be arranged, by changing a sign of $q_1$ or $q_2$.

The triangle inequality is known to hold for $\theta$ (Berger et al. 1984). Therefore, with $q_1, q_2, q_3$ being quaternions for $U_1, U_2, U_3$, and with the signs of $q_1$ and $q_3$ chosen so that $q_1 \cdot q_2 \geq 0$ and $q_2 \cdot q_3 \geq 0$,

$$d(U_1, U_3) + d(U_2, U_3) = 2 \theta(q_1, q_2) + 2 \theta(q_2, q_3)$$

$$\geq 2 \theta(q_2, q_3)$$

$$\geq d(U_1, U_3).$$

(C3)

The last line is an inequality rather than an equality, since $q_1 \cdot q_3$ might be negative.

**APPENDIX D: TRIANGLE INEQUALITY FOR $\xi_0$**

In this appendix we verify eq. (27c) of Section 3.3, the triangle inequality for $\xi_0$.

For frames $U_1, U_2, U_3$, we choose frames $V_1, V_3 \in \mathbb{B}_w$ so that $T_1 = U_2^{-1} U_1$ and $T_3 = U_2^{-1} U_1$. Then

$$\xi_0(U_1, U_2) + \xi_0(U_2, U_3) = \xi_0(U_2^{-1} U_1, I) + \xi_0(I, U_2^{-1} U_1)$$

$$= \xi_0(V_1, I) + \xi_0(I, V_3)$$

$$= d(V_1, I) + d(I, V_3)$$

$$\geq d(V_1, V_3)$$

$$\geq \xi_0(V_1, V_3)$$

$$= \xi_0(U_2^{-1} U_1, U_2^{-1} U_3)$$

$$= \xi_0(U_1, U_3).$$

(D1)

The corresponding fact for principal axis triples then follows from eq. (31):

$$\xi_0(\overline{U}_1, \overline{U}_2) + \xi_0(\overline{U}_2, \overline{U}_3) = \xi_0(U_1, U_2) + \xi_0(U_2, U_3)$$

$$\geq \xi_0(U_1, U_3)$$

$$= \xi_0(\overline{U}_1, \overline{U}_3).$$

(D2)

**APPENDIX E: SAMPLE CALCULATION OF $\xi_0$**

In this appendix we calculate the separation $\xi_0$ between two earthquakes in New Guinea. These same events were also used as an illustration by Kagan (1991). The first was on 1976 January 6. Its moment tensor, from the GCMT catalog (Dziewonski et al. 1981) but with a scale factor omitted, is

$$M_1 = \begin{pmatrix}
-0.412 & 0.398 & -1.239 \\
0.398 & 0.084 & 1.058 \\
-1.239 & 1.058 & 0.328
\end{pmatrix}.$$

We calculate the eigenvectors of $M_1$, then normalize them and order them according to decreasing size of the eigenvalues. If necessary, the direction of one vector should be reversed to make the triple of vectors right handed. The three vectors become the columns of the matrix

$$U_1 = \begin{pmatrix}
0.404 & 0.642 & 0.651 \\
-0.456 & 0.758 & -0.465 \\
-0.793 & -0.109 & 0.600
\end{pmatrix}.$$
which is then one of the four eigenframes associated with \( M \). We have displayed numbers to three decimal places, but we retain all places to full machine precision during calculations.

The second event was on 1980 September 26. Its eigenframe \( U_2 \) is found as was done for \( U_1 \). Then

\[
U_1^{-1} U_2 = U_1^T U_2 = \begin{pmatrix}
0.187 & 0.879 & -0.439 \\
-0.651 & -0.224 & -0.725 \\
-0.735 & 0.421 & 0.531
\end{pmatrix}.
\]

A quaternion for \( U_1^{-1} U_2 \), from eqs (F1a) of Appendix F, is

\[
(0.611, 0.469, 0.121, -0.626).
\]

Of the four entries, the largest in absolute value is \( z = -0.626 \). From eq. (34), the principal axis angle is therefore

\[
\phi_0 = 2 \cos^{-1} |z| = 102.5^\circ.
\]

The two earthquakes were spatially close. Had they not been close, the principal axis angle would not have been meaningful. Since a sphere is not parallelizable, there is no reasonable way to decide when frames at two widely separated points on Earth should be considered the ‘same’. (One can of course regard Earth as a subset of \( \mathbb{R}^3 \) and define the frames to be the same if one frame is the translate of the other in \( \mathbb{R}^3 \), but we cannot envision any seismological application where this point of view would be useful.)

Other approaches

A slight variation would be to calculate with quaternions rather than matrices. Eqs (F1) can be used to find unit quaternions \( q_1 \) and \( q_2 \) for \( U_1 \) and \( U_2 \). Then the quaternion for \( U_1^{-1} U_2 \) is \( q_1 q_2 \), which again gives eq. (E1).

Other approaches are also possible. For example, instead of starting with the moment tensors, one can compute unit vectors \( t \) and \( p \) in the \( T \) and \( P \) directions for each event, using the catalog values of plunge and azimuth for those directions. Ideally, the frame for an event is the matrix \( U \) with column vectors \( t, b = p \times t \), and \( p \). The vectors \( t \) and \( p \), however, are apt to be not quite orthogonal—due to rounding, for example—and so \( U \), though right-handed, will not be exactly orthogonal. Calculation of the quaternion for \( U \) becomes less straightforward, as explained in Appendix F.

There are various ways to orthogonalize \( U \), that is, to replace it with a nearby matrix \( U' \) that is orthogonal. A naive but adequate method is to use the matrix \( U' \) whose columns are

\[
t' = t, \quad b' = b / |b|, \quad p' = t \times b / |t \times b|.
\]

A slightly better result—that is, closer to \( U \)—is obtained by using the appropriate one of \( q_0, q_1, q_2, q_3 \) from eqs (F1) as a (non-normalized) quaternion for \( U \). Once normalized, the quaternion determines \( U \) by eq. (7). A still better result—in fact theoretically best—is obtained through singular value decomposition (Szabo 2000).

APPENDIX F: QUATERNION FROM ROTATION

A unit quaternion \( q \) for a rotation matrix \( U = (u_i) \) can be computed directly, avoiding eq. (2). For most rotations \( U \), any of the functions \( q_x, q_y, q_z, q_w \) in eqs (F1) will give a satisfactory quaternion for \( U \). When dealing with real data, however, for which \( U \) may be only approximately orthogonal, the function \( q_w \) (respectively, \( q_x, q_y, q_z \)) should be avoided when the trace of \( U (UX_x, UX_y, UX_z) \) is close to \(-1\). The trace test becomes less relevant if \( U \) has been orthogonalized.

\[
q_w = \left( w, \frac{u_{12} - u_{23}}{4w}, \frac{u_{13} - u_{31}}{4w}, \frac{u_{21} - u_{12}}{4w} \right), \quad w = \frac{1}{2} (1 + u_{11} + u_{22} + u_{33})^{1/2},
\]

\[
q_x = \left( \frac{u_{32} - u_{23}}{4x}, x, \frac{u_{12} + u_{21}}{4x}, \frac{u_{13} + u_{31}}{4x} \right), \quad x = \frac{1}{2} (1 + u_{11} - u_{22} - u_{33})^{1/2},
\]

\[
q_y = \left( \frac{u_{13} - u_{31}}{4y}, \frac{u_{12} + u_{21}}{4y}, y, \frac{u_{23} + u_{32}}{4y} \right), \quad y = \frac{1}{2} (1 - u_{11} + u_{22} - u_{33})^{1/2},
\]

\[
q_z = \left( \frac{u_{21} - u_{12}}{4z}, \frac{u_{13} + u_{31}}{4z}, \frac{u_{23} + u_{32}}{4z}, z \right), \quad z = \frac{1}{2} (1 - u_{11} - u_{22} + u_{33})^{1/2}.
\]

To verify eqs (F1), first note that

\[
w = \frac{1}{2} (1 + \text{trace} \ U)^{1/2}, \quad x = \frac{1}{2} (1 + \text{trace} \ UX_x)^{1/2}, \quad y = \text{etc.}
\]

Next, for any \( 3 \times 3 \) matrix \( U = (u_{ij}) \), define the vector \( K \), which determines the skew symmetric part of \( U \), by

\[
K(U) = (u_{32} - u_{23}, u_{13} - u_{31}, u_{21} - u_{12}).
\]

If \( U \) is a rotation matrix with unit quaternion \( q = (w, x, y, z) \), then, from eq. (7),

\[
4w^2 = 1 + \text{trace} \ U, \quad 4w (x, y, z) = K(U).
\]

If trace \( U \) \( \neq -1 \), then eqs (F4), together with a choice of sign for \( w \), determine \( w, x, y, z \), and hence \( q \). The quaternion \( q \) in eqs (F1a) is the solution \( (w, x, y, z) \) of eqs (F4) with \( w > 0 \).

Applying eqs (F4) to \( UX_x \), rather than to \( U \), gives

\[
4x^2 = 1 + \text{trace} \ UX_x, \quad 4x (-w, -z, y) = K(UX_x),
\]

where \( (w, x, y, z) \) is still the quaternion for \( U \) (not \( UX_x \)). The quaternion \( q \) in eqs (F1b) is the solution of eqs (F5) with \( x > 0 \). Eqs (F1c) and (F1d) are verified similarly.

To see why one of \( q_x, q_y, q_z, q_w \) (eqs F1) is sometimes preferable to the others, we find from eqs (F1a) that

\[
\frac{\partial q_w}{\partial u_{11}} = \frac{i_w}{8w^2},
\]

where the bar denotes quaternion conjugate (Section 2.2). The partial derivatives of \( q_w \) with respect to the other \( u_i \) are like \( 1/(8w^2) \) or \( 1/(4w) \). A perturbation in \( U \) therefore can produce a perturbation in \( q_w \) that is larger by a factor of about

\[
f_w(U) = \frac{1}{8w^2} = \frac{1}{2(1 + \text{trace} \ U)} = \frac{1}{8 \cos^2(\xi/2)}.
\]
where $\xi$ is the rotation angle of $U$. The parameter $f_w$ is therefore a measure of the numerical instability of $q_w$. For rotation angle $\xi = 179^\circ$, the parameter $f_w$ is about 1600, whereas for $\xi = 120^\circ$ it is 0.5. When dealing with real data, one clearly should avoid using $q_w$ when $\text{trace}(U) \approx -1$ or, equivalently, when $\xi$ is large. (Our talk of rotation angle is somewhat loose, since the matrices in question might not be precisely rotations.)

For $q_x, q_y, q_z$, the instability factors are, respectively,

$$f_x(U) = \frac{1}{8x^2} = \frac{1}{2(1 + \text{trace } UX_x)} = \frac{1}{8 \cos^2(\xi_x/2)}$$

$$f_y(U) = \frac{1}{8y^2} = \frac{1}{2(1 + \text{trace } UY_y)} = \frac{1}{8 \cos^2(\xi_y/2)}$$

$$f_z(U) = \frac{1}{8z^2} = \frac{1}{2(1 + \text{trace } UZ_z)} = \frac{1}{8 \cos^2(\xi_z/2)}$$

where as in Section 3.7, the parameters $\xi_x, \xi_y, \xi_z$ are the rotation angles for $UX_x, UY_y, UZ_z$. For a given $U$ the best choice $q_0$ among $q_w, q_x, q_y, q_z$ is therefore determined by which of $U, UX_x, UY_y, UZ_z$ has the largest trace or, equivalently, by which has the smallest rotation angle. The instability factor for $q_0$ is therefore

$$f_0(U) = \frac{1}{8 \cos^2(\xi_0/2)}$$

where $\xi_0$ is the principal axis angle. Thus the level surfaces for $f_0$ are the same as those for $\xi_0$ (Fig. 10). The parameter $f_0$ varies from 0 to 0.5. The minimum occurs at the identity and at the $180^\circ$ rotations about the coordinate axes, and the maximum occurs at the $120^\circ$ rotations about axes in the $(\pm 1, \pm 1, \pm 1)$ directions.

In practice, if $U$ has been orthogonalized, then we usually need not worry about $q_0$; any of $q_w, q_x, q_y, q_z$ will work, unless the relevant trace happens to be extremely close to $-1$. The problems arise when $U$ is only approximately orthogonal; then the choice among $q_w, q_x, q_y, q_z$ can matter.

Eqs (F1) are found in, for example, Hanson (2006). Bar-Itzhack (2000) describes other treatments of this material.

### Mathematical perspective

To avoid having to treat multiple sign possibilities, we define the set $U^+ \subset U$ to consist of the rotations that have a unit quaternion $q = (w, x, y, z)$ for which all of $w, x, y, z$ are positive. From a purely mathematical point of view, all of the functions $q_w, q_x, q_y, q_z$ coincide on $U^+$. Their directional derivatives at $U \in U^+$ must also agree, if the directions are tangent to $U$. But the directional derivatives at $U$ may not agree in directions transverse to $U$, since the functions themselves in general do not agree outside $U$. This is how, for example, $\partial q_w/\partial \xi_1 \neq \partial q_x/\partial \xi_1$, even though $q_w$ and $q_x$ agree on $U^+$. And this is why the choice of $q_w, q_x, q_y, q_z$ can matter.