Appendix C

Basic Equations

This appendix introduces mathematical tools that are needed to understand the topics in this book.

C.1 Vectors

A vector has both magnitude and direction. Vectors are used to represent directional data such as force and velocity. Such a physical entity does not change by the choice of coordinates if we choose, for example, the x-axis in the north or east direction. However, vector components depend on the choice. Consider that $e^{(1)}$, $e^{(2)}$, and $e^{(3)}$ are the unit vectors parallel to the coordinate axes. If a vector $a$ can be written as

$$a = a_1 e^{(1)} + a_2 e^{(2)} + a_3 e^{(3)},$$  \hspace{1cm} (C.1)

the coefficients $a_1$, $a_2$, and $a_3$ are the components of the vector. The direction of the unit vectors depend on the choice of coordinates so that the components are also affected by the choice, although the entity $a$ is constant.

Scalar product

There are in general two types of products between two vectors. They are scalar and vector products. Let $a = (a_1, a_2, a_3)^T$ and $b = (b_1, b_2, b_3)^T$ be arbitrary vectors, then $a \cdot b \equiv |a||b| \cos \theta$ is their scalar product, where $\theta$ is the angle between them.

If we write the angle between $e^{(i)}$ and $e^{(j)}$ as

$$\theta_{ij} = \left( e^{(i)}, e^{(j)} \right),$$  \hspace{1cm} (C.2)

the axes of Cartesian coordinates cross at right angles, so that $\theta_{ij} = 90^\circ$ and $0^\circ$ for $i = j$ and $i \neq j$, respectively. Therefore,

$$e^{(i)} \cdot e^{(j)} = \delta_{ij},$$  \hspace{1cm} (C.3)
where $\delta_{ij}$ is the Kronecker delta. Using this relationship and the expansion in Eq. (C.1), we obtain

$$a \cdot b = a_1 b_1 + a_2 b_2 + a_3 b_3.$$ 

This equation contains components in the right-hand side, and the scalar product does not depend on the orientation of the coordinates because the product is also written only with the physical quantities $|a|$, $|b|$, and $\theta$.

Orthogonal projection of a vector onto a line is often used. Let $a$ and $e$ be an arbitrary vector and the unit vector parallel to the line (Fig. C.1). The component $a$ parallel to $e$ is equal to their scalar product.

**Vector product**

The other type of product between the vectors $a$ and $b$ is defined by the equation

$$a \times b \equiv \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}. \quad (C.4)$$

The result is a vector perpendicular to the plane defined by the vectors $a$ and $b$, and $|a \times b| = |a| |b| \sin \theta$, where $\theta$ is the angle between the vectors. The absolute value is equal to the area of the parallelogram that is spanned by the vectors. The direction of $(a \times b)$ may be remembered by the right-hand rule: $a$ and $b$ are represented by the first and second fingers of the right hand. The thumb then points in the direction of $(a \times b)$. The vector product can also be written as

$$a \times b = \sum_i \left( \sum_{j,k} \epsilon_{ijk} a_j b_k \right) e^{(i)} = \sum_{i,j,k} \epsilon_{ijk} a_j b_k e^{(i)}. \quad (C.5)$$

The coefficient $\epsilon_{ijk}$ is the permutation symbol with the value $\pm 1$ or zero as

$$\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1, \quad \epsilon_{213} = \epsilon_{321} = \epsilon_{132} = -1, \quad \epsilon_{ijk} = 0 \quad \text{(otherwise)}.$$

The symbol satisfies the identity [134]:

$$\sum_k \epsilon_{ijk} \epsilon_{kmn} = \delta_{im} \delta_{jn} - \delta_{in} \delta_{jm} \quad \text{or} \quad \sum_k \epsilon_{ijk} \epsilon_{mkn} = \delta_{im} \delta_{jn} - \delta_{in} \delta_{jm}. \quad (C.6)$$

**Triple products**

Consider a parallelepiped with the sides defined by three vectors $a$, $b$, and $c$. The magnitude of $b \times c$ is identical to the area of one face parallel to the two vectors, and the distance between the face and the other side of the parallelepiped is equal to $a \cdot e$ where $e$ is the unit vector parallel to the vector.
Figure C.1: The scalar product between an arbitrary vector $\mathbf{a}$ and a unit vector $\mathbf{e}$ is equal to the length of the shadow of $\mathbf{a}$ projected onto the line parallel to $\mathbf{e}$.

product: $\mathbf{e} = (\mathbf{b} \times \mathbf{c})/|\mathbf{b} \times \mathbf{c}|$. Therefore, the triple scalar product $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ is equal to the volume of the solid figure. On the other hand, we call $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ a triple vector product.

Let us write the triple scalar product as $\|\mathbf{a}, \mathbf{b}, \mathbf{c}\| \equiv \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$, then

$$\|\mathbf{a}, \mathbf{b}, \mathbf{c}\| = \sum_{i,j,k} \epsilon_{ijk} a_i b_j c_k = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$ 

Accordingly, the triple scalar product is equivalent to the determinant of the matrix that is composed by the vector components. The symbol $\epsilon_{ijk}$ indicate that the sign of product depends on the order of the three vectors:

$$\|\mathbf{a}, \mathbf{b}, \mathbf{c}\| = \|\mathbf{b}, \mathbf{c}, \mathbf{a}\| = \|\mathbf{c}, \mathbf{a}, \mathbf{b}\| = -\|\mathbf{b}, \mathbf{a}, \mathbf{c}\| = -\|\mathbf{c}, \mathbf{b}, \mathbf{a}\| = -\|\mathbf{a}, \mathbf{c}, \mathbf{b}\|.$$ 

C.2 Dumy and free indices

There are two types in indices. Let us use $\mathbf{a} \times \mathbf{b} + \mathbf{c}$ as an example. We have the $i$th component of the terms as

$$\sum_{j,k} \epsilon_{ijk} a_j b_k + c_i.$$ 

Since the symbols $k$ and $j$ in the first term are placed only for the summation, they can be replaced by any other symbol:

$$\sum_{j,k} \epsilon_{ijk} a_j b_k + c_i = \sum_{p,q} \epsilon_{ipq} a_p b_q + c_i.$$ 

The replacement does not affect terms out of the summation. In this case, $c_i$ is not affected. Such a index is called a dummy index. By contrast, the replacement of $i$ in the first term affects the second one. Such an index is called a free index. Replacement of dummy indices sometimes helps to simplify equations$^1$.

$^1$An example is shown in the process to derive Eq. (11.11).
Figure C.2: (a) Rotation of coordinates with fixed vector $\mathbf{a}$. (b) Rotation of the vector with fixed coordinates.

### C.3 Rotation of coordinates

Vector components depend on the orientation of coordinate axes. Let us consider the dependence. Let $O-12$ and $O-1'2'$ be two Cartesian coordinates on a plane with the identical origin, and the components of the vector $\mathbf{a}$ are $a_i$ and $a'_i$ for the coordinates, respectively. If the vector is fixed and the $O-1'$ axis is rotated counterclockwise by the angle $\theta$, the transformation between the components is written by the famous formula

$$
\begin{pmatrix}
a'_1 \\
a'_2
\end{pmatrix} =
\begin{pmatrix}
\cos \theta & \sin \theta \\
-\sin \theta & \cos \theta
\end{pmatrix}
\begin{pmatrix}
a_1 \\
a_2
\end{pmatrix}.
$$

We write the coefficient matrix as

$$
\mathbf{Q} =
\begin{pmatrix}
\cos \theta & \sin \theta \\
-\sin \theta & \cos \theta
\end{pmatrix}.
$$

Now consider the same amount of rotation. However, the coordinates are fixed but the vector $\mathbf{a}$ is rotated counterclockwise. This is the opposite situation to the above discussion, but the components of the vector are also changed by this procedure. Their change is described by Eq. (C.7) with the opposite sense of rotation. Namely, the matrix

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

describes the rotation of the vector.

It is important to distinguish whether coordinates or vectors are rotated. We distinguish them by the symbols $\mathbf{Q}$ and $\mathbf{R}$. They are related by the equations

$$
\mathbf{Q} = \mathbf{R}^{-1}, \quad \mathbf{R} = \mathbf{Q}^{-1}.
$$

The equation

$$
a' = \mathbf{Q} \cdot \mathbf{a}
$$

(C.11)
C.4. MATRICES

represents, accordingly, the rotation of coordinates. The physical entities of \( a \) and \( a' \) are identical in this case. The equation \( a' = Q \cdot a \) indicates the rotation of the entity.

On the other hand, the equation \( a' = R \cdot a \) (C.12)

indicates the rotation of the entity.

The angles between the axes are written as in Eq. (C.2). For the two-dimensional case (Fig. C.2(a)), we have

\[
\begin{align*}
\theta_{11} &= \theta, & \theta_{12} &= \pi/2 - \theta, \\
\theta_{21} &= \pi/2 + \theta, & \theta_{22} &= \theta.
\end{align*}
\]

Using the formula \( \cos(\pi/2 \pm \theta) = \mp \sin \theta \), we obtain

\[
\begin{pmatrix}
\cos \theta_{11} & \cos \theta_{12} \\
\cos \theta_{21} & \cos \theta_{22}
\end{pmatrix} = \begin{pmatrix}
\cos \theta & \sin \theta \\
-\sin \theta & \cos \theta
\end{pmatrix} = Q. \tag{C.13}
\]

\[Q = (\cos \theta_{ij}) = \left( e^{i(i)} \cdot e^{(j)} \right). \tag{C.14}\]

This relationship also holds for rotations in three-dimensional space. \( Q \) and \( R \) are called orthogonal matrices which satisfy the following formulas:

\[
\begin{align*}
Q^{-1} &= Q^T, & Q \cdot Q^T &= I, \\
R^{-1} &= R^T, & R \cdot R^T &= I. \tag{C.15}
\end{align*}
\]

Combining Eqs. (C.10), (C.13) and (C.15), we see that \( R \) has the form

\[
R = \begin{pmatrix}
\cos \theta & -\sin \theta & 0 \\
\sin \theta & \cos \theta & 0 \\
0 & 0 & 1
\end{pmatrix}. \tag{C.16}
\]

C.4 Matrices

Many physical quantities are represented by matrices in the models of tectonics. Let \( A \) and \( B \) be arbitrary square matrices, then their products are defined by the equations:

\[
A \cdot B = \left( \sum_k A_{ik} B_{kj} \right), & A : B = \left( \sum_{ij} A_{ij} B_{ij} \right).
\]

The following formulas of matrices are used in this book:

\[
A \cdot I = A, \tag{C.17}
\]

\[
|A \cdot B| = |A||B|, \tag{C.18}
\]

\[
(A \cdot B)^T = B^T \cdot A^T. \tag{C.19}
\]

\[
|A^T| = |A|. \tag{C.20}
\]

\[
(A^T)^{-1} = (A^{-1})^T. \tag{C.21}
\]

\[
(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}. \tag{C.22}
\]
Symmetry and antisymmetry

Consider a square matrix with the shape
\[
A = \begin{pmatrix}
a & b & c \\
d & e & f \\
g & h & i \\
\end{pmatrix}.
\]

This can be divided into two parts: \( A = B + C \), where
\[
B = \frac{A + A^T}{2} = \frac{1}{2} \begin{pmatrix}
2a & b + d & c + g \\
d + b & 2e & f + h \\
g + c & h + f & 2i \\
\end{pmatrix},
\]
\[
C = \frac{A - A^T}{2} = \frac{1}{2} \begin{pmatrix}
0 & b - d & c - g \\
d - b & 0 & f - h \\
g - c & h - f & 0 \\
\end{pmatrix}.
\]

As \( B = B^T \), \( B \) is a symmetric matrix. On the other hand, \( C = -C^T \). This is called an antisymmetric matrix. Every square matrix can be decomposed into symmetric and antisymmetric matrices.

C.5 Eigenvalues and eigenvectors

If \( A \) is a square matrix and \( u \) is a vector, the dot product \( A \cdot u \) is a vector which is parallel to \( u \) for special \( u \). Those are the eigenvectors of the matrix. If they are parallel, \( \lambda \) refers to the ratio between \( u \) and \( A \cdot u \), and the ratio is called an eigenvalue. Linear algebra says that an \( n \times n \) matrix has \( n \) pairs of eigenvectors and eigenvalues, and the latter is obtained by solving the following algebraic equation, called a characteristic equation,
\[
|A - \lambda I| = 0. \tag{C.23}
\]

The matrix has real eigenvalues if the matrix components are real and \( A \) is symmetric. Then all the eigenvectors are perpendicular to each other. The above equation does not explicitly contain components that depend on the choice of coordinates, so that neither the eigenvalues nor the eigenvectors depend on the choice. Specifically, the characteristic equation of the \( 2 \times 2 \) matrix
\[
A = \begin{pmatrix}
a & c \\
c & b \\
\end{pmatrix} \tag{C.24}
\]
is \( \lambda^2 - (a + b)\lambda + ab - c^2 = 0 \), so that
\[
\lambda_1 = \frac{(a + b) + \sqrt{(a - b)^2 + 4c^2}}{2},
\]
\[
\lambda_2 = \frac{(a + b) - \sqrt{(a - b)^2 + 4c^2}}{2}.
\]
are the eigenvalues, and the eigenvectors are
\[
\left( \frac{-a + b + \sqrt{(a-b)^2 + 4c^2}}{2c}, 1 \right)^T, \\
\left( \frac{-a + b - \sqrt{(a-b)^2 + 4c^2}}{2c}, 1 \right)^T.
\]

How about the eigenvalues and eigenvectors of \( A' = A - pl \), where \( p \) is an arbitrary scalar parameter? Let \( \lambda' \) be the eigenvalues of \( A' \), then we have
\[
|A' - \lambda'I| = |A - (p + \lambda)I|.
\]
Therefore, \( A \) and \( A' \) have the same eigenvectors, but their eigenvalues differ by the constant \( p \).

Let \( \lambda_1, \lambda_2, \) and \( \lambda_3 \) be the real eigenvectors of the \( 3 \times 3 \) symmetric matrix \( A \), and \( u^{(1)}, u^{(2)}, \) and \( u^{(3)} \) are the unit vectors parallel to the eigenvectors, then the unit vectors make right angles to each other. The unit vectors can be used as base vectors of the new orthogonal coordinates. Their components make up an orthogonal matrix
\[
Q = \begin{pmatrix}
\begin{pmatrix} u_1^{(1)} \\ u_2^{(1)} \\ u_3^{(1)} \end{pmatrix} & \begin{pmatrix} u_1^{(2)} \\ u_2^{(2)} \\ u_3^{(2)} \end{pmatrix} & \begin{pmatrix} u_1^{(3)} \\ u_2^{(3)} \\ u_3^{(3)} \end{pmatrix}
\end{pmatrix} = \begin{pmatrix}
[u^{(1)}]_T \\ [u^{(2)}]_T \\ [u^{(3)}]_T
\end{pmatrix},
\]
where \( u_i^{(j)} \) is the \( i \)th component of the \( j \)th vector that is evaluated in the old coordinates. Then the coordinate transformation from the old to new is represented by the matrix \( Q \) (see Exercise A1). Using this, we can reshape the matrix \( A \) to a diagonal matrix
\[
Q \cdot A \cdot Q^T = \begin{pmatrix}
\lambda_1 & 0 & 0 \\
0 & \lambda_2 & 0 \\
0 & 0 & \lambda_3
\end{pmatrix}, \tag{C.25}
\]
in which the diagonal components are identical to the eigenvalues of \( A \). If we use the new coordinates, the matrix becomes the diagonal form. In this case, the power of \( A \) is
\[
A^n = \begin{pmatrix}
\lambda_1^n & 0 & 0 \\
0 & \lambda_2^n & 0 \\
0 & 0 & \lambda_3^n
\end{pmatrix}. \tag{C.26}
\]
Accordingly, the diagonal form holds for any \( n \), indicating that eigenvectors do not change.

If \( f(A) \) is a \( 3 \times 3 \) matrix and is a function of \( A \), and if the two matrices have the same eigenvectors, the \( n \)th term of Taylor expansion of this function encompasses \( A^n \) with the above shape. Therefore,
\[
f(A) = \begin{pmatrix}
f(\lambda_1) & 0 & 0 \\
0 & f(\lambda_2) & 0 \\
0 & 0 & f(\lambda_3)
\end{pmatrix}. \tag{C.27}
\]
APPENDIX C. BASIC EQUATIONS

Given positive eigenvalues, we have

\[
\log A = \begin{pmatrix}
\log \lambda_1 & 0 & 0 \\
0 & \log \lambda_2 & 0 \\
0 & 0 & \log \lambda_3
\end{pmatrix}, \tag{C.28}
\]

\[
\sqrt{A} = \begin{pmatrix}
\sqrt[2]{\lambda_1} & 0 & 0 \\
0 & \sqrt[2]{\lambda_2} & 0 \\
0 & 0 & \sqrt[2]{\lambda_3}
\end{pmatrix} \tag{C.29}
\]

Given a non-diagonal matrix \( A \), we get the matrix-valued function of \( A \) from the formula

\[
f(A) = Q^T \cdot \begin{pmatrix}
f(\lambda_1) & 0 & 0 \\
0 & f(\lambda_2) & 0 \\
0 & 0 & f(\lambda_3)
\end{pmatrix} \cdot Q,
\]

if \( A \) is real and symmetric, or simply diagonalizable [148].

**Ellipse and Ellipsoid**

The equation of an ellipsoid is

\[
\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} + \frac{x_3^2}{c^2} = 1.
\]

This is rewritten to the simple form

\[
x \cdot A \cdot x = 1, \tag{C.30}
\]

where \( x = (x_1, x_2, x_3)^T \) and

\[
A = \begin{pmatrix}
1/a^2 & 0 & 0 \\
0 & 1/b^2 & 0 \\
0 & 0 & 1/c^2
\end{pmatrix}.
\]

Equation (C.30) does not explicitly contain components so the equation does not depend on the choice of coordinates. The matrix \( A \) prescribes the shape and principal axis of the ellipsoid. If \( A = I \), the equation represents a unit sphere. If a 3 \( \times \) 3 matrix \( B \) with positive eigenvalues,

\[
B = \begin{pmatrix}
\lambda_1 & 0 & 0 \\
0 & \lambda_2 & 0 \\
0 & 0 & \lambda_3
\end{pmatrix}
\]

represents an ellipsoid, its principal radii \( a, b, \) and \( c \) are determined by the equations

\[
a = 1/\sqrt{\lambda_1}, \quad b = 1/\sqrt{\lambda_2}, \quad c = 1/\sqrt{\lambda_3}. \tag{C.31}
\]

The principal axes are identical with the eigenvectors of the matrix \( B^{-1/2} \).

Consider two vectors \( x \) and \( y \) that are related through the equation \( y = Q \cdot x \), where \( Q \) is an orthogonal matrix. Then \( x = Q^T \cdot y \). Substituting this one into the equation of ellipse \( x \cdot A \cdot x = 1 \), we obtain

\[
y \cdot (Q \cdot A \cdot Q^T) \cdot y = 1.
\]
This equation represents the same ellipsoid with the equation \( x \cdot A \cdot x = 1 \). Therefore, the matrix components are transformed by the rotation of coordinates as

\[
A' = Q \cdot A \cdot Q^T. \tag{C.32}
\]

If the coordinates are fixed and the entity that the matrix represents is rotated, the result is the entity represented by the matrix

\[
A' = R^T \cdot A \cdot R.
\]

### C.6 Tensors

Mathematically, a tensor is defined as a linear transformation from between vectors. Within this book we are able to regard tensors as physical entities that are represented by 3 \( \times \) 3 or 2 \( \times \) 2 matrices, and the equations shown in the previous sections in this appendix apply to tensors. Symmetric and antisymmetric tensors are represented by symmetric and antisymmetric matrices. Zero and identical matrices, \( O \) and \( I \), represent zero and unit tensors. Orthogonal matrices \( Q \) and \( R \) stands for orthogonal tensors.

#### Higher-order tensors

Consider vectors \( a \) and \( b \) and the product of their components \( T_{ij} = a_i b_j \). Let us write this product as \( T = ab \). According to Eq. (C.11), the vectors are transformed by the rotation of coordinates as

\[
a'_i = \sum_{p=1}^{3} Q_{ip} a_p, \quad b'_j = \sum_{q=1}^{3} Q_{jq} b_q.
\]

Therefore, the product \( T \) is transformed as

\[
T'_{ij} = a'_i b'_j = \sum_{p,q} Q_{ip} Q_{jq} a_p b_q = \sum_{p,q} Q_{ip} T_{pq} \left( Q_{jq} \right)^T. \tag{C.33}
\]

This is \( T' = Q \cdot T \cdot Q^T \), and is in the same form as Eq. (C.32)—the quantity is a tensor. One can see by comparing Eqs. (C.33) and (C.11) that a vector with a single index is transformed by one \( Q \) and a tensor with two indices is transformed by two \( Q \)s. Generalizing this observation, we define a \( n \)th-order tensor \( T_{ijk...} \) by the behavior when the coordinates are rotated as

\[
T'_{ijk...} = \sum_{p,q,r...} Q_{ip} Q_{jq} Q_{kr} \cdots T_{pq...}, \tag{C.34}
\]

where \( T_{ijk...} \) has \( n \) indices and the right-hand side of this equation has \( n \) \( Q \)s. Scalar quantities like temperature are not changed by the rotation, so that the quantities are 0th-order tensors. Vectors are first-order tensors, and tensors that are represented by square matrices are second-order tensors.
This is the definition of tensors with their behavior for rotation. The product of a vector and tensor makes a third-order tensor such as \( T_{ij}a_k \) and \( a_i T_{jk} \).

Tensors of the second-order are transformed as
\[
T' = Q \cdot T \cdot Q^T.
\] (C.35)
when coordinates are rotated. The equation
\[
T' = R \cdot T \cdot R^T
\]
stands for the rotation of the second order tensor \( T \) itself.

**Projector**

In general, the product of two vectors becomes a second-order tensor
\[
ab = \begin{pmatrix}
a_1b_1 & a_1b_2 & a_1b_3 \\
a_2b_1 & a_2b_2 & a_2b_3 \\
a_3b_1 & a_3b_2 & a_3b_3
\end{pmatrix}.
\]

Using the base vectors of coordinates \( e^{(i)} \), we have \( a = a_1e^{(1)} + a_2e^{(2)} + a_3e^{(3)} \) and \( b = b_1e^{(1)} + b_2e^{(2)} + b_3e^{(3)} \), so that
\[
ab = \sum_{ij} (a_ib_j) e^{(i)}e^{(j)}.
\]
The products of the base vectors, for example
\[
e^{(1)}e^{(2)} = \begin{pmatrix}
0 & 1 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{pmatrix},
\]
act as the bases for second-order tensors:
\[
T = \sum_{i,j} T_{ij}e^{(i)}e^{(j)}.
\] (C.36)

For a unit vector \( e = (1, 0, 0)^T \),
\[
ee = \begin{pmatrix}
1 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{pmatrix}, \quad l - ee = \begin{pmatrix}
0 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{pmatrix}.
\]

Therefore, the tensor-valued functions
\[
P^\parallel(e) = ee, \quad P^\perp(e) = l - ee
\] (C.37) (C.38)
are called projects: if \( a \) is an arbitrary vector, \( P^\parallel(e) \cdot a \) and \( P^\perp(e) \cdot a \) are the vector quantities that are the orthogonal projection of \( a \) onto the line parallel to \( e \) and onto the plane normal to \( e \), respectively (Fig. C.1). \( P^\perp \) is also called an elementary orthogonal projector [148]. Note that \( P^\parallel(e) = l - P^\perp(e) \) and \( |P^\parallel(e) \cdot a| = e \cdot a \).
Isotropic and anisotropic tensors

There are tensors that are invariant for rotation. They are the scalar multiplication of the unit tensor \(T = pl\). For any rotation \(Q\), \(T' = Q \cdot T \cdot Q^T = pQ \cdot Q^T = pl\). This indicates that the tensors look identical regardless of viewpoint. Such tensors are called isotropic. Most tensors are not isotropic and are called anisotropic. Quartz single crystals are examples that have crystallographic anisotropy.

Dot and colon products

Tensors with an order greater than two have free indices. The operation of identifying two indices of a tensor and so summing on them is known as contraction. Linear transformation of vectors \(T \cdot a = \sum_i T_{ij} a_j\) is a contraction of the third-order tensor \(T_{ij} a_j\). The contraction of this tensor is not unique: \(\sum_j T_{ji} a_j\) is the other one. The scalar product of vectors is a contraction and is obviously commutative: \(a \cdot b = b \cdot a\). If \(a\) and \(b\) are matrices with 3 rows and 1 column, this is \(a^T b\). The matrix components depend on the frame of reference. The vectors are physical quantities independent of the frames so that the superscript “T” is not included in \(a \cdot b\). It is also an important formula that \(x \cdot A \cdot y = y \cdot A \cdot x\) if \(A\) is symmetric.

The matrix multiplication \(A \cdot B\) can be regarded as a contraction of the fourth-order tensor \(A_{ij} B_{kl}\) by identifying the inner and adjacent suffixes \(j\) and \(k\), and a derivation of a lower-order tensor. In this respect, the dot is placed similarly to the scalar product of vectors. The contraction of tensors results in lower tensors. If two indices in the fourth-order tensor are identified, we write the contraction as \(A : B\), resulting in a zeroth-order tensor (scalar),

\[A : B = \sum_{i,j} A_{ij} B_{ji}.\]

Second-order tensors only have the contraction

\[\text{trace } T \equiv \sum_i T_{ii} = T : I = I : T.\]

Let us look over the relationship between the trace of tensors and triple scalar products. We write the triple scalar product of vectors \(a, b,\) and \(c\) as \(\|a, b, c\| \equiv a \cdot (b \times c)\), while the following equation holds for an arbitrary second-order tensor \(T\):

\[\|(T \cdot a) \cdot b, c\| = \left\| T \cdot \left( \sum_i a_i e^{(i)} \right) \cdot \sum_j b_j e^{(j)} \cdot \sum_k c_k e^{(k)} \right\| = \sum_{i,j,k} a_i b_j c_k \left\| T \cdot e^{(i)} \cdot e^{(j)} \cdot e^{(k)} \right\|\]

\[= \sum_{i,j,k} a_i b_j c_k \sum_{pqr} T_{ip} e^{(i)}_p e^{(j)}_q e^{(k)}_r = \sum_{i,j,k} a_i b_j c_k T_{ij} \epsilon_{ijk} e^{(i)}_i e^{(j)}_j e^{(k)}_k\]

\[= \sum_{i,j,k} \epsilon_{ijk} T_{ij} a_i b_j c_k\]
APPENDIX C. BASIC EQUATIONS

\[ \| (\mathbf{T} \cdot \mathbf{a}) \cdot \mathbf{b}, \mathbf{c} \| + \| \mathbf{a} \cdot (\mathbf{T} \cdot \mathbf{b}) \cdot \mathbf{c} \| + \| \mathbf{a} \cdot \mathbf{b} \cdot (\mathbf{T} \cdot \mathbf{c}) \| = \sum_{i,j,k} (T_{ii} + T_{jj} + T_{kk}) \epsilon_{ijk} a_i b_j c_k \]

\[ = (T_{11} + T_{22} + T_{33}) a_1 b_2 c_3 + (T_{22} + T_{33} + T_{11}) a_2 b_3 c_1 \]
\[ + (T_{33} + T_{11} + T_{22}) a_3 b_1 c_2 - (T_{22} + T_{11} + T_{33}) a_2 b_1 c_3 \]
\[ - (T_{33} + T_{22} + T_{11}) a_3 b_2 c_1 - (T_{11} + T_{33} + T_{22}) a_1 b_3 c_2 \]
\[ = \text{trace} \mathbf{T} \sum_{ijk} \epsilon_{ijk} a_i b_j c_k \]

Accordingly, we arrive at an expression of the trace,
\[ \text{trace} \mathbf{T} = \frac{1}{\| \mathbf{a}, \mathbf{b}, \mathbf{c} \|} \left[ \| (\mathbf{T} \cdot \mathbf{a}) \cdot \mathbf{b}, \mathbf{c} \| + \| \mathbf{a} \cdot (\mathbf{T} \cdot \mathbf{b}) \cdot \mathbf{c} \| + \| \mathbf{a} \cdot \mathbf{b} \cdot (\mathbf{T} \cdot \mathbf{c}) \| \right]. \quad (C.39) \]

**Invariants**

Let \( A_i \) be the eigenvalues of tensor \( \mathbf{A} \). It should be noted that \( A_i \)'s are not the components of a vector, because coordinate rotations do not affect eigenvalues. Namely, \( A_i \)'s are invariant for coordinate rotations.

Now suppose that the cubic equation
\[ \lambda^3 - A_1 \lambda^2 + A_2 \lambda - A_3 = 0 \quad (C.40) \]
is the characteristic equation of the tensor \( \mathbf{A} \). The eigenvalues \( A_i \)'s are the solutions of this equation. According to the relationship between the coefficients and solutions of cubic equation, we have
\[ A_1 = A_1 + A_2 + A_3, \quad (C.41) \]
\[ A_2 = A_1 A_2 + A_2 A_3 + A_3 A_1, \quad (C.42) \]
\[ A_3 = A_1 A_2 A_3. \quad (C.43) \]

Given a symmetric tensor, all the eigenvalues are real and the coefficients \( A_1, A_2, \) and \( A_3 \) become real. The above three equations indicate that the coefficients are independent of the coordinate system. Such scalar quantities are called the invariants of the tensor. Since \( A_i \)'s are invariants, any scalar-valued function of \( A_i \)'s is invariant. Especially, those defined by Eqs. (C.41)–(C.43) are called the first, second, and third basic invariants. They can be also written as
\[ A_1 = A_{11} + A_{22} + A_{33} = \text{trace} \mathbf{A}, \]
\[ A_2 = \left[ \begin{array}{cc} A_{22} & A_{23} \\ A_{32} & A_{33} \end{array} \right] + \left[ \begin{array}{cc} A_{33} & A_{31} \\ A_{13} & A_{11} \end{array} \right] + \left[ \begin{array}{cc} A_{11} & A_{12} \\ A_{21} & A_{22} \end{array} \right], \]
\[ = \frac{1}{2} \left[ (\text{trace} \mathbf{A})^2 - \text{trace} \mathbf{A}^2 \right] \]
\[ A_3 = |\mathbf{A}|. \quad (C.44) \]

---

\(^2\)Compare the sign of the third term in the left-hand side of Eq. (C.40) with that of Eq. (C.48).
The cubic equation
\[ -\lambda^3 + A_{I} \lambda^2 - A_{II} \lambda + A_{III} = 0 \]  \quad (C.45)
is identical to the characteristic equation of \( A \) [148].

**Cayley-Hamilton theorem**

The *Cayley-Hamilton theorem* provides the basis to consider the relation between stress and strain tensors. The cubic equation
\[ A^3 - A_I A^2 + A_{II} A - A_{III} I = 0 \]  \quad (C.46)
stands for the theorem, where the coefficients in the left-hand side are the basic invariants of the tensor \( A \). Note that by replacing \( A, I \) and \( O \) in the above equation with \( \lambda, 1 \) and 0, respectively, we obtain the characteristic equation (Eq. (C.40)).

The theorem is demonstrated as follows. According to Eq. (C.27), the tensor
\[ T = A^3 - A_I A^2 + A_{II} A - A_{III} I \]has the eigenvalues \( T_i = \lambda_i^3 - A_I \lambda_i^2 + A_{II} \lambda_i - A_{III} \), where \( i = 1, 2 \) and 3. Equation (C.45) indicates that all the eigenvalues \( T_i \) vanish. Thus \( T = O \).

**Deviatoric tensors**

Symmetric tensors with zero trace are called *deviatoric tensors*. Let \( A \) be an arbitrary \( 3 \times 3 \) symmetric tensor, then
\[ A' = A - \left( \frac{\text{trace} A}{3} \right) I \]  \quad (C.47)
is a deviatoric tensor. This equation indicates that symmetric tensors can be decomposed into isotropic and anisotropic tensors, \( A = (\text{trace} A/3)I + A' \). Therefore, deviatoric tensors derived from isotropic tensors vanish.

It is a convention of continuum mechanics that the characteristic equation of deviatoric tensors is written as
\[ \lambda^3 - A'_I \lambda^2 - A'_{II} \lambda - A'_{III} = 0. \]  \quad (C.48)
Contrary to the characteristic equation of ordinary tensors (Eq. (C.40)), the third term in the left-hand side of this equation has negative sign. Accordingly, the second basic invariant of deviatoric tensors has the opposite sign to that of ordinary tensors. The first invariant of \( A' \) always vanishes, \( A'_I = \text{trace} A' = 0 \). The second invariant satisfies
\[ A'_{II} = - \left( A'_{11} A'_{22} + A'_{21} A'_{31} + A'_{31} A'_{12} \right) \quad \text{(C.49)} \]
\[ = - \frac{1}{2} \left[ (\text{trace} A')^2 - (A')^2 \right] = \frac{1}{2} (A' : A') \quad \text{(C.50)} \]
\[ = \frac{1}{2} \left( A_{11}'^2 + (A_{22}')^2 + (A_{33}')^2 \right) + (A_{12}')^2 + (A_{23}')^2 + (A_{31}')^2 \geq 0. \]  \quad (C.51)
APPENDIX C. BASIC EQUATIONS

Figure C.3: Schematic illustration to explain the sign of the second basic invariant of deviatoric tensors. Closed circles lie in the gray zone that runs from upper left to lower right.

where $A'_1$, $A'_2$, and $A'_3$ are the eigenvalues of $A'$ and, according to Eq. (C.47),

$$A'_i = A_i - \frac{\text{trace } A}{3}. \quad (C.52)$$

One can see this inequality by considering the correlation coefficient of two variables $x$ and $y$. Suppose that six pairs $(A'_1, A'_2), (A'_2, A'_1), (A'_2, A'_3), (A'_3, A'_2), (A'_3, A'_1), (A'_1, A'_3)$ represent six measurements of two phenomena. If we regard the pairs as points on a coordinate plane, we have a picture like Fig. C.3. The figure is drawn taking into account the constraint $A'_1 = 0$ which means that the mean of $A'_1$, $A'_2$, and $A'_3$ is zero. The points are plotted on both sides of the abscissa and ordinate, and in the gray zone in the figure. The zone indicates a negative correlation between the observed phenomena. If the averages of $x$ and $y$ are both zero, the correlation coefficient is (constant) $\times \sum_i x_i y_i$, where $x_i$ and $y_i$ are the $i$th measurements of $x$ and $y$. This summation has the same form as in Eq. (C.49). The correlation is negative so we conclude that $A''_1 \geq 0$.

The above comparison of $A''_1$ with the correlation coefficient indicates that the second invariant of a deviatoric tensor is a measure of how the eigenvalues of the tensor disperse. Due to Eq. (C.52), the invariant is also the dispersion of eigenvalues of $A$. If the eigenvalues are identical, then $A''_1 = 0$ and $A' = 0$, meaning the tensor $A$ is isotropic.

### C.7 Polar decomposition

If a real second-order tensor $F$ satisfies $|F| > 0$, then there are the orthogonal tensor $^3R$ and symmetric tensors $U$ and $V$ such that

$$F = R \cdot U = V \cdot R.$$

\(^3\)This should not be represented by $Q$ which indicates the rotation of coordinates.
It is important that \( R, U \) and \( V \) are uniquely determined from \( F \). These statements are collectively called the polar decomposition theorem. \( R \cdot U \) and \( V \cdot R \) are called right and left polar decomposition, respectively.

To understand the theorem, let \( v = (v_1, v_2, v_3)^T \) be an arbitrary vector. Using a given \( F \), we define another vector \( w = F \cdot v \). As \( |F| > 0 \), \( w \) vanishes only if \( v = 0 \). We assume that \( v \neq 0 \). Considering Eq. (C.19),

\[
\mathbf{w} \cdot \mathbf{w} = (F \cdot \mathbf{v}) \cdot (F \cdot \mathbf{v}) = \mathbf{v} \cdot (F^T \cdot F) \cdot \mathbf{v} > 0.
\]

Accordingly, we define the tensor \( C \equiv F^T \cdot F \). This is symmetric, because \( C^T = (F^T \cdot F)^T = F^T \cdot F = C \). In this derivation, we have used the distributive property of transposition (Eq. (C.19)). The real symmetric tensor \( C \) has real eigenvalues and three eigenvectors, so that we choose the Cartesian coordinates parallel to the eigenvectors, and the tensor has the diagonal form

\[
C = \begin{pmatrix}
C_1 & 0 & 0 \\
0 & C_2 & 0 \\
0 & 0 & C_3 \\
\end{pmatrix}.
\]

Equation (C.53) becomes \( C_1 v_1^2 + C_2 v_2^2 + C_3 v_3^2 > 0 \). This inequality holds only if all the diagonal components are positive. So, we introduce the tensors

\[
\mathbf{U} = \sqrt{C} \quad \text{and} \quad \mathbf{R} = F \cdot U^{-1}.
\]

The former is a real symmetric tensor and satisfies \( C = U^2 \), and the latter is a second-order tensor. If this is shown to be an orthogonal tensor, the right polar decomposition may be proved. Using the definition of \( R \), we have

\[
\mathbf{R} \cdot \mathbf{R}^T = \left( U^{-1} \right)^T \cdot F^T \cdot F \cdot U^{-1} = U^{-1} \cdot C \cdot U^{-1} = U^{-1} \cdot \mathbf{U} \cdot \mathbf{U}^{-1} = I.
\]

Therefore, \( R \) is indeed an orthogonal tensor. The left polar decomposition is proved in the same way.

There are two remaining problems: (1) the uniqueness of the decomposition and (2) the relation between \( U \) and \( V \). Suppose that it is not unique and that we have \( F = R \cdot U = \tilde{R} \cdot \tilde{U} \). Using the first decomposition, we obtain

\[
F^T \cdot F = U^T \cdot R^T \cdot R \cdot U = U \cdot U.
\]

Similarly, \( F^T \cdot F = \tilde{U} \cdot \tilde{U} \), so that \( U \cdot U = \tilde{U} \cdot \tilde{U} \). This indicates \( U = \tilde{U} \). \( U \) is unique. Hence, \( R = F \cdot U^{-1} \) is also unique. The uniqueness of the left polar decomposition is proved in the same way. Their uniqueness leads to the relationship between \( U \) and \( V \):

\[
\mathbf{R}^T \cdot \mathbf{V} \cdot \mathbf{R} = \mathbf{U}, \quad \mathbf{V} = \mathbf{R} \cdot \mathbf{U} \cdot \mathbf{R}^T.
\]
C.8 Calculus

Vectors and tensors can depend on position and form vector and tensor fields. We use the nabla operator

\[ \nabla \equiv \sum_i e_i \frac{\partial}{\partial x_i} \]

to get the spatial derivatives of the field. \( \nabla \cdot \) and \( \nabla \times \) are divergence and curl operators, respectively, and are defined as

\[ \nabla \cdot \equiv \sum_i \frac{\partial}{\partial x_i}, \quad \nabla \times \equiv \sum_j \epsilon_{ijk} e_j \frac{\partial}{\partial x_j} . \]

Applying the \( \nabla \) and \( \nabla^2 \) operators to a scalar function \( F(x) \), we obtain

\[ \nabla F = \sum_i e_i \frac{\partial F}{\partial x_i}, \quad \nabla^2 F = \sum_i \frac{\partial^2 F}{\partial x_i^2} = \sum_{i,j} \delta_{ij} \frac{\partial a_i}{\partial x_j} = \sum_i \frac{\partial^2}{\partial x_i^2} \]

where

\[ \nabla^2 = \nabla \cdot \nabla = \sum_i \left( e_i \frac{\partial}{\partial x_i} \right)^2 = \sum_i \frac{\partial^2}{\partial x_i^2} \]

is the Laplacian operator. The vector quantity \( \nabla F \) is the gradient of \( F \), namely, \( F \) becomes larger in the direction indicated by the vector. The operator \( \nabla^4 \) appears when the flexure of the lithosphere and viscous flow are considered. In the two dimensional case, it reads

\[ \nabla^4 = \frac{\partial^4}{\partial x_1^4} + 2 \frac{\partial^2}{\partial x_1^2 \partial x_2^2} + \frac{\partial^4}{\partial x_2^4} . \]  

We use the relationship \( e_i \cdot e_j = \delta_{ij} \) and the independence of the base vectors on the position \( (\partial e_i / \partial x_j = 0) \) to differentiate vector field \( a(x) = \sum_i a_i(x) e_i \):

\[ \nabla \cdot a = \sum_i e_i \frac{\partial a_i}{\partial x_i}, \quad \nabla \times a = \sum_j e_j \frac{\partial a_i}{\partial x_j} \times \sum_k e_k (\frac{\partial a_j}{\partial x_k}) = \sum_{i,j,k} \epsilon_{ijk} \frac{\partial a_k}{\partial x_j} (e_i \times e_j) \]

\[ = \left( \frac{\partial a_3}{\partial x_2} - \frac{\partial a_2}{\partial x_3} - \frac{\partial a_2}{\partial x_1} + \frac{\partial a_1}{\partial x_2} \right) ^T, \]

\[ \nabla a = \sum_i e_i \frac{\partial}{\partial x_i} \sum_j e_j a_j = \sum_{i,j} \left( \frac{\partial a_i}{\partial x_j} \right) e_i e_j . \]
The results are a scalar, vector, and second-order tensor field, respectively. Similarly, the nabla operator is contracted with a second-order tensor \( \mathbf{\sigma} = (\sigma_{ij}) \), to yield a vector

\[
\nabla \cdot \mathbf{\sigma} = \sum_i e^{(i)} \frac{\partial}{\partial x_i} \cdot \left[ \sum_{j,k} \sigma_{jk} e^{(j)} e^{(k)} \right] = \sum_{i,j} e^{(i)} \frac{\partial \sigma_{ji}}{\partial x_j}. \tag{C.56}
\]

The curl of the \( \nabla F \) always vanishes:

\[
\nabla \times \nabla F = 0, \tag{C.57}
\]

because its first component is

\[
\frac{\partial}{\partial x_2} \frac{\partial F}{\partial x_3} - \frac{\partial}{\partial x_3} \frac{\partial F}{\partial x_2} = 0
\]

and, similarly, the other components vanish. Applying the Laplacian operator to a vector \( \mathbf{a} \) yields a vector

\[
\nabla^2 \mathbf{a} = \sum_i e^{(i)} \frac{\partial^2 a_j}{\partial x_i^2}.
\]

Since \( \mathbf{x} \cdot \mathbf{A} \cdot \mathbf{x} \) is a scalar, \( \nabla \mathbf{x} \cdot \mathbf{A} \cdot \mathbf{x} \) is a vector. If \( \mathbf{x} \) is an independent variable, its components are also independent of each other. Therefore, we have \( \partial x_i / \partial x_j = \delta_{ij} \) or

\[
\nabla \mathbf{x} = \frac{\partial \mathbf{x}}{\partial \mathbf{x}} = \mathbf{I}.
\]

The \( i \)th component of \( \nabla \mathbf{x} \cdot \mathbf{A} \cdot \mathbf{x} \) is

\[
\frac{\partial}{\partial x_i} \sum_{j,k} x_j A_{jk} x_k = \sum_{j,k} \left( \delta_{ij} A_{jk} x_k + x_j A_{jk} \delta_{ki} \right) = \sum_k A_{ik} x_k + \sum_j x_j A_{ji} \tag{C.58}
\]

\[
= \sum_j \left[ A_{ij} x_j + (\mathbf{A}^T)_{ij} x_j \right]. \tag{C.59}
\]

Therefore, we have

\[
\nabla \left( \mathbf{x} \cdot \mathbf{A} \cdot \mathbf{x} \right) = (\mathbf{A} + \mathbf{A}^T) \cdot \mathbf{x}. \tag{C.60}
\]

If \( \mathbf{A} \) is symmetric, then

\[
\nabla \left( \mathbf{x} \cdot \mathbf{A} \cdot \mathbf{x} \right) = 2\mathbf{A} \cdot \mathbf{x}. \tag{C.61}
\]

We use formulae that exchange line, surface, and volume integrals\(^4\). *Gauss’s divergence theorem* provides a formula to interchange volume and surface integrals. Suppose \( V \) is a volume with a smooth closed surface \( S \) (Fig. C.4). On the surface, we define the surface element \( dS \) where the unit vector \( \mathbf{N} \) is the outward normal. If \( \mathbf{A} \) is a differentiable quantity, the theorem says that

\[
\int_V \nabla \mathbf{A} \, dV = \int_S \mathbf{N} \mathbf{A} \, dS. \tag{C.62}
\]

\(^4\)See [134] for derivation of integral formulae.
Figure C.4: Schematic picture for an explanation of Gauss’s divergence theorem and Stokes’ theorem. (a) A rock mass with a volume $V$ is surrounded by a closed and smooth surface $S$. The unit vector $N$ is the outward normal on the surface element $dS$. (b) Closed curve $C$ bounds a surface $S$. Unit vector $t$ is tangent to $C$.

Any scalar, vector, or tensor field that $A$ represents satisfies this equation. In the case of a vector or tensor field, we have

$$\int_V \nabla \cdot A \, dV = \int_S N \cdot A \, dS.$$  \hspace{1cm} \text{(C.63)}

We also obtain the following formula from Gauss’s divergence theorem:

$$\int_V (\phi \nabla^2 \psi + \nabla \phi \cdot \nabla \psi) \, dV = \int_S \phi \frac{\partial \psi}{\partial N} \, dS$$

$$= \int_V (\phi \nabla^2 \psi + \phi \nabla^2 \psi) \, dV = \int_S \left( \phi \frac{\partial \psi}{\partial N} - \psi \frac{\partial \phi}{\partial N} \right) \, dS,$$

where $\phi$ and $\psi$ are scalar fields and $\partial / \partial N \equiv (NN) \cdot \nabla$ is the gradient operator in the direction of $N$.

Stokes’ theorem exchanges a surface integral along a closed curve $C$ for a surface integral over the open surface $S$ that is bordered by the closed curve (Fig. C.4). For the vector field $a$, it reads

$$\oint_C (a \cdot t) \, ds = \int_S (\nabla \times a) \cdot N \, dS,$$

where the unit vector $t$ is tangent to the curve and $ds$ is an element of the curve. The unit $N$ in the above equation is locally normal to the surface.

C.9 Rotation

Tectonic rotations are observable phenomena via geologic methods such as paleomagnetism and geologic mapping of twisted rock masses. Here we consider how to treat rotations.

Suppose the Cartesian coordinates $O-123$ and $O-1'2'3'$ that have the origin in common. Their relationship is described by three angles, $\theta, \phi, \psi$ (Fig. C.5). These angles are called Euler angles$^5$.

$^5$There are several definitions on Euler angles. See reference [69].
C.9. ROTATION

Figure C.5: Cartesian coordinates with the origin in common and the Euler angles describing their relationship.

The orthogonal tensor that stands for the transformation from vector and tensor components of $O$-123 to those of $O$-1’2’3’ is

$$Q = \begin{pmatrix} 
\cos \psi & \sin \psi & 0 \\
-\sin \psi & \cos \psi & 0 \\
0 & 0 & 1 
\end{pmatrix} \cdot \begin{pmatrix} 
1 & 0 & 0 \\
0 & \cos \theta & \sin \theta \\
0 & -\sin \theta & \cos \theta 
\end{pmatrix} \cdot \begin{pmatrix} 
\cos \phi & \sin \phi & 0 \\
-\sin \phi & \cos \phi & 0 \\
0 & 0 & 1 
\end{pmatrix}$$

(C.64)

$$= \begin{pmatrix} 
\cos \psi \cos \phi - \cos \theta \sin \phi \sin \psi & \cos \psi \sin \phi + \cos \theta \cos \phi \sin \psi & \sin \psi \sin \theta \\
\sin \psi \cos \phi - \cos \theta \sin \phi \cos \psi & -\sin \psi \sin \phi + \cos \theta \cos \phi \cos \psi & \cos \psi \sin \theta \\
\sin \theta \sin \phi & \sin \theta \cos \phi & \cos \theta 
\end{pmatrix}.$$  

(C.65)

The three matrices from left to right in Eq. (C.64) represent the rotation about O3’, O1’, and O3 axes, respectively. The equation indicates that the three rotations are equal to the single rotation shown by the matrix in Eq. (C.65). If an axis of rotation is specified with $\theta$ and $\phi$, $\psi$ indicates the angle of rotation.

If a rotation matrix is given, how can we calculate the direction of the pivot and rotation angle $\alpha$ from the matrix? If we choose the third coordinate axis as being parallel to the pivot, the orthogonal matrix has the form

$$Q = \begin{pmatrix} 
\cos \alpha & \sin \alpha & 0 \\
-\sin \alpha & \cos \alpha & 0 \\
0 & 0 & 1 
\end{pmatrix}.$$  

Therefore, we have the trace of this matrix:

$$\text{trace } Q = 1 + 2 \cos \alpha.$$  

(C.66)

The trace of a matrix is an invariant, so that the value of this trace is not affected by the choice of coordinates. The rotation angle $\alpha$ is, therefore, obtained by substituting the trace of the given orthogonal tensor into Eq. (C.66). If a vector $u$ is parallel to the pivot of rotation, we have $u = I \cdot u = Q \cdot u$, indicating that Eq. (C.23) holds. Accordingly, the pivot is determined as an eigenvector of $Q$. The vector corresponds to the only real eigenvalue of the non-symmetric tensor $Q$. 
Rotation is also represented by the antisymmetric tensor
\[
\mathbf{W} = \begin{pmatrix}
0 & -W_3 & W_2 \\
W_3 & 0 & -W_1 \\
-W_2 & W_1 & 0
\end{pmatrix}.
\] (C.67)

Let us define a vector by these matrix components as \( \mathbf{W} = \sum_i W_i e^{(i)} \), then we have
\[
W_{ij} = -\sum_k \epsilon_{ijk} W_k.
\] (C.68)

Contrary, we obtain the components of the vector from the matrix by the equation
\[
W_i = -\frac{1}{2} \sum_{j,k} \epsilon_{ijk} W_{jk}.
\] (C.69)

Such vectors that correspond to antisymmetric tensors are called *axial vectors*. In the case that the axial vector \( \mathbf{W} \) represents angular velocity, \( \mathbf{W} \times \mathbf{a} \) indicates the velocity of the end point of \( \mathbf{a} \) whose initial point is fixed on the axis of rotation (Fig. 2.3). Accordingly, \( \mathbf{W} \) should be related to the rotation. For an arbitrary vector \( \mathbf{a} = \sum_i a_i e^{(i)} \), we have \( \mathbf{Q} \cdot \mathbf{a} = -e \cdot \mathbf{W} \cdot \mathbf{a} \) from Eq. (C.68):
\[
W_{ij} = -\sum_k \epsilon_{ijk} W_k.
\]

However, \( j \) and \( k \) are dummy indices. Replacing them by \( k \) and \( j \), respectively, we have
\[
\sum_j W_{ij} a_j = -\sum_{k,j} \epsilon_{ikj} W_{jk} a_k = \sum_{j,k} \epsilon_{ijk} W_{jk} a_k.
\]

This indicates that \( \mathbf{W} \cdot \mathbf{a} = \mathbf{W} \times \mathbf{a} \). The angle of rotation during a time interval \( t \) is
\[
\psi = |\mathbf{W}| t.
\] (C.70)

Then a question arises: how is the antisymmetric tensor \( \mathbf{W} \) related to the orthogonal tensor \( \mathbf{R} \)? Suppose that the terminal point of a vector \( \mathbf{a} \) is rotated to that of \( \mathbf{b} \) during the time interval, and the rotation is represented by \( \mathbf{R} \), namely
\[
\mathbf{b} = \mathbf{R} \cdot \mathbf{a}.
\] (C.71)

A broken line approximates the circular orbit of the point (Fig. 2.3). The arc length of one segment is \( |\mathbf{W}| dt \). Let \( t \) be equal to \( n \cdot dt \), where \( n \) is a natural number. Taking the limit \( n \to \infty \), the broken line approaches the circular orbit. Rotation during the interval \( dt \) is written as
\[
\mathbf{a}' = \mathbf{a} + (\mathbf{W} \times \mathbf{a}) dt = \mathbf{a} + (\mathbf{W} \cdot \mathbf{a}) dt = (I + \mathbf{W} dt) \cdot \mathbf{a}.
\]
\[
\mathbf{a}'' = (I + \mathbf{W} dt) \cdot \mathbf{a}' = (I + \mathbf{W} dt) \cdot (I + \mathbf{W} dt) \cdot \mathbf{a}.
\]

The above two equations indicate that
\[
\mathbf{b} = \lim_{n \to \infty} \left( I + \frac{t}{n} \mathbf{W} \right)^n \cdot \mathbf{a}.
\]
Comparing this and Eq. (C.71), we have

\[
R = \lim_{n \to \infty} \left( I + \frac{t}{n} W \right)^n.
\]

Taylor expansion of the \( n \)th power is

\[
\left( I + \frac{t}{n} W \right)^n = I + tW + \frac{n(n-1)}{2!} \left( \frac{t}{n} W \right)^2 + \frac{n(n-1)(n-2)}{3!} \left( \frac{t}{n} W \right)^3 + \cdots
\]

Taking the limit \( n \to \infty \), the right-hand side is simplified, and we have

\[
R = I + tW + \frac{1}{2} (tW)^2 + \frac{1}{3!} (tW)^3 + \cdots = \exp(tW)
\]

(C.72)

Recalling the Cayley-Hamilton theorem (Eq. (C.46)), the higher-order terms in Eq. (C.72) are replaced by lower terms, and we obtain

\[
R = I + \left( \sin \frac{\psi}{\psi} \right) tW + \frac{1}{2} \left[ \sin \left( \frac{\psi}{\psi} \right) \right] (tW)^2,
\]

(C.73)

where \( \psi \) is defined by Eq. (C.70). This is called Rodrigues’s equation. This is also written as

\[
R = I + \sin \alpha \begin{pmatrix} 0 & -\sigma_3 & \sigma_2 \\ \sigma_3 & 0 & -\sigma_1 \\ -\sigma_2 & \sigma_1 & 0 \end{pmatrix} + (1 - \cos \alpha) \begin{pmatrix} 0 & -\sigma_3 & \sigma_2 \\ \sigma_3 & 0 & -\sigma_1 \\ -\sigma_2 & \sigma_1 & 0 \end{pmatrix}^2,
\]

(C.74)

when an angle of rotation \( \alpha \) about the fixed axis denoted by the unit vector \( \sigma = (\sigma_1, \sigma_2, \sigma_3)^T \) is given\(^6\).

C.10 Exercise

B.1 Fault-slip data are represented by two unit vectors, one for the perpendicular to fault plane, and the other for slip direction. Let the unit vector \( u = (u_1, u_2, u_3)^T \) point to the slip direction of the hangingwall, and let \( v = (v_1, v_2, v_3)^T \) be the unit vector normal to the fault plane. If the plane is horizontal, the vector points upward. The vector components are evaluated in the Cartesian coordinates \( O-123 \). The third unit vector \( w \) is defined to be normal to the two vectors and to form a right-hand system in the order of \( u, v, \) and \( w \). Show the orthogonal tensor that stands for the coordinate transformation from \( O-123 \) to \( O-uvw \).

\(^6\)See Exercise 2.5 in p. 56 and its answer in p. 355.