Rotation in the Space*

(Com S 477/577 Notes)

Yan-Bin Jia

Sep 6, 2016

The position of a point after some rotation about the origin can simply be obtained by multiplying its coordinates with a matrix. One reason for introducing homogeneous coordinates is to be able to describe translation with a matrix so that multiple transformations, whether each is a rotation or a translation, can be concatenated into one described by the product of their respective matrices. However, in some applications (such as spaceship tracking), we need only be concerned with rotations of an object, or at least independently from other transformations. In such a situation, we often need to extract the rotation axis and angle from a matrix which represents the concatenation of multiple rotations. The homogeneous transformation matrix, however, is not well suited for the purpose.

1 Euler Angles

A rigid body in the space has a coordinate frame attached to itself and located often at the center of mass. This frame is referred to as the *body frame* or *local frame*. The position, orientation, and motion of the body can be described using the body frame relative to a fixed reference frame, called the *world frame*.

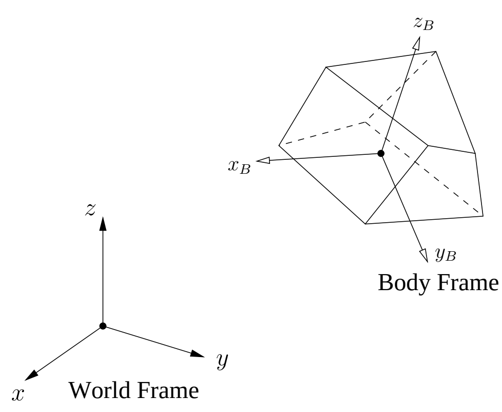

Figure 1: Body frame (local) vs. world frame (global).

The rigid body has six degrees of freedom: its position given by the x, y, z coordinates of its center of mass (i.e., the origin of the local frame) in the world frame, and its orientation described

*The material is partially based on Chapters 3–6 of the book [2]

Figure 2: Roll, pitch, yaw of an aircraft.

by three angles of rotation of its body frame from the world frame. To describe this rotation, one choice is to use three angle about the x-, y-, and z-axes. A more convenient choice, from the perspective of the rotating body, is to use three angles about the axes of the body frame. The latter three angles are called *Euler angles*.

For example, an aircraft in flight can perform three independent rotations: *roll*, about an axis from nose to tail; *pitch*, nose up or down about an axis from wing to wing; and *yaw*, nose left or right about an axis from top to bottom.

There are several conventions for Euler angles, depending on the axes about which the rotations are carried out. Here we introduce the Z-Y-X Euler angles. The body frame starts in the same orientation as the world frame. To achieve its final orientation, the first rotation is by an angle ϕ, about the body z-axis, the second rotation by an angle $\theta \in [0, \pi]$ about the body y-axis, and the third rotation by an angle ψ about the body x-axis. Here, ϕ, θ, and ψ correspond to as yaw, pitch, and roll, respectively. The three rotations are illustrated in Figure 3, where x'_B-y'_B-z'_B and x''_B-y''_B-z''_B are the two intermediate configurations of the body frame x_B-y_B-z_B.

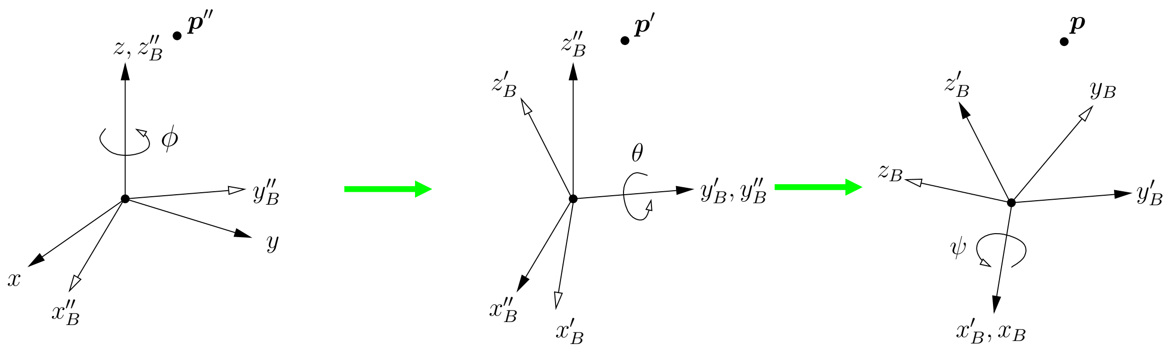

Figure 3: Z-Y-X Euler angles ϕ (yaw), θ (pitch), ψ (roll). Here, $\boldsymbol{p}$, $\boldsymbol{p}'$, and $\boldsymbol{p}''$ are the same point of coordinates in the frames x_B-y_B-z_B, x'_B-y'_B-z'_B, and x''_B-y''_B-z''_B, respectively.

Now we derive the homogeneous transformation matrix. Consider a point $\boldsymbol{p}$ in the body frame $x_{\mathcal{B}}$-$y_{\mathcal{B}}$-$z_{\mathcal{B}}$. Let us determine its coordinates in the original frame x-y-z by considering the four frames backward: $x_{\mathcal{B}}$-$y_{\mathcal{B}}$-$z_{\mathcal{B}}$, x'_B-y'_B-z'_B, x''_B-y''_B-z''_B, and x-y-z. Consider a point $\boldsymbol{p}$ with its coordinates given in the body frame. Since the frame $x_{\mathcal{B}}$-$y_{\mathcal{B}}$-$z_{\mathcal{B}}$ is obtained from the frame x'_B-y'_B-z'_B by a rotation about the x'_B-axis through an angle ψ, the coordinates of $\boldsymbol{p}$ in the latter frame is

$$\boldsymbol{p}' = \mathrm{Rot}_x(\psi)\boldsymbol{p}.$$

Next, the frame x'_B-y'_B-z'_B is obtained from the frame x''_B-y''_B-z''_B after a rotation about the y''_B-axis through an angle θ. So the same point has coordinates

$$\boldsymbol{p}'' = \mathrm{Rot}_y(\theta)\boldsymbol{p}' = \mathrm{Rot}_y(\theta)(\mathrm{Rot}_x(\psi)\boldsymbol{p}$$

in the frame x''_B-y''_B-z''_B. Similarly, the coordinates of the point in the frame x-y-z is

$$\mathrm{Rot}_z(\phi)\boldsymbol{p}'' = \mathrm{Rot}_z(\phi)(\mathrm{Rot}_y(\theta)(\mathrm{Rot}_x(\psi)\boldsymbol{p}.$$

Thus the transformation matrix associated with the Z-Y-Z Euler angles are

$$\begin{aligned}
\mathrm{Rot}_{zyx}(\phi,\theta,\psi) &= \mathrm{Rot}_z(\phi)\cdot\mathrm{Rot}_y(\theta)\cdot\mathrm{Rot}_x(\psi) \\
&= \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\psi & -\sin\psi \\ 0 & \sin\psi & \cos\psi \end{pmatrix} \\
&= \begin{pmatrix} \cos\phi\cos\theta & \cos\phi\sin\theta\sin\psi - \sin\phi\cos\psi & \cos\phi\sin\theta\cos\psi + \sin\phi\sin\psi \\ \sin\phi\cos\theta & \sin\phi\sin\theta\sin\psi + \cos\phi\cos\psi & \sin\phi\sin\theta\cos\psi - \cos\phi\sin\psi \\ -\sin\theta & \cos\theta\sin\psi & \cos\theta\cos\psi \end{pmatrix} \quad (1)
\end{aligned}$$

Euler angles are defined in terms of three moving axes. Singularities happen when two of these axes coincide. Consider the z-x-z angles α, β, γ with $\beta = 0$. Here, the first and the third rotations are about the same axis. All the Euler angles $(\alpha, 0, \beta)$ with the same $\alpha+\beta$ value thus describe one rotation. Namely, one degree of freedom is lost. Linear interpolation from one orientation to another is not well-behaved. Imagine, when the latitude and longitude values are interpolated. What will happen when longitude goes towards 90 degrees to reach the north pole? All latitude values there make no difference as they end up describing the same point! This is a phenomenon referred to as "gimbal lock", drawing its name from certain orientation with three nested moving gimbals in which two of the three axes become collinear — restricting the available rotations to only two axes. (Watch a nice video tutorial on gimbal lock at `https://www.youtube.com/watch?v=zc8b2Jo7mno`.)

The above singularity issue with Euler angles is because they form a 3D box $[0, 2\pi]^3$, while the rotations constitute a 3D projective space. The mapping between the two spaces cannot be a "continuous both way" (or, strictly speaking, a homeomorphism[1]).

2 Arbitrary Rotation Axis

Let $\boldsymbol{v}$ be a vector that is undergoing a rotation of the amount θ about some axis through the origin with an arbitrary orientation. This can be viewed as a rotation about a line that was treated in Appendix A in the notes titled "Transformations in Homogeneous Coordinates". The resulting vector can be found by first rotating the axis to the z-axis, performing the rotation, and rotating the axis back to its original orientation. Now let us present a simpler approach with direct geometric meaning.

[1]a bijective and continuous function whose inverse is also continuous

Let $\hat{\boldsymbol{k}} = (k_x, k_y, k_z)^T$ be the unit vector along the axis. To obtain the resulting $\boldsymbol{v}'$ from the rotation, we decompose $\boldsymbol{v}$ into two parts: one along $\hat{\boldsymbol{k}}$ and the other in the plane Π containing the origin and perpendicular to $\hat{\boldsymbol{k}}$:

$$\begin{aligned} \boldsymbol{v}_{\parallel} &= (\boldsymbol{v}\cdot\hat{\boldsymbol{k}})\hat{\boldsymbol{k}}, && (2) \\ \boldsymbol{v}_{\perp} &= \boldsymbol{v} - \boldsymbol{v}_{\parallel} \\ &= \boldsymbol{v} - (\boldsymbol{v}\cdot\hat{\boldsymbol{k}})\hat{\boldsymbol{k}}. && (3) \end{aligned}$$

As shown in Figure 4, the component $\boldsymbol{v}_{\parallel}$ is not affected by rotation. We need only determine the vector $\boldsymbol{v}'_{\perp}$ that results from applying the same rotation to $\boldsymbol{v}_{\perp}$.

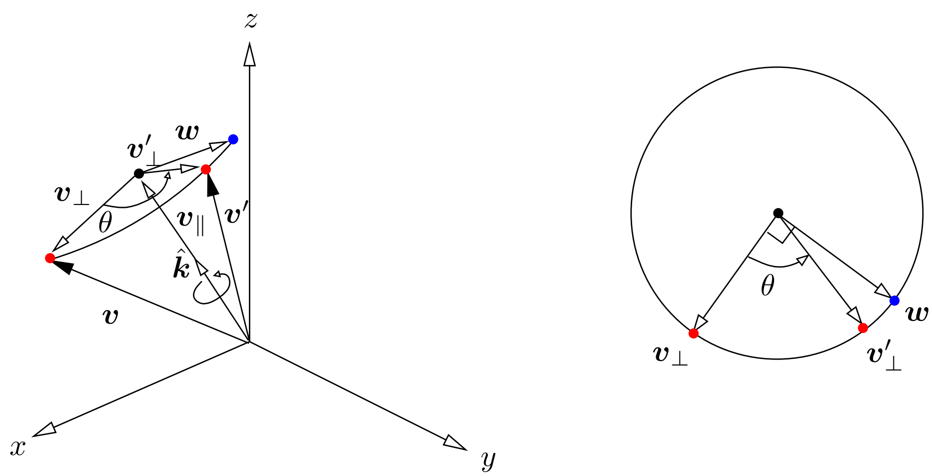

Figure 4: The vector $\boldsymbol{v}$ is rotated to $\boldsymbol{v}'$ about the axis $\hat{\boldsymbol{k}}$.

The plane Π is spanned by two orthogonal vectors: $\boldsymbol{v}_{\perp}$ and

$$\boldsymbol{w} = \hat{\boldsymbol{k}} \times \boldsymbol{v}_{\perp} = \hat{\boldsymbol{k}} \times \boldsymbol{v}. \tag{4}$$

It is easy to see that $\boldsymbol{v}_{\perp}, \boldsymbol{w}, \hat{\boldsymbol{k}}$ are the axes of a right-handed system. Since $\hat{\boldsymbol{k}}$ is orthogonal to $\boldsymbol{v}_{\perp}$, $\boldsymbol{v}_{\perp}$ and $\boldsymbol{w}$ have the same length. The vector $\boldsymbol{v}'_{\perp}$ lies in the plane Π, and has the form

$$\boldsymbol{v}'_{\perp} = \boldsymbol{v}_{\perp}\cos\theta + \boldsymbol{w}\sin\theta. \tag{5}$$

Finally, we have the rotated vector from $\boldsymbol{v}$:

$$\begin{aligned} \boldsymbol{v}' &= \boldsymbol{v}_{\parallel} + \boldsymbol{v}'_{\perp} \\ &= (\boldsymbol{v}\cdot\hat{\boldsymbol{k}})\hat{\boldsymbol{k}} + \boldsymbol{v}_{\perp}\cos\theta + \boldsymbol{w}\sin\theta && \text{(by (2) and (5))} \\ &= (\boldsymbol{v}\cdot\hat{\boldsymbol{k}})\hat{\boldsymbol{k}} + (\boldsymbol{v} - (\boldsymbol{v}\cdot\hat{\boldsymbol{k}})\hat{\boldsymbol{k}})\cos\theta + (\hat{\boldsymbol{k}}\times\boldsymbol{v})\sin\theta && \text{(by (3) and (4))} \\ &= \boldsymbol{v}\cos\theta + (\hat{\boldsymbol{k}}\times\boldsymbol{v})\sin\theta + \hat{\boldsymbol{k}}(\hat{\boldsymbol{k}}\cdot\boldsymbol{v})(1-\cos\theta). && (6) \end{aligned}$$

Equation (6) is called *Rodrigues' rotation formula*.

Note that

$$
\begin{aligned}
\boldsymbol{v}\cos\theta &= \begin{pmatrix} \cos\theta & 0 & 0 \\ 0 & \cos\theta & 0 \\ 0 & 0 & \cos\theta \end{pmatrix} \boldsymbol{v}; \\
\hat{\boldsymbol{k}} \times \boldsymbol{v} &= \begin{pmatrix} 0 & -k_z & k_y \\ k_z & 0 & -k_x \\ -k_y & k_x & 0 \end{pmatrix} \boldsymbol{v}; \\
\hat{\boldsymbol{k}}(\hat{\boldsymbol{k}} \cdot \boldsymbol{v}) &= (\hat{\boldsymbol{k}}\hat{\boldsymbol{k}}^T)\boldsymbol{v} \\
&= \begin{pmatrix} k_x^2 & k_x k_y & k_x k_z \\ k_y k_x & k_y^2 & k_y k_z \\ k_z k_x & k_z k_y & k_z^2 \end{pmatrix} \boldsymbol{v}.
\end{aligned}
$$

Substituting the above into (6), we express $\boldsymbol{v}'$ as the product of the following 3×3 rotation matrix with $\boldsymbol{v}$:

$$
\mathrm{Rot}_{\hat{\boldsymbol{k}}}(\theta) = \begin{pmatrix} k_x k_x(1-\cos\theta)+\cos\theta & k_x k_y(1-\cos\theta)-k_z\sin\theta & k_x k_z(1-\cos\theta)+k_y\sin\theta \\ k_x k_y(1-\cos\theta)+k_z\sin\theta & k_y k_y(1-\cos\theta)+\cos\theta & k_y k_z(1-\cos\theta)-k_x\sin\theta \\ k_x k_z(1-\cos\theta)-k_y\sin\theta & k_y k_z(1-\cos\theta)+k_x\sin\theta & k_z k_z(1-\cos\theta)+\cos\theta \end{pmatrix}. \tag{7}
$$

3 Rotation Matrix

We have seen the use of a matrix to represent a rotation. Such a matrix is referred to as a *rotation matrix*. In this section we look at the properties of rotation matrix. Below let us first review some concepts from linear algebra.

3.1 Eigenvalues

An $n \times n$ matrix A is *orthogonal* if its columns are unit vectors and orthogonal to each other, namely, if $A^T A = I_n$. An orthogonal matrix has determinant $\det(A) = \pm 1$.

A complex number λ is called the *eigenvalue* of A if there exists a vector $\boldsymbol{x} \neq 0$ such that $A\boldsymbol{x} = \lambda\boldsymbol{x}$. It tells whether the vector $\boldsymbol{x}$ is stretched or shrunk or reversed or left unchanged — when it is multiplied by A. The vector $\boldsymbol{x}$ is called an *eigenvector* of A associated with the eigenvalue λ. The set of all eigenvalues is called the *spectrum* of A.

The matrix A has exactly n eigenvalues (multiplicities included). They are the roots of the nth degree polynomial $\det(A - \lambda I)$, called the *characteristic polynomial*. Let $\lambda_1, \dots, \lambda_k$ be the distinct eigenvalues of this polynomial, then

$$
\det(A - \lambda I) = (-1)^n (\lambda - \lambda_1)^{\delta_1} \cdots (\lambda - \lambda_k)^{\delta_k}, \qquad \text{for some } \delta_1, \dots, \delta_k > 0.
$$

Here δ_i, $1 \leq i \leq k$, is the *algebraic multiplicity* of λ_i.

The *geometric multiplicity* of λ_i is $n - \mathrm{rank}(A - \lambda_i I)$. It specifies the maximum number of linearly independent eigenvectors associated with λ_i.

Example 1. The diagonal matrix aI, where I is the $n \times n$ identity matrix, has the characteristic polynomial $(a - \lambda)^n$. So a is the only eigenvalue while every vector $\boldsymbol{x} \in C^n$, $\boldsymbol{x} \neq 0$ is an eigenvector. Both algebraic and geometric multiplicities of a are n.

Example 2. The nth order matrix

$$A = \begin{pmatrix} a & 1 & 0 & \cdots & 0 \\ 0 & a & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & 1 \\ 0 & 0 & 0 & \cdots & a \end{pmatrix}$$

also has the characteristic polynomial $(a-\lambda)^n$ and a as its only eigenvalue. The algebraic multiplicity of a is n. However, the rank of $A - aI$ is $n-1$; hence the geometric multiplicity of a is $n-(n-1)=1$.

The product of the n eigenvalues equals the determinant of A. The sum of the n eigenvalues equals the sum of the n diagonal entries; this number is called the *trace* of A and denoted $\mathrm{Tr}(A)$.

The matrix A is *invertible* if and only if it does not have a zero eigenvalue. A symmetric matrix has real eigenvalues. A symmetric and positive definite matrix has positive eigenvalues.

3.2 Matrices Describing Rotations

All rotation matrices applying to points in the space are 3×3 .

Theorem 1 *A 3×3 matrix describes a rotation about some axis through the origin if and only if it is orthogonal and has determinant 1.*

Proof ($\Rightarrow$) We first show that any rotation matrix A is orthogonal and $\det(A) = 1$. Let $\boldsymbol{v}$ be the corresponding rotation axis and θ the angle of rotation about the axis. The inner product of two vectors, say, $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$, must be preserved under the rotation; namely,

$$(A\boldsymbol{v}_1)^T(A\boldsymbol{v}_2) = \boldsymbol{v}_1^T\boldsymbol{v}_2.$$

The above implies that

$$\boldsymbol{v}_1^T(A^TA - I)\boldsymbol{v}_2 = 0,$$

where I is the 3×3 identity matrix. Because $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are arbitrarily chosen, it must hold that

$$A^TA - I = 0 \qquad \text{and thus} \qquad A^TA = I.$$

Therefore, A is orthogonal. Further more, from the equation

$$\det(A^TA) = \det(A)^T\det(A) = \Big(\det(A)\Big)^2 = 1$$

we infer that $\det(A)$ must be 1 or -1. But a rotation preserves the right-handedness of a coordinate frame, so we cannot have $\det(A) = -1$. For example, the diagonal matrix with diagonal entries 1, 1, and -1 would describe a reflection about the x-y plane instead of a rotation. We have thus shown that $\det(A) = 1$.

($\Leftarrow$) Suppose A is a 3×3 orthogonal matrix with determinant 1. We want to show that it represents a rotation. First, we show that there exists a vector $\boldsymbol{u}$ such that $A\boldsymbol{u} = \boldsymbol{u}$. This vector will be along the rotation axis. It suffices to establish that the matrix A has an eigenvalue of 1, which is true if and only if $\det(A - I) = 0$. We have

$$\begin{aligned} \det(A - I) &= \det(A^T)\det(A - I) \\ &= \det(A^T A - A^T) \\ &= \det(I - A^T) \qquad\qquad (A \text{ orthogonal}) \\ &= \det((I - A)^T) \\ &= \det(I - A) \\ &= -\det(A - I). \end{aligned}$$

Thus $\det(A - I) = 0$. Next, we show that the plane containing the origin and orthogonal to $\boldsymbol{u}$ maps to itself under the rotation. Consider an arbitrary vector $\boldsymbol{v}$ in this plane, i.e., $\boldsymbol{v}^T\boldsymbol{u} = 0$. The vector $A\boldsymbol{v}$ is orthogonal to $\boldsymbol{u}$ because

$$\begin{aligned} (A\boldsymbol{v})^T\boldsymbol{u} &= \boldsymbol{v}^T A^T A\boldsymbol{u} \\ &= \boldsymbol{v}^T I\boldsymbol{u} \\ &= \boldsymbol{v}^T\boldsymbol{u} \\ &= 0. \end{aligned}$$

Finally, the inner product of two vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ is invariant under the transformation because A is orthogonal. Therefore, the lengths of $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ and the angles between them are preserved under A. Hence the transformation represented by A is a rotation. □

Note that a rotation about an axis not through the origin cannot be represented by a 3×3 orthogonal matrix with determinant 1. To see this, let $\boldsymbol{b} \neq 0$ be a point on the rotation axis. Then a point $\boldsymbol{x}$ is rotated into $R(\boldsymbol{x} - \boldsymbol{b}) + \boldsymbol{b}$, where R is a rotation matrix (orthogonal and having determinant 1). However, there does not exist a rotation matrix R' such that $R'\boldsymbol{x} = R(\boldsymbol{x} - \boldsymbol{b}) + \boldsymbol{b}$ unless $R = I$.

Two successive rotations about the same axis in $\mathbb{R}^3$ is equivalent to one rotation about this axis through an angle which is the sum of the two rotation angles. Even if they are not about the same axis through the origin, their composition is still a rotation through some angle about some axis. The reason is that the product of their rotation matrices, each orthogonal with determinant 1, is still orthogonal with determinant 1. Thus, by Theorem 1 the product matrix describes a rotation. To state this formally, *the composition of any two rotations is equivalent to a rotation.*

Given a rotation matrix, we can extract the Euler angles ϕ, θ, and ψ based on the form (1).

3.3 Recovery of Rotation Axis and Angle

Consider the following orthogonal matrix with determinant 1:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Theorem 1 says that it represents a rotation about some axis $\boldsymbol{v}$ in the space. How to find this axis of rotation? It follows from the proof of Theorem 1 that $\boldsymbol{v}$ must be an eigenvector of A which corresponds to the eigenvalue 1. Namely,

$$(A - I)\boldsymbol{v} = 0,$$

from which we can obtain a solution vector by setting one of the non-zero components in $\boldsymbol{v}$ to 1.

We makes use of $Q = A - A^T$. Note that $\boldsymbol{v}$ is also an eigenvector of A^T for

$$\begin{aligned} A^T\boldsymbol{v} &= A^T A\boldsymbol{v} \\ &= I\boldsymbol{v} \\ &= \boldsymbol{v}. \end{aligned}$$

We have

$$\begin{aligned} (A - A^T)\boldsymbol{v} &= Q\boldsymbol{v} \\ &= \begin{pmatrix} 0 & q_3 & -q_2 \\ -q_3 & 0 & q_1 \\ q_2 & -q_1 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \end{aligned} \tag{8}$$

where

$$\begin{aligned} q_1 &= a_{23} - a_{32}, \\ q_2 &= a_{31} - a_{13}, \\ q_3 &= a_{12} - a_{21}. \end{aligned}$$

The matrix $A - A^T$ in (8) is not of full rank because $\det(A - A^T) = 0$. So the three linear equations are not independent. First, we assume v_3 to be non-zero and let $v_3 = 1$, obtaining

$$\begin{aligned} q_3 v_2 - q_2 &= 0, \\ -q_3 v_1 + q_1 &= 0. \end{aligned}$$

Assume $q_3 \neq 0$ first. Solving the two equations above gives us the axis of rotation

$$\boldsymbol{v} = \left(\frac{q_1}{q_3}, \frac{q_2}{q_3}, 1 \right)^T.$$

Equivalently, we have $\boldsymbol{v} = (q_1, q_2, q_3)^T$ (which also works for $q_3 = 0$, i.e., $v_3 = 0$). Thus, if a rotation is defined by the 3×3 matrix $A = (a_{ij})$, then the axis of rotation is

$$\boldsymbol{v} = \begin{pmatrix} a_{23} - a_{32} \\ a_{31} - a_{13} \\ a_{12} - a_{21} \end{pmatrix}. \tag{9}$$

Having determined the rotation axis $\boldsymbol{v}$, we look at how to recover the rotation angle. If the rotation happens to be about the z-axis through an angle θ, then we have the rotation matrix

$$A = \mathrm{Rot}_z(\theta) \equiv \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

And the rotation angle θ satisfies $\cos\theta = a_{11}$. We can also utilize the trace

$$\mathrm{Tr}(A) = a_{11} + a_{22} + a_{33} = 1 + 2\cos\theta$$

and obtain

$$\cos\theta = \frac{\mathrm{Tr}(A) - 1}{2}, \tag{10}$$

$$\begin{aligned} \sin\theta &= \pm\sqrt{1 - \sin^2\cos\theta} \\ &= \pm\sqrt{\frac{3 - \mathrm{Tr}(A)^2 + 2\mathrm{Tr}(A)}{4}}. \end{aligned} \tag{11}$$

The sign of $\sin\theta$ must agree with that of a_{21}.

In fact, the solution of θ from (10)–(11) generalizes to the situation with an arbitrary rotation axis. To see this, we use the old trick of applying a sequence of rotations, represented by a matrix Q, to transform the axis $\boldsymbol{v}$ into the z-axis. Then we carry out the rotation about the z-axis through the angle θ. Finally, we apply a sequence of reverse rotations Q^T to transform the z-axis back to $\boldsymbol{v}$. The rotation matrix is represented as

$$A = Q^T \mathrm{Rot}_z(\theta) Q. \tag{12}$$

The trace function is commutative, that is, $\mathrm{Tr}(CD) = \mathrm{Tr}(DC)$ for any two square matrices of the same dimensions.[2] This property allows us to derive the following from (12):

$$\begin{aligned} \mathrm{Tr}(A) &= \mathrm{Tr}\Big(Q^T \mathrm{Rot}_z(\theta) Q\Big) \\ &= \mathrm{Tr}\Big(\Big(Q^T \mathrm{Rot}_z(\theta)\Big) Q\Big) \\ &= \mathrm{Tr}\Big(Q\Big(Q^T \mathrm{Rot}_z(\theta)\Big)\Big) \\ &= \mathrm{Tr}\Big(\Big(QQ^T\Big) \mathrm{Rot}_z(\theta)\Big) \\ &= \mathrm{Tr}\Big(I \cdot \mathrm{Rot}_z(\theta)\Big) \\ &= \mathrm{Tr}\Big(\mathrm{Rot}_z(\theta)\Big) \\ &= 1 + 2\cos\theta. \end{aligned}$$

Hence the rotation angle θ still satisfies equation (10) when the rotation axis is arbitrary. How do we determine the sign in (11)? First, we normalize the rotation axis $\boldsymbol{v}$ from (9) and obtain a unit vector $\hat{\boldsymbol{v}} = (v_x, v_y, v_z)$. The rotation about $\hat{\boldsymbol{v}}$ through an angle θ is described by the following rotation matrix according to (7):

$$\mathrm{Rot}_{\hat{\boldsymbol{v}}}(\theta) = \begin{pmatrix} v_x v_x(1-\cos\theta) + \cos\theta & v_x v_y(1-\cos\theta) - v_z\sin\theta & v_x v_z(1-\cos\theta) + v_y\sin\theta \\ v_x v_y(1-\cos\theta) + v_z\sin\theta & v_y v_y(1-\cos\theta) + \cos\theta & v_y v_z(1-\cos\theta) - v_x\sin\theta \\ v_x v_z(1-\cos\theta) - v_y\sin\theta & v_y v_z(1-\cos\theta) + v_x\sin\theta & v_z v_z(1-\cos\theta) + \cos\theta \end{pmatrix}.$$

[2] The proof is directly based on the definition of trace.

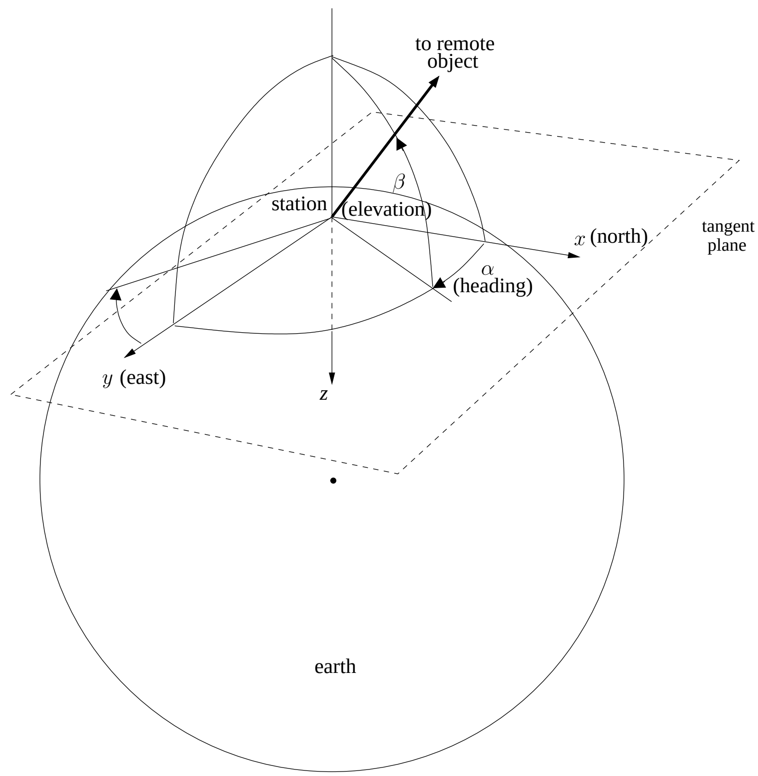

Figure 5: Tracking an aircraft from a ground station.

Comparing an off-diagonal element in the matrix A, say a_{21}, with its corresponding entry in $\mathrm{Rot}_{\hat{\boldsymbol{v}}}(\theta)$ gives us

$$
\begin{aligned}
\sin\theta &= \frac{a_{21} - v_x v_y (1 - \cos\theta)}{v_z} \\
&= \frac{2a_{21} - v_x v_y (3 - \mathrm{Tr}(A))}{2v_z}.
\end{aligned}
\tag{13}
$$

The angle $\theta \in [0, 2\pi)$ of rotation is thus completely determined from (10) and (13).

4 Application — Tracking

In this section, we describe an application of rotations in $\mathbb{R}^3$. Consider a remote object, such as an aircraft, which is being tracked from a station on the earth. We define a local coordinate system at the station with the xy-plane being the tangent plane to the earth. The x- and y-axes point in the directions of the North and East, respectively, and the z-axis points toward the center of the earth. The coordinate system is drawn in the figure below.

The *heading* α is the angle in the tangent plane between the North and the projection of the direction to the remote object being tracked. The *elevation* is the angle between the tangent plane and the direction to the object. The tracking transformation is a rotation about the z-axis through the angle α, followed by a rotation about the new y-axis through the angle β. After these two rotations, the x-axis is pointing toward the object.

Let R be the 3×3 matrix to represent the composite rotation; it relates the tracking coordinate system to the station's coordinate system. Namely, a point $\boldsymbol{p} = (p_x, p_y, p_z)^T$ in the (re-oriented) tracking frame is $R\boldsymbol{p}$ in the station's frame. Then we have

$$\begin{aligned} R &= \mathrm{Rot}_z(\alpha) \cdot \mathrm{Rot}_y(\beta) \\ &= \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix} \\ &= \begin{pmatrix} \cos\alpha\cos\beta & -\sin\alpha & \cos\alpha\sin\beta \\ \sin\alpha\cos\beta & \cos\alpha & \sin\alpha\sin\beta \\ -\sin\beta & 0 & \cos\beta \end{pmatrix}. \end{aligned}$$

This composite tracking transformation can be represented as an equivalent rotation through some angle θ about some axis $\boldsymbol{v}$. By (9) we know that the rotation axis is given by

$$\boldsymbol{v} = \begin{pmatrix} \sin\alpha\sin\beta \\ -\cos\alpha\sin\beta - \sin\beta \\ -\sin\alpha\cos\beta - \sin\alpha \end{pmatrix} \equiv \begin{pmatrix} \sin\frac{\alpha}{2}\sin\frac{\beta}{2} \\ -\cos\frac{\alpha}{2}\sin\frac{\beta}{2} \\ -\sin\frac{\alpha}{2}\cos\frac{\beta}{2} \end{pmatrix}. \tag{14}$$

Let $\hat{\boldsymbol{v}} = (v_x, v_y, v_z)$ be the normalization of $\boldsymbol{v}$. By (10) and (13) the angle of rotation satisfies the following equations:

$$\begin{aligned} \cos\theta &= \frac{\cos\alpha\cos\beta + \cos\alpha + \cos\beta - 1}{2}, \\ \sin\theta &= \frac{2\sin\alpha\cos\beta - v_x v_y (3 - \mathrm{Tr}(A))}{2v_z}. \end{aligned}$$

The domain for the tracking angles is $\alpha \in (-\pi, \pi]$ and $\beta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$. As β increases from $\frac{\pi}{2} - \epsilon$ to $\frac{\pi}{2} + \epsilon$, for small ϵ, α instantaneously jumps to $\alpha + \pi$. For this reason, $(\alpha, \frac{\pi}{2})$ is a singular point, and the phenomenon is gimbal lock.

References

[1] J. Craig. *Introduction to Robotics: Mechanics and Control.* 2nd ed., Addison-Wesley, 1989.

[2] J. B. Kuipers. *Quaternions and Rotation Sequences.* Princeton University Press, 1999.

[3] `http://sundaram.wordpress.com/2013/03/08/mathematical-reason-behind-gimbal-lock-in-euler-angles/`.