Chapter 2
Representation of a Three-Dimensional Moving Scene

I will not define time, space, place and motion, as being well known to all.

– Isaac Newton, *Principia Mathematica, 1687*

The study of the geometric relationship between a three-dimensional (3-D) scene and its two-dimensional (2-D) images taken from a moving camera is at heart the interplay between two fundamental sets of transformations: *Euclidean motion*, also called *rigid-body motion*, which models how the camera moves, and *perspective projection*, which describes the image formation process. Long before these two transformations were brought together in computer vision, their theory had been developed independently. The study of the principles of motion of a material body has a long history belonging to the foundations of mechanics. For our purpose, more recent noteworthy insights to the understanding of the motion of rigid objects came from Chasles and Poinsot in the early 1800s. Their findings led to the current treatment of this subject, which has since been widely adopted.

In this chapter, we will start with an introduction to three-dimensional Euclidean space as well as to rigid-body motions. The next chapter will then focus on the perspective projection model of the camera. Both chapters require familiarity with some basic notions from linear algebra, many of which are reviewed in Appendix A at the end of this book.

2.1 Three-dimensional Euclidean space

We will use $\mathbb{E}^3$ to denote the familiar three-dimensional Euclidean space. In general, a Euclidean space is a set whose elements satisfy the five axioms of Euclid. Analytically, three-dimensional Euclidean space can be represented globally by a Cartesian coordinate frame: every point $p \in \mathbb{E}^3$ can be identified with a point in $\mathbb{R}^3$ with three coordinates

$$\boldsymbol{X} \doteq [X_1, X_2, X_3]^T = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \quad \in \mathbb{R}^3.$$

Sometimes, we may also use $[X, Y, Z]^T$ to indicate individual coordinates instead of $[X_1, X_2, X_3]^T$. Through such an assignment of a Cartesian frame, one establishes a one-to-one correspondence between $\mathbb{E}^3$ and $\mathbb{R}^3$, which allows us to safely talk about points and their coordinates as if they were the same thing.

Cartesian coordinates are the first step toward making it possible measure distances and angles. In order to do so, $\mathbb{E}^3$ must be endowed with a *metric*. A precise definition of metric relies on the notion of *vector*.

Definition 2.1 (Vector). *In Euclidean space, a* vector v *is determined by a pair of points* $p, q \in \mathbb{E}^3$ *and is defined as a directed arrow connecting* p *to* q*, denoted* $v = \overrightarrow{pq}$.

The point p is usually called the base point of the vector v. In coordinates, the vector v is represented by the triplet $[v_1, v_2, v_3]^T \in \mathbb{R}^3$, where each coordinate is the difference between the corresponding coordinates of the two points: if p has coordinates $\boldsymbol{X}$ and q has coordinates $\boldsymbol{Y}$, then v has coordinates[1]

$$v \doteq \boldsymbol{Y} - \boldsymbol{X} \quad \in \mathbb{R}^3.$$

The preceding definition of a vector is referred to as a *bound vector*. One can also introduce the concept of a *free vector*, a vector whose definition does not depend on its base point. If we have two pairs of points (p, q) and (p', q') with coordinates satisfying $\boldsymbol{Y} - \boldsymbol{X} = \boldsymbol{Y}' - \boldsymbol{X}'$, we say that they define the same free vector. Intuitively, this allows a vector v to be transported in parallel anywhere in $\mathbb{E}^3$. In particular, without loss of generality, one can assume that the base point is the origin of the Cartesian frame, so that $\boldsymbol{X} = 0$ and $\boldsymbol{Y} = v$. Note, however, that this notation is confusing: $\boldsymbol{Y}$ here denotes the coordinates of a vector that happen to be the same as the coordinates of the point q just because we have chosen the point p to be the origin. The reader should keep in mind that points and vectors are different geometric objects. This will be important, as we will see shortly, since a rigid-body motion acts differently on points and vectors.

[1]Note that we use the same symbol v for a vector and its coordinates.

The set of all free vectors forms a *linear vector space*[2] (Appendix A), with the linear combination of two vectors $v, u \in \mathbb{R}^3$ defined by

$$\alpha v + \beta u = [\alpha v_1 + \beta u_1, \alpha v_2 + \beta u_2, \alpha v_3 + \beta u_3]^T \in \mathbb{R}^3, \quad \forall \alpha, \beta \in \mathbb{R}.$$

The *Euclidean metric* for $\mathbb{E}^3$ is then defined simply by an inner product[3] (Appendix A) on the vector space $\mathbb{R}^3$. It can be shown that by a proper choice of Cartesian frame, any inner product in $\mathbb{E}^3$ can be converted to the following canonical form

$$\langle u, v \rangle \doteq u^T v = u_1 v_1 + u_2 v_2 + u_3 v_3, \quad \forall u, v \in \mathbb{R}^3. \tag{2.1}$$

This inner product is also referred to as the standard Euclidean metric. In most parts of this book (but not everywhere!) we will use the canonical inner product $\langle u, v \rangle = u^T v$. Consequently, the norm (or length) of a vector v is $\|v\| \doteq \sqrt{\langle v, v \rangle} = \sqrt{v_1^2 + v_2^2 + v_3^2}$. When the inner product between two vectors is zero, i.e. $\langle u, v \rangle = 0$, they are said to be *orthogonal*.

Finally, Euclidean space $\mathbb{E}^3$ can be formally described as a space that, with respect to a Cartesian frame, can be identified with $\mathbb{R}^3$ and has a metric (on its vector space) given by the above inner product. With such a metric, one can measure not only distances between points or angles between vectors, but also calculate the length of a curve[4] or the volume of a region.

While the inner product of two vectors is a real scalar, the so-called *cross product* of two vectors is a vector as defined below.

Definition 2.2 (Cross product). *Given two vectors $u, v \in \mathbb{R}^3$, their* cross product *is a third vector with coordinates given by*

$$u \times v \doteq \begin{bmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{bmatrix} \in \mathbb{R}^3.$$

It is immediate from this definition that the cross product of two vectors is linear in each of its arguments: $u \times (\alpha v + \beta w) = \alpha u \times v + \beta u \times w$, $\forall \alpha, \beta \in \mathbb{R}$. Furthermore, it is immediate to verify that

$$\langle u \times v, u \rangle = \langle u \times v, v \rangle = 0, \quad u \times v = -v \times u.$$

Therefore, the cross product of two vectors is orthogonal to each of its factors, and the order of the factors defines an *orientation* (if we change the order of the factors, the cross product changes sign).

---

[2] Note that the set of points does not.

[3] In some literature, the inner product is also referred to as the "dot product."

[4] If the trajectory of a moving particle p in $\mathbb{E}^3$ is described by a curve $\gamma(\cdot) : t \mapsto \boldsymbol{X}(t) \in \mathbb{R}^3, t \in [0,1]$, then the total length of the curve is given by

$$l(\gamma(\cdot)) = \int_0^1 \|\dot{\boldsymbol{X}}(t)\| \, dt,$$

where $\dot{\boldsymbol{X}}(t) = \frac{d}{dt}(\boldsymbol{X}(t)) \in \mathbb{R}^3$ is the so-called tangent vector to the curve.

If we fix u, the cross product can be represented by a map from $\mathbb{R}^3$ to $\mathbb{R}^3$: $v \mapsto u \times v$. This map is linear in v and therefore can be represented by a matrix (Appendix A). We denote this matrix by $\widehat{u} \in \mathbb{R}^{3\times3}$, pronounced "$u$ hat." It is immediate to verify by substitution that this matrix is given by[5]

$$\widehat{u} \doteq \begin{bmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{bmatrix} \in \mathbb{R}^{3\times3}. \tag{2.2}$$

Hence, we can write $u \times v = \widehat{u}v$. Note that $\widehat{u}$ is a 3×3 skew-symmetric matrix, i.e. $\widehat{u}^T = -\widehat{u}$ (see Appendix A).

Example 2.3 (Right-hand rule). It is immediate to verify that for $e_1 \doteq [1,0,0]^T, e_2 \doteq [0,1,0]^T \in \mathbb{R}^3$, we have $e_1 \times e_2 = [0,0,1]^T \doteq e_3$. That is, for a standard Cartesian frame, the cross product of the principal axes X and Y gives the principal axis Z. The cross product therefore conforms to the *right-hand rule*. See Figure 2.1. ■

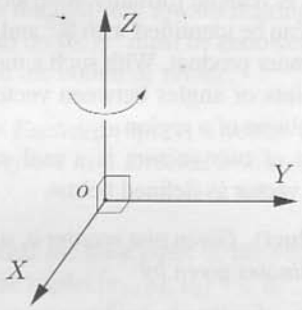

Figure 2.1. A right-handed (X, Y, Z) coordinate frame.

The cross product, therefore, naturally defines a map between a vector u and a 3×3 skew-symmetric matrix $\widehat{u}$. By inspection, the converse of this statement is clearly true, since we can easily identify a three-dimensional vector associated with every 3×3 skew-symmetric matrix (just extract u_1, u_2, u_3 from (2.2)).

Lemma 2.4 (Skew-symmetric matrix). *A matrix $M \in \mathbb{R}^{3\times3}$ is skew-symmetric if and only if $M = \widehat{u}$ for some $u \in \mathbb{R}^3$.*

Therefore, the vector space $\mathbb{R}^3$ and the space of all skew-symmetric 3×3 matrices, called $so(3)$,[6] are isomorphic (i.e. there exists a one-to-one map that preserves the vector space structure). The isomorphism is the so-called *hat operator*

$$\wedge : \mathbb{R}^3 \to so(3); \quad u \mapsto \widehat{u},$$

[5] In some literature, the matrix $\widehat{u}$ is denoted by $u_\times$ or $[u]_\times$.

[6] We will explain the reason for this name later in this chapter.

and its inverse map, called the *vee operator*, which extracts the components of the vector u from a skew-symmetric matrix $\widehat{u}$, is given by

$$\vee : so(3) \to \mathbb{R}^3; \quad \widehat{u} \mapsto \widehat{u}^\vee = u.$$

2.2 Rigid-body motion

Consider an object moving in front of a camera. In order to describe its motion one should, in principle, specify the trajectory of every single point on the object, for instance, by specifying coordinates of a point as a function of time $\boldsymbol{X}(t)$. Fortunately, for rigid objects we do not need to specify the motion of every point. As we will see shortly, it is sufficient to specify the motion of one (instead of every) point, and the motion of three coordinate axes attached to that point. The reason is that for every rigid object, the distance between any two points on it does not change over time as the object moves. See Figure 2.2.

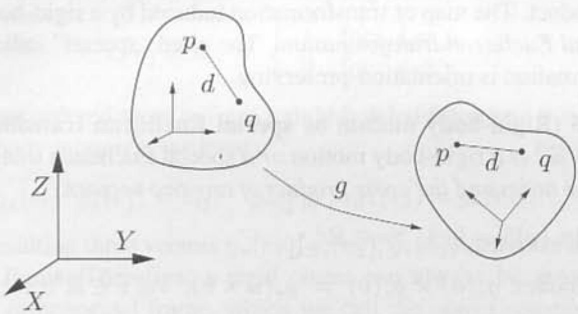

Figure 2.2. A motion of a rigid body preserves the distance d between any pair of points (p, q) on it.

Thus, if $\boldsymbol{X}(t)$ and $\boldsymbol{Y}(t)$ are the coordinates of any two points p and q on the object, respectively, the distance between them is constant:

$$\|\boldsymbol{X}(t) - \boldsymbol{Y}(t)\| \equiv \text{constant}, \quad \forall\, t \in \mathbb{R}. \tag{2.3}$$

A *rigid-body motion* (or rigid-body transformation) is then a family of maps that describe how the coordinates of every point on a rigid object change in time while satisfying (2.3). We denote such a map by

$$g(t) : \mathbb{R}^3 \to \mathbb{R}^3; \quad \boldsymbol{X} \mapsto g(t)(\boldsymbol{X}).$$

If instead of looking at the entire continuous path of the moving object, we concentrate on the map between its initial and final configuration, we have a *rigid-body displacement*, denoted by

$$g : \mathbb{R}^3 \to \mathbb{R}^3; \quad \boldsymbol{X} \mapsto g(\boldsymbol{X}).$$

Besides transforming the coordinates of points, g also induces a transformation on vectors. Suppose that v is a vector defined by two points p and q with coordinates

$v = \boldsymbol{Y} - \boldsymbol{X}$; then, after the transformation g, we obtain a new vector[7]

$$u = g_*(v) \doteq g(\boldsymbol{Y}) - g(\boldsymbol{X}).$$

Since g preserves the distance between points, we have that $\|g_*(v)\| = \|v\|$ for all free vectors $v \in \mathbb{R}^3$.

A map that preserves the distance is called a *Euclidean transformation*. In the 3-D space, the set all Euclidean transformations is denoted by $E(3)$. Note that preserving distances between points is not sufficient to characterize a rigid object moving in space. In fact, there are transformations that preserve distances, and yet they are not physically realizable. For instance, the map

$$f : [X_1, X_2, X_3]^T \mapsto [X_1, X_2, -X_3]^T$$

preserves distances but not orientations. It corresponds to a reflection of points in the XY-plane as a double-sided mirror. To rule out this kind of maps,[8] we require that any rigid-body motion, besides preserving distances, preserves orientations as well. That is, in addition to preserving the norm of vectors, it must also preserve their cross product. The map or transformation induced by a rigid-body motion is called a *special Euclidean transformation*. The word "special" indicates the fact that a transformation is orientation-preserving.

Definition 2.5 (Rigid-body motion or special Euclidean transformation). *A map $g : \mathbb{R}^3 \to \mathbb{R}^3$ is a* rigid-body motion *or a* special Euclidean transformation *if it preserves the norm and the cross product of any two vectors,*

1. *norm:* $\|g_*(v)\| = \|v\|, \ \forall v \in \mathbb{R}^3$,
2. *cross product:* $g_*(u) \times g_*(v) = g_*(u \times v), \ \forall u, v \in \mathbb{R}^3$.

The collection of all such motions or transformations is denoted by $SE(3)$.

In the above definition of rigid-body motions, it is not immediately obvious that the angles between vectors are preserved. However, the inner product $\langle \cdot, \cdot \rangle$ can be expressed in terms of the norm $\| \cdot \|$ by the *polarization identity*

$$\boxed{\langle u, v \rangle = \frac{1}{4}\left(\|u+v\|^2 - \|u-v\|^2\right)} \tag{2.4}$$

and, since $\|u + v\| = \|g_*(u) + g_*(v)\|$, one can conclude that, for any rigid-body motion g,

$$\langle u, v \rangle = \langle g_*(u), g_*(v) \rangle, \quad \forall u, v \in \mathbb{R}^3. \tag{2.5}$$

In other words, a rigid-body motion can also be defined as one that preserves both the inner product and the cross product.

[7]The use of g_* here is consistent with the so-called *push-forward map* or *differential operator* of g in differential geometry, which denotes the action of a differentiable map on the tangent spaces of its domains.

[8]In Chapter 10, however, we will study the important role of reflections in multiple-view geometry.

Example 2.6 (Triple product and volume). From the definition of a rigid-body motion, one can show that it also preserves the so-called *triple product* among three vectors:

$$\langle g_*(u), g_*(v) \times g_*(w) \rangle = \langle u, v \times w \rangle.$$

Since the triple product corresponds to the volume of the parallelepiped spanned by the three vectors, rigid-body motion also preserves volumes. ■

How do these properties help us describe a rigid-body motion concisely? The fact that distances and orientations are preserved by a rigid-body motion means that individual points cannot move relative to each other. As a consequence, a rigid-body motion can be described by the motion of a chosen point on the body and the rotation of a coordinate frame attached to that point. In order to see this, we represent the *configuration* of a rigid body by attaching a Cartesian coordinate frame to some point on the rigid body, and we will keep track of the motion of this coordinate frame relative to a fixed *world (reference) frame*.

To this end, consider a coordinate frame, with its principal axes given by three *orthonormal* vectors $e_1, e_2, e_3 \in \mathbb{R}^3$; that is, they satisfy

$$e_i^T e_j = \delta_{ij} \doteq \begin{cases} 1 & \text{for} \quad i = j, \\ 0 & \text{for} \quad i \neq j. \end{cases} \tag{2.6}$$

The vectors are ordered so as to form a right-handed frame: $e_1 \times e_2 = e_3$. Then, after a rigid-body motion g, we have

$$g_*(e_i)^T g_*(e_j) = \delta_{ij}, \quad g_*(e_1) \times g_*(e_2) = g_*(e_3). \tag{2.7}$$

That is, the resulting three vectors $g_*(e_1), g_*(e_2), g_*(e_3)$ still form a right-handed orthonormal frame. Therefore, a rigid object can always be associated with a right-handed orthonormal frame, which we call the *object coordinate frame* or the *body coordinate frame*, and its rigid-body motion can be entirely specified by the motion of such a frame.

In Figure 2.3 we show an object, in this case a camera, moving relative to a world reference frame $W : (X, Y, Z)$ selected in advance. In order to specify the configuration of the camera relative to the world frame W, one may pick a fixed point o on the camera and attach to it an object frame, in this case called a camera frame,[9] $C : (x, y, z)$. When the camera moves, the camera frame also moves along with the camera. The configuration of the camera is then determined by two components:

1. the vector between the origin o of the world frame and that of the camera frame, $g(o)$, called the "translational" part and denoted by T;

2. the relative orientation of the camera frame C, with coordinate axes (x, y, z), relative to the fixed world frame W with coordinate axes (X, Y, Z), called the "rotational" part and denoted by R.

[9]Here, to distinguish the two coordinate frames, we use lower-case x, y, z for coordinates in the camera frame.

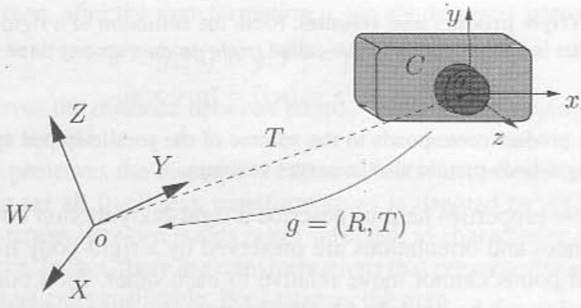

Figure 2.3. A rigid-body motion between a camera frame C: (x, y, z) and a world coordinate frame W: (X, Y, Z).

In the problems we consider in this book, there is no obvious choice of the world reference frame and its origin o. Therefore, we can choose the world frame to be attached to the camera and specify the translation and rotation of the scene relative to that frame (as long as it is rigid), or we could attach the world frame to the scene and specify the motion of the camera relative to that frame. All that matters is the *relative* motion between the scene and the camera; the choice of the world reference frame is, from the point of view of geometry, arbitrary.[10]

If we can move a rigid object (e.g., a camera) from one place to another, we can certainly reverse the action and put it back to its original position. Similarly, we can combine several motions to generate a new one. Roughly speaking, this property of invertibility and composition can be mathematically characterized by the notion of "group" (Appendix A). As we will soon see, the set of rigid-body motions is indeed a group, the so-called *special Euclidean group*. However, the abstract notion of group is not useful until we can give it an explicit representation and use it for computation. In the next few sections, we will focus on studying in detail how to represent rigid-body motions in terms of matrices.[11] More specifically, we will show that any rigid-body motion can be represented as a 4×4 matrix. For simplicity, we start with the rotational component of a rigid-body motion.

2.3 Rotational motion and its representations

2.3.1 *Orthogonal matrix representation of rotations*

Suppose we have a rigid object rotating about a fixed point $o \in \mathbb{E}^3$. How do we describe its orientation relative to a chosen coordinate frame, say W? Without loss of generality, we may always assume that the origin of the world frame is

[10]The human vision literature, on the other hand, debates whether the primate brain maintains a view-centered or an object-centered representation of the world.

[11]The notion of matrix representation for a group is introduced in Appendix A.

the center of rotation o. If this is not the case, simply translate the origin to the point o. We now attach another coordinate frame, say C, to the rotating object, say a camera, with its origin also at o. The relation between these two coordinate frames is illustrated in Figure 2.4.

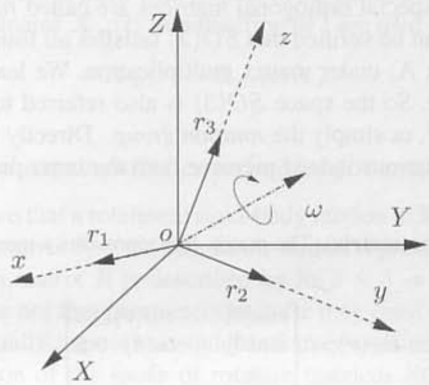

Figure 2.4. Rotation of a rigid body about a fixed point o and along the axis ω. The coordinate frame W (solid line) is fixed, and the coordinate frame C (dashed line) is attached to the rotating rigid body.

The configuration (or "orientation") of the frame C relative to the frame W is determined by the coordinates of the three orthonormal vectors $r_1 = g_*(e_1), r_2 = g_*(e_2), r_3 = g_*(e_3) \in \mathbb{R}^3$ relative to the world frame W, as shown in Figure 2.4. The three vectors r_1, r_2, r_3 are simply the unit vectors along the three principal axes x, y, z of the frame C, respectively. The configuration of the rotating object is then completely determined by the 3×3 matrix

$$R_{wc} \doteq [r_1, r_2, r_3] \in \mathbb{R}^{3\times 3},$$

with r_1, r_2, r_3 stacked in order as its three columns. Since r_1, r_2, r_3 form an orthonormal frame, it follows that

$$r_i^T r_j = \delta_{ij} \doteq \begin{cases} 1 & \text{for} \quad i = j, \\ 0 & \text{for} \quad i \neq j, \end{cases} \qquad \forall i, j \in \{1, 2, 3\}.$$

This can be written in matrix form as

$$R_{wc}^T R_{wc} = R_{wc} R_{wc}^T = I.$$

Any matrix that satisfies the above identity is called an *orthogonal matrix*. It follows from the above definition that the inverse of an orthogonal matrix is simply its transpose: $R_{wc}^{-1} = R_{wc}^T$. Since r_1, r_2, r_3 form a right-handed frame, we further have the condition that the determinant of R_{wc} must be $+1$.[12] Hence R_{wc} is a *special orthogonal matrix*, where as before, the word "special" indicates that it is

[12]This can easily be seen by computing the determinant of the rotation matrix $\det(R) = r_1^T(r_2 \times r_3)$, which is equal to $+1$.

orientation-preserving. The space of all such special orthogonal matrices in $\mathbb{R}^{3\times 3}$ is usually denoted by

$$\boxed{SO(3) \doteq \left\{R \in \mathbb{R}^{3\times 3} \mid R^T R = I, \det(R) = +1\right\}.}$$

Traditionally, 3×3 special orthogonal matrices are called *rotation matrices* for obvious reasons. It can be verified that $SO(3)$ satisfies all four axioms of a group (defined in Appendix A) under matrix multiplication. We leave the proof to the reader as an exercise. So the space $SO(3)$ is also referred to as the *special orthogonal group* of $\mathbb{R}^3$, or simply the *rotation group*. Directly from the definition, one can show that rotations indeed preserve both the inner product and the cross product of vectors.

Example 2.7 (A rotation matrix). The matrix that represents a rotation about the Z-axis by an angle θ is

$$R_Z(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The reader can similarly derive matrices for rotation about the X-axis or the Y-axis. In the next section we will study how to represent a rotation about any axis. ∎

Going back to Figure 2.4, every rotation matrix $R_{wc} \in SO(3)$ represents a possible configuration of the object rotated about the point o. Besides this, R_{wc} takes another role as the matrix that represents the coordinate transformation from the frame C to the frame W. To see this, suppose that for a given a point $p \in \mathbb{E}^3$, its coordinates with respect to the frame W are $\boldsymbol{X}_w = [X_{1w}, X_{2w}, X_{3w}]^T \in \mathbb{R}^3$. Since r_1, r_2, r_3 also form a basis for $\mathbb{R}^3$, $\boldsymbol{X}_w$ can be expressed as a linear combination of these three vectors, say $\boldsymbol{X}_w = X_{1c}r_1 + X_{2c}r_2 + X_{3c}r_3$ with $[X_{1c}, X_{2c}, X_{3c}]^T \in \mathbb{R}^3$. Obviously, $\boldsymbol{X}_c = [X_{1c}, X_{2c}, X_{3c}]^T$ are the coordinates of the same point p with respect to the frame C. Therefore, we have

$$\boldsymbol{X}_w = X_{1c}r_1 + X_{2c}r_2 + X_{3c}r_3 = R_{wc}\boldsymbol{X}_c.$$

In this equation, the matrix R_{wc} transforms the coordinates $\boldsymbol{X}_c$ of a point p relative to the frame C to its coordinates $\boldsymbol{X}_w$ relative to the frame W. Since R_{wc} is a rotation matrix, its inverse is simply its transpose,

$$\boldsymbol{X}_c = R_{wc}^{-1}\boldsymbol{X}_w = R_{wc}^T\boldsymbol{X}_w.$$

That is, the inverse transformation of a rotation is also a rotation; we call it R_{cw}, following an established convention, so that

$$R_{cw} = R_{wc}^{-1} = R_{wc}^T.$$

The configuration of a continuously rotating object can then be described as a trajectory $R(t) : t \mapsto SO(3)$ in the space $SO(3)$. When the starting time is not $t = 0$, the relative motion between time t_2 and time t_1 will be denoted as $R(t_2, t_1)$. The composition law of the rotation group (see Appendix A) implies

$$R(t_2, t_0) = R(t_2, t_1)R(t_1, t_0), \quad \forall t_0 < t_1 < t_2 \in \mathbb{R}.$$

For a rotating camera, the world coordinates $\boldsymbol{X}_w$ of a fixed 3-D point p are transformed to its coordinates relative to the camera frame C by

$$\boldsymbol{X}_c(t) = R_{cw}(t)\boldsymbol{X}_w.$$

Alternatively, if a point p is fixed with respect to the camera frame has coordinates $\boldsymbol{X}_c$, its world coordinates $\boldsymbol{X}_w(t)$ as a function of t are then given by

$$\boldsymbol{X}_w(t) = R_{wc}(t)\boldsymbol{X}_c.$$

2.3.2 Canonical exponential coordinates for rotations

So far, we have shown that a rotational rigid-body motion in $\mathbb{E}^3$ can be represented by a 3×3 rotation matrix $R \in SO(3)$. In the matrix representation that we have so far, each rotation matrix R is described by its $3 \times 3 = 9$ entries. However, these nine entries are not free parameters because they must satisfy the constraint $R^T R = I$. This actually imposes six independent constraints on the nine entries. Hence, the dimension of the space of rotation matrices $SO(3)$ should be only three, and six parameters out of the nine are in fact redundant. In this subsection and Appendix 2.A, we will introduce a few explicit parameterizations for the space of rotation matrices.

Given a trajectory $R(t) : \mathbb{R} \to SO(3)$ that describes a continuous rotational motion, the rotation must satisfy the following constraint

$$R(t)R^T(t) = I.$$

Computing the derivative of the above equation with respect to time t and noticing that the right-hand side is a constant matrix, we obtain

$$\dot{R}(t)R^T(t) + R(t)\dot{R}^T(t) = 0 \quad \Rightarrow \quad \dot{R}(t)R^T(t) = -(\dot{R}(t)R^T(t))^T.$$

The resulting equation reflects the fact that the matrix $\dot{R}(t)R^T(t) \in \mathbb{R}^{3\times 3}$ is a skew-symmetric matrix. Then, as we have seen in Lemma 2.4, there must exist a vector, say $\omega(t) \in \mathbb{R}^3$, such that

$$\dot{R}(t)R^T(t) = \widehat{\omega}(t).$$

Multiplying both sides by $R(t)$ on the right yields

$$\dot{R}(t) = \widehat{\omega}(t)R(t). \tag{2.8}$$

Notice that from the above equation, if $R(t_0) = I$ for $t = t_0$, we have $\dot{R}(t_0) = \widehat{\omega}(t_0)$. Hence, around the identity matrix I, a skew-symmetric matrix gives a first-order approximation to a rotation matrix:

$$R(t_0 + dt) \approx I + \widehat{\omega}(t_0)\, dt.$$

As we have anticipated, the space of all skew-symmetric matrices is denoted by

$$\boxed{so(3) \doteq \left\{\widehat{\omega} \in \mathbb{R}^{3\times 3} \mid \omega \in \mathbb{R}^3\right\},} \tag{2.9}$$

and following the above observation, it is also called the *tangent space* at the identity of the rotation group $SO(3)$.[13] If $R(t)$ is not at the identity, the tangent space at $R(t)$ is simply $so(3)$ transported to $R(t)$ by a multiplication by $R(t)$ on the right: $\dot{R}(t) = \widehat{\omega}(t)R(t)$. This also shows that, locally, elements of $SO(3)$ depend on only three parameters, $(\omega_1, \omega_2, \omega_3)$.

Having understood its local approximation, we will now use this knowledge to obtain a useful representation for the rotation matrix. Let us start by assuming that the matrix $\widehat{\omega}$ in (2.8) is constant,

$$\dot{R}(t) = \widehat{\omega}R(t). \tag{2.10}$$

In the above equation, $R(t)$ can be interpreted as the *state transition matrix* for the following linear ordinary differential equation (ODE):

$$\dot{x}(t) = \widehat{\omega}x(t), \quad x(t) \in \mathbb{R}^3. \tag{2.11}$$

It is then immediate to verify that the solution to the above ODE is given by

$$x(t) = e^{\widehat{\omega}t}x(0), \tag{2.12}$$

where $e^{\widehat{\omega}t}$ is the matrix exponential

$$e^{\widehat{\omega}t} = I + \widehat{\omega}t + \frac{(\widehat{\omega}t)^2}{2!} + \cdots + \frac{(\widehat{\omega}t)^n}{n!} + \cdots. \tag{2.13}$$

The exponential $e^{\widehat{\omega}t}$ is also often denoted by $\exp(\widehat{\omega}t)$. Due to the uniqueness of the solution to the ODE (2.11), and assuming $R(0) = I$ is the initial condition for (2.10), we must have

$$\boxed{R(t) = e^{\widehat{\omega}t}.} \tag{2.14}$$

To verify that the matrix $e^{\widehat{\omega}t}$ is indeed a rotation matrix, one can directly show from the definition of the matrix exponential that

$$(e^{\widehat{\omega}t})^{-1} = e^{-\widehat{\omega}t} = e^{\widehat{\omega}^T t} = (e^{\widehat{\omega}t})^T.$$

Hence $(e^{\widehat{\omega}t})^T e^{\widehat{\omega}t} = I$. It remains to show that $\det(e^{\widehat{\omega}t}) = +1$, and we leave this fact to the reader as an exercise (see Exercise 2.12). *A physical interpretation of equation (2.14) is that if $\|\omega\| = 1$, then $R(t) = e^{\widehat{\omega}t}$ is simply a rotation around the axis $\omega \in \mathbb{R}^3$ by an angle of t radians.*[14] In general, t can be absorbed into ω, so we have $R = e^{\widehat{\omega}}$ for ω with arbitrary norm. So, the matrix exponential (2.13) indeed defines a map from the space $so(3)$ to $SO(3)$, the so-called *exponential map*

$$\exp : so(3) \to SO(3); \quad \widehat{\omega} \mapsto e^{\widehat{\omega}}.$$

Note that we obtained the expression (2.14) by assuming that the $\omega(t)$ in (2.8) is constant. This is, however, not always the case. A question naturally arises:

---

[13] Since $SO(3)$ is a Lie group, $so(3)$ is called its Lie algebra.

[14] We can use either $e^{\widehat{\omega}\theta}$, where θ encodes explicitly the rotation angle and $\|\omega\| = 1$, or more simply $e^{\widehat{\omega}}$ where $\|\omega\|$ encodes the rotation angle.

can every rotation matrix $R \in SO(3)$ be expressed in an exponential form as in (2.14)? The answer is yes, and the fact is stated as the following theorem.

Theorem 2.8 (Logarithm of $SO(3)$). *For any $R \in SO(3)$, there exists a (not necessarily unique) $\omega \in \mathbb{R}^3$ such that $R = \exp(\widehat{\omega})$. We denote the inverse of the exponential map by $\widehat{\omega} = \log(R)$.*

Proof. The proof of this theorem is by construction: if the rotation matrix $R \neq I$ is given as

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix},$$

the corresponding ω is given by

$$\boxed{\|\omega\| = \cos^{-1}\left(\frac{\mathrm{trace}(R) - 1}{2}\right), \quad \frac{\omega}{\|\omega\|} = \frac{1}{2\sin(\|\omega\|)} \begin{bmatrix} r_{32} - r_{23} \\ r_{13} - r_{31} \\ r_{21} - r_{12} \end{bmatrix}.} \tag{2.15}$$

If $R = I$, then $\|\omega\| = 0$, and $\frac{\omega}{\|\omega\|}$ is not determined (and therefore can be chosen arbitrarily). □

The significance of this theorem is that any rotation matrix can be realized by rotating around some fixed axis ω by a certain angle $\|\omega\|$. However, the exponential map from $so(3)$ to $SO(3)$ is not one-to-one, since any vector of the form $2k\pi\omega$ with k integer would give rise to the same R. This will become clear after we have introduced the so-called *Rodrigues' formula* for computing $R = e^{\widehat{\omega}}$.

From the constructive proof of Theorem 2.8, we know how to compute the exponential coordinates ω for a given rotation matrix $R \in SO(3)$. On the other hand, given ω, how do we effectively compute the corresponding rotation matrix $R = e^{\widehat{\omega}}$? One can certainly use the series (2.13) from the definition. The following theorem, however, provides a very useful formula that simplifies the computation significantly.

Theorem 2.9 (Rodrigues' formula for a rotation matrix). *Given $\omega \in \mathbb{R}^3$, the matrix exponential $R = e^{\widehat{\omega}}$ is given by*

$$\boxed{e^{\widehat{\omega}} = I + \frac{\widehat{\omega}}{\|\omega\|}\sin(\|\omega\|) + \frac{\widehat{\omega}^2}{\|\omega\|^2}(1 - \cos(\|\omega\|)).} \tag{2.16}$$

Proof. Let $t = \|\omega\|$ and redefine ω to be of unit length. Then, it is immediate to verify that powers of $\widehat{\omega}$ can be reduced by the following two formulae

$$\widehat{\omega}^2 = \omega\omega^T - I, \quad \widehat{\omega}^3 = -\widehat{\omega}.$$

Hence the exponential series (2.13) can be simplified as

$$e^{\widehat{\omega}t} = I + \left(t - \frac{t^3}{3!} + \frac{t^5}{5!} - \cdots\right)\widehat{\omega} + \left(\frac{t^2}{2!} - \frac{t^4}{4!} + \frac{t^6}{6!} - \cdots\right)\widehat{\omega}^2.$$

The two sets of parentheses contain the Taylor series for $\sin(t)$ and $(1-\cos(t))$, respectively. Thus, we have $e^{\widehat{\omega}t} = I + \widehat{\omega}\sin(t) + \widehat{\omega}^2(1-\cos(t))$. □

Using Rodrigues' formula, it is immediate to see that if $\|\omega\| = 1, t = 2k\pi$, we have

$$e^{\widehat{\omega}2k\pi} = I$$

for all $k \in \mathbb{Z}$. Hence, for a given rotation matrix $R \in SO(3)$, there are infinitely many exponential coordinates $\omega \in \mathbb{R}^3$ such that $e^{\widehat{\omega}} = R$. The exponential map $\exp : so(3) \to SO(3)$ is therefore *not* one-to-one. It is also useful to know that the exponential map is not *commutative*, i.e. for two $\widehat{\omega}_1, \widehat{\omega}_2 \in so(3)$,

$$e^{\widehat{\omega}_1}e^{\widehat{\omega}_2} \neq e^{\widehat{\omega}_2}e^{\widehat{\omega}_1} \neq e^{\widehat{\omega}_1+\widehat{\omega}_2},$$

unless $\widehat{\omega}_1\widehat{\omega}_2 = \widehat{\omega}_2\widehat{\omega}_1$.

Remark 2.10. *In general, the difference between $\widehat{\omega}_1\widehat{\omega}_2$ and $\widehat{\omega}_2\widehat{\omega}_1$ is called the* Lie bracket *on $so(3)$, denoted by*

$$[\widehat{\omega}_1, \widehat{\omega}_2] = \widehat{\omega}_1\widehat{\omega}_2 - \widehat{\omega}_2\widehat{\omega}_1, \quad \forall\, \widehat{\omega}_1, \widehat{\omega}_2 \in so(3).$$

From the definition above it can be verified that $[\widehat{\omega}_1, \widehat{\omega}_2]$ is also a skew-symmetric matrix in $so(3)$. The linear structure of $so(3)$ together with the Lie bracket form the Lie algebra *of the (Lie) group $SO(3)$. For more details on the Lie group structure of $SO(3)$, the reader may refer to [Murray et al., 1993]. Given $\widehat{\omega}$, the set of all rotation matrices $e^{\widehat{\omega}t}, t \in \mathbb{R}$, is then a* one-parameter *subgroup of $SO(3)$, i.e. the planar rotation group $SO(2)$. The multiplication in such a subgroup is always commutative, since for the same $\omega \in \mathbb{R}^3$, we have*

$$e^{\widehat{\omega}t_1}e^{\widehat{\omega}t_2} = e^{\widehat{\omega}t_2}e^{\widehat{\omega}t_1} = e^{\widehat{\omega}(t_1+t_2)}, \quad \forall t_1, t_2 \in \mathbb{R}.$$

The exponential coordinates introduced above provide a local parameterization for rotation matrices. There are also other ways to parameterize rotation matrices, either globally or locally, among which *quaternions* and *Euler angles* (or more formally, *Lie-Cartan coordinates*) are two popular choices. We leave more detailed discussions to Appendix 2.A at the end of this chapter. We use exponential coordinates because they are simpler and more intuitive.

2.4 Rigid-body motion and its representations

In the previous section, we studied purely rotational rigid-body motions and how to represent and compute a rotation matrix. In this section, we will study how to represent a rigid-body motion in general, a motion with both rotation and translation.

Figure 2.5 illustrates a moving rigid object with a coordinate frame C attached to it. To describe the coordinates of a point p on the object with respect to the world frame W, it is clear from the figure that the vector $\boldsymbol{X}_w$ is simply the sum of

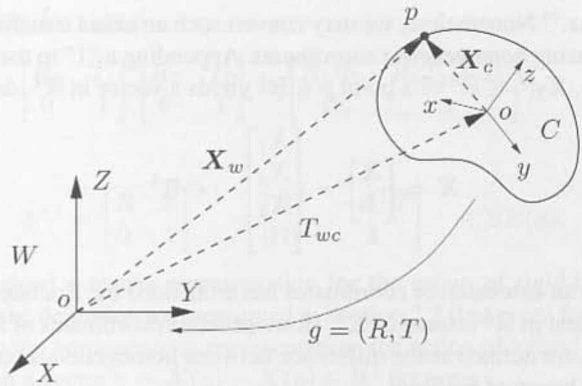

Figure 2.5. A rigid-body motion between a moving frame C and a world frame W.

the translation $T_{wc} \in \mathbb{R}^3$ of the origin of the frame C relative to that of the frame W and the vector $\boldsymbol{X}_c$ but expressed relative to the frame W. Since $\boldsymbol{X}_c$ are the coordinates of the point p relative to the frame C, with respect to the world frame W, it becomes $R_{wc}\boldsymbol{X}_c$, where $R_{wc} \in SO(3)$ is the relative rotation between the two frames. Hence, the coordinates $\boldsymbol{X}_w$ are given by

$$\boldsymbol{X}_w = R_{wc}\boldsymbol{X}_c + T_{wc}. \tag{2.17}$$

Usually, we denote the full rigid-body motion by $g_{wc} = (R_{wc}, T_{wc})$, or simply $g = (R, T)$ if the frames involved are clear from the context. Then g represents not only a description of the configuration of the rigid-body object but also a transformation of coordinates between the two frames. In compact form, we write

$$\boldsymbol{X}_w = g_{wc}(\boldsymbol{X}_c).$$

The set of all possible configurations of a rigid body can then be described by the space of rigid-body motions or special Euclidean transformations

$$\boxed{SE(3) \doteq \left\{g = (R, T) \mid R \in SO(3), T \in \mathbb{R}^3\right\}.}$$

Note that $g = (R, T)$ is not yet a matrix representation for $SE(3)$.[15] To obtain such a representation, we need to introduce the so-called homogeneous coordinates. We will introduce only what is needed to carry our study of rigid-body motions.

2.4.1 *Homogeneous representation*

One may have already noticed from equation (2.17) that in contrast to the pure rotation case, the coordinate transformation for a full rigid-body motion is not

[15]For this to be the case, the composition of two rigid-body motions needs to be the multiplication of two matrices. See Appendix A.

linear but *affine*.[16] Nonetheless, we may convert such an affine transformation to a linear one by using *homogeneous coordinates*. Appending a "1" to the coordinates $\boldsymbol{X} = [X_1, X_2, X_3]^T \in \mathbb{R}^3$ of a point $p \in \mathbb{E}^3$ yields a vector in $\mathbb{R}^4$, denoted by

$$\bar{\boldsymbol{X}} \doteq \begin{bmatrix} \boldsymbol{X} \\ 1 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ 1 \end{bmatrix} \in \mathbb{R}^4.$$

In effect, such an extension of coordinates has embedded the Euclidean space $\mathbb{E}^3$ into a hyperplane in $\mathbb{R}^4$ instead of $\mathbb{R}^3$. Homogeneous coordinates of a vector $v = \boldsymbol{X}(q) - \boldsymbol{X}(p)$ are defined as the difference between homogeneous coordinates of the two points hence of the form

$$\bar{v} \doteq \begin{bmatrix} v \\ 0 \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}(q) \\ 1 \end{bmatrix} - \begin{bmatrix} \boldsymbol{X}(p) \\ 1 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ 0 \end{bmatrix} \in \mathbb{R}^4.$$

Notice that in $\mathbb{R}^4$, vectors of the above form give rise to a subspace, and all linear structures of the original vectors $v \in \mathbb{R}^3$ are perfectly preserved by the new representation. Using the new notation, the (affine) transformation (2.17) can then be rewritten in a "linear" form

$$\bar{\boldsymbol{X}}_w = \begin{bmatrix} \boldsymbol{X}_w \\ 1 \end{bmatrix} = \begin{bmatrix} R_{wc} & T_{wc} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{X}_c \\ 1 \end{bmatrix} \doteq \bar{g}_{wc} \bar{\boldsymbol{X}}_c,$$

where the 4×4 matrix $\bar{g}_{wc} \in \mathbb{R}^{4\times 4}$ is called the *homogeneous representation* of the rigid-body motion $g_{wc} = (R_{wc}, T_{wc}) \in SE(3)$. In general, if $g = (R, T)$, then its homogeneous representation is

$$\bar{g} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{4\times 4}. \tag{2.18}$$

Notice that by introducing a little redundancy into the notation, we can represent a rigid-body transformation of coordinates by a linear matrix multiplication. The homogeneous representation of g in (2.18) gives rise to a natural matrix representation of the special Euclidean transformations

$$\boxed{SE(3) \doteq \left\{ \bar{g} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \middle| R \in SO(3), T \in \mathbb{R}^3 \right\} \quad \subset \mathbb{R}^{4\times 4}.}$$

Using this representation, it is then straightforward to verify that the set $SE(3)$ indeed satisfies all the requirements of a group (Appendix A). In particular, $\forall g_1, g_2$

[16] We say that two vectors u, v are related by a *linear* transformation if $u = Av$ for some matrix A, and by an *affine* transformation if $u = Av + b$ for some matrix A and vector b. See Appendix A.

and $g \in SE(3)$, we have

$$\bar{g}_1\bar{g}_2 = \begin{bmatrix} R_1 & T_1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} R_2 & T_2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R_1R_2 & R_1T_2 + T_1 \\ 0 & 1 \end{bmatrix} \quad \in SE(3)$$

and

$$\bar{g}^{-1} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} R^T & -R^TT \\ 0 & 1 \end{bmatrix} \quad \in SE(3).$$

Thus, $\bar{g}$ is indeed a matrix representation for the group of rigid-body motions according to the definition we mentioned in Section 2.2 (but given formally in Appendix A). In the homogeneous representation, the action of a rigid-body motion $g \in SE(3)$ on a vector $v = \boldsymbol{X}(q) - \boldsymbol{X}(p) \in \mathbb{R}^3$ becomes

$$\bar{g}_*(\bar{v}) = \bar{g}\bar{\boldsymbol{X}}(q) - \bar{g}\bar{\boldsymbol{X}}(p) = \bar{g}\bar{v}.$$

That is, the action is also simply represented by a matrix multiplication. In the 3-D coordinates, we have $g_*(v) = Rv$, since only rotational part affects vectors. The reader can verify that such an action preserves both the inner product and the cross product. As can be seen, *rigid motions act differently on points (rotation and translation) than they do on vectors (rotation only).*

2.4.2 Canonical exponential coordinates for rigid-body motions

In Section 2.3.2, we studied exponential coordinates for a rotation matrix $R \in SO(3)$. Similar coordinatization also exists for the homogeneous representation of a full rigid-body motion $g \in SE(3)$. For the rest of this section, we demonstrate how to extend the results we have developed for the rotational motion to a full rigid-body motion. The results developed here will be extensively used throughout the book. The derivation parallels the case of a pure rotation in Section 2.3.2.

Consider the motion of a continuously moving rigid body described by a trajectory on $SE(3)$: $g(t) = (R(t), T(t))$, or in the homogeneous representation

$$g(t) = \begin{bmatrix} R(t) & T(t) \\ 0 & 1 \end{bmatrix} \quad \in \mathbb{R}^{4\times 4}.$$

From now on, for simplicity, whenever there is no ambiguity, we will remove the bar "$-$" to indicate a homogeneous representation and simply use g. We will use the same convention for points, $\boldsymbol{X}$ for $\bar{\boldsymbol{X}}$, and for vectors, v for $\bar{v}$, whenever their correct dimension is clear from the context.

In analogy with the case of a pure rotation, let us first look at the structure of the matrix

$$\dot{g}(t)g^{-1}(t) = \begin{bmatrix} \dot{R}(t)R^T(t) & \dot{T}(t) - \dot{R}(t)R^T(t)T(t) \\ 0 & 0 \end{bmatrix} \quad \in \mathbb{R}^{4\times 4}. \tag{2.19}$$

From our study of the rotation matrix, we know that $\dot{R}(t)R^T(t)$ is a skew-symmetric matrix; i.e. there exists $\widehat{\omega}(t) \in so(3)$ such that $\widehat{\omega}(t) = \dot{R}(t)R^T(t)$.

Define a vector $v(t) \in \mathbb{R}^3$ such that $v(t) = \dot{T}(t) - \widehat{\omega}(t)T(t)$. Then the above equation becomes

$$\dot{g}(t)g^{-1}(t) = \begin{bmatrix} \widehat{\omega}(t) & v(t) \\ 0 & 0 \end{bmatrix} \quad \in \mathbb{R}^{4\times 4}.$$

If we further define a matrix $\widehat{\xi} \in \mathbb{R}^{4\times 4}$ to be

$$\widehat{\xi}(t) = \begin{bmatrix} \widehat{\omega}(t) & v(t) \\ 0 & 0 \end{bmatrix},$$

then we have

$$\dot{g}(t) = (\dot{g}(t)g^{-1}(t))g(t) = \widehat{\xi}(t)g(t), \tag{2.20}$$

where $\widehat{\xi}$ can be viewed as the "tangent vector" along the curve of $g(t)$ and can be used to approximate $g(t)$ locally:

$$g(t+dt) \approx g(t) + \widehat{\xi}(t)g(t)dt = \left(I + \widehat{\xi}(t)dt\right)g(t).$$

A 4×4 matrix of the form of $\widehat{\xi}$ is called a *twist*. The set of all twists is denoted by

$$\boxed{se(3) \doteq \left\{ \widehat{\xi} = \begin{bmatrix} \widehat{\omega} & v \\ 0 & 0 \end{bmatrix} \middle| \ \widehat{\omega} \in so(3), v \in \mathbb{R}^3 \right\} \quad \subset \mathbb{R}^{4\times 4}.}$$

The set $se(3)$ is called the tangent space (or Lie algebra) of the matrix group $SE(3)$. We also define two operators "$\vee$" and "$\wedge$" to convert between a twist $\widehat{\xi} \in se(3)$ and its *twist coordinates* $\xi \in \mathbb{R}^6$ as follows:

$$\begin{bmatrix} \widehat{\omega} & v \\ 0 & 0 \end{bmatrix}^{\vee} \doteq \begin{bmatrix} v \\ \omega \end{bmatrix} \quad \in \mathbb{R}^6, \qquad \begin{bmatrix} v \\ \omega \end{bmatrix}^{\wedge} \doteq \begin{bmatrix} \widehat{\omega} & v \\ 0 & 0 \end{bmatrix} \quad \in \mathbb{R}^{4\times 4}.$$

In the twist coordinates ξ, we will refer to v as the *linear velocity* and ω as the *angular velocity*, which indicates that they are related to either the translational or the rotational part of the full motion. Let us now consider a special case of equation (2.20) when the twist $\widehat{\xi}$ is a constant matrix

$$\dot{g}(t) = \widehat{\xi}g(t).$$

We have again a time-invariant linear ordinary differential equation, which can be integrated to give

$$g(t) = e^{\widehat{\xi}t}g(0).$$

Assuming the initial condition $g(0) = I$, we may conclude that

$$\boxed{g(t) = e^{\widehat{\xi}t},}$$

where the twist exponential is

$$e^{\widehat{\xi}t} = I + \widehat{\xi}t + \frac{(\widehat{\xi}t)^2}{2!} + \cdots + \frac{(\widehat{\xi}t)^n}{n!} + \cdots. \tag{2.21}$$

By Rodrigues' formula (2.16) introduced in the previous section and additional properties of the matrix exponential, the following relationship can be established:

$$\boxed{e^{\widehat{\xi}} = \begin{bmatrix} e^{\widehat{\omega}} & \frac{(I-e^{\widehat{\omega}})\widehat{\omega}v+\omega\omega^T v}{\|\omega\|} \\ 0 & 1 \end{bmatrix}, \quad \text{if} \quad \omega \neq 0.} \tag{2.22}$$

If $\omega = 0$, the exponential is simply $e^{\widehat{\xi}} = \begin{bmatrix} I & v \\ 0 & 1 \end{bmatrix}$. It is clear from the above expression that the exponential of $\widehat{\xi}$ is indeed a rigid-body transformation matrix in $SE(3)$. Therefore, the exponential map defines a transformation from the space $se(3)$ to $SE(3)$,

$$\exp : se(3) \to SE(3); \quad \widehat{\xi} \mapsto e^{\widehat{\xi}},$$

and the twist $\widehat{\xi} \in se(3)$ is also called the *exponential coordinates* for $SE(3)$, as is $\widehat{\omega} \in so(3)$ for $SO(3)$.

Can every rigid-body motion $g \in SE(3)$ be represented in such an exponential form? The answer is yes and is formulated in the following theorem.

Theorem 2.11 (Logarithm of $SE(3)$). *For any $g \in SE(3)$, there exist (not necessarily unique) twist coordinates $\xi = (v, \omega)$ such that $g = \exp(\widehat{\xi})$. We denote the inverse to the exponential map by $\widehat{\xi} = \log(g)$.*

Proof. The proof is constructive. Suppose $g = (R, T)$. From Theorem 2.8, for the rotation matrix $R \in SO(3)$ we can always find ω such that $e^{\widehat{\omega}} = R$. If $R \neq I$, i.e. $\|\omega\| \neq 0$, from equation (2.22) we can solve for $v \in \mathbb{R}^3$ from the linear equation

$$\boxed{\frac{(I-e^{\widehat{\omega}})\widehat{\omega}v+\omega\omega^T v}{\|\omega\|} = T.} \tag{2.23}$$

If $R = I$, then $\|\omega\| = 0$. In this case, we may simply choose $\omega = 0, v = T$. $\square$

As with the exponential coordinates for rotation matrices, the exponential map from $se(3)$ to $SE(3)$ is not one-to-one. There are usually infinitely many exponential coordinates (or twists) that correspond to every $g \in SE(3)$.

Remark 2.12. *As in the rotation case, the linear structure of $se(3)$, together with the closure under the Lie bracket operation*

$$[\widehat{\xi}_1, \widehat{\xi}_2] = \widehat{\xi}_1\widehat{\xi}_2 - \widehat{\xi}_2\widehat{\xi}_1 = \begin{bmatrix} \widehat{\omega_1 \times \omega_2} & \omega_1 \times v_2 - \omega_2 \times v_1 \\ 0 & 0 \end{bmatrix} \in se(3),$$

makes $se(3)$ the Lie algebra for $SE(3)$. The two rigid-body motions $g_1 = e^{\widehat{\xi}_1}$ and $g_2 = e^{\widehat{\xi}_2}$ commute with each other, $g_1 g_2 = g_2 g_1$, if and only if $[\widehat{\xi}_1, \widehat{\xi}_2] = 0$.

Example 2.13 (Screw motions). Screw motions are a specific class of rigid-body motions. A screw motion consists of rotation about an axis in space through an angle of θ radians, followed by translation along the same axis by an amount d. Define the *pitch* of the screw motion to be the ratio of translation to rotation, $h = d/\theta$ (assuming $\theta \neq 0$). If we choose a point $\boldsymbol{X}$ on the axis and $\omega \in \mathbb{R}^3$ to be a unit vector specifying the direction, the axis is the set of points $L = \{\boldsymbol{X} + \mu\omega\}$. Then the rigid-body motion given by the screw is

$$g = \begin{bmatrix} e^{\widehat{\omega}\theta} & (I - e^{\widehat{\omega}\theta})\boldsymbol{X} + h\theta\omega \\ 0 & 1 \end{bmatrix} \in SE(3). \tag{2.24}$$

The set of all screw motions along the same axis forms a subgroup $SO(2) \times \mathbb{R}$ of $SE(3)$, which we will encounter occasionally in later chapters. A statement, also known as *Chasles' theorem*, reveals a rather remarkable fact that any rigid-body motion can be realized as a rotation around a particular axis in space and translation along that axis. ■

2.5 Coordinate and velocity transformations

In this book, we often need to know how the coordinates of a point and its velocity change as the camera moves. This is because it is usually more convenient to choose the camera frame as the reference frame and describe both camera motion and 3-D points relative to it. Since the camera may be moving, we need to know how to transform quantities such as coordinates and velocities from one camera frame to another. In particular, we want to know how to correctly express the location and velocity of a point with respect to a moving camera. Here we introduce a convention that we will be using for the rest of this book.

Rules of coordinate transformations

The time $t \in \mathbb{R}$ will typically be used to index camera motion. Even in the discrete case in which a few snapshots are given, we will take t to be the index of the camera position and the corresponding image. Therefore, we will use $g(t) = (R(t), T(t)) \in SE(3)$ or

$$g(t) = \begin{bmatrix} R(t) & T(t) \\ 0 & 1 \end{bmatrix} \in SE(3)$$

to denote the relative displacement between some fixed world frame W and the camera frame C at time $t \in \mathbb{R}$. Here we will ignore the subscript "cw" from the notation $g_{cw}(t)$ as long as it is clear from the context. By default, we assume $g(0) = I$, i.e. at time $t = 0$ the camera frame coincides with the world frame. So if the coordinates of a point $p \in \mathbb{E}^3$ relative to the world frame are $\boldsymbol{X}_0 = \boldsymbol{X}(0)$, its coordinates relative to the camera at time t are given by

$$\boxed{\boldsymbol{X}(t) = R(t)\boldsymbol{X}_0 + T(t),} \tag{2.25}$$

or in the homogeneous representation,

$$\boldsymbol{X}(t) = g(t)\boldsymbol{X}_0. \tag{2.26}$$

If the camera is at locations $g(t_1), g(t_2), \ldots, g(t_m)$ at times $t_1, t_2, \ldots, t_m$, respectively, then the coordinates of the point p are given as $\boldsymbol{X}(t_i) = g(t_i)\boldsymbol{X}_0, i =$

$1, 2, \ldots, m$, correspondingly. If it is only the position, not the time, that matters, we will often use g_i as a shorthand for $g(t_i)$, and similarly R_i for $R(t_i)$, T_i for $T(t_i)$, and $\boldsymbol{X}_i$ for $\boldsymbol{X}(t_i)$. We hence have

$$\boxed{\boldsymbol{X}_i = R_i \boldsymbol{X}_0 + T_i.} \tag{2.27}$$

When the starting time is not $t = 0$, the relative motion between the camera at time t_2 and time t_1 will be denoted by $g(t_2, t_1) \in SE(3)$. Then we have the following relationship between coordinates of the same point p at different times:

$$\boldsymbol{X}(t_2) = g(t_2, t_1)\boldsymbol{X}(t_1), \quad \forall t_1, t_2 \in \mathbb{R}.$$

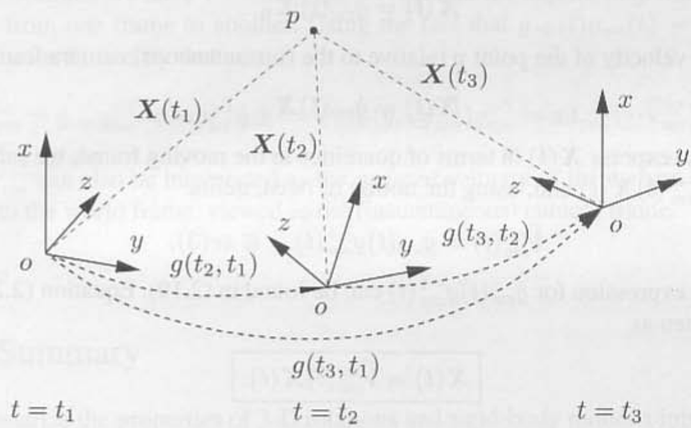

Figure 2.6. Composition of rigid-body motions. $\boldsymbol{X}(t_1), \boldsymbol{X}(t_2), \boldsymbol{X}(t_3)$ are the coordinates of the point p with respect to the three camera frames at time $t = t_1, t_2, t_3$, respectively.

Now consider a third position of the camera at $t = t_3 \in \mathbb{R}$, as shown in Figure 2.6. The relative motion between the camera at t_3 and t_2 is $g(t_3, t_2)$, and that between t_3 and t_1 is $g(t_3, t_1)$. We then have the following relationship among the coordinates:

$$\boldsymbol{X}(t_3) = g(t_3, t_2)\boldsymbol{X}(t_2) = g(t_3, t_2)g(t_2, t_1)\boldsymbol{X}(t_1).$$

Comparing this with the direct relationship between the coordinates at t_3 and t_1,

$$\boldsymbol{X}(t_3) = g(t_3, t_1)\boldsymbol{X}(t_1),$$

we see that the following *composition rule* for consecutive motions must hold:

$$g(t_3, t_1) = g(t_3, t_2)g(t_2, t_1).$$

The composition rule describes the coordinates $\boldsymbol{X}$ of the point p relative to any camera position if they are known with respect to a particular one. The same composition rule implies the *rule of inverse*

$$g^{-1}(t_2, t_1) = g(t_1, t_2),$$

since $g(t_2,t_1)g(t_1,t_2) = g(t_2,t_2) = I$. In cases in which time is of no physical meaning, we often use g_{ij} as a shorthand for $g(t_i,t_j)$. The above composition rules then become (in the homogeneous representation)

$$\boxed{\boldsymbol{X}_i = g_{ij}\boldsymbol{X}_j, \quad g_{ik} = g_{ij}g_{jk}, \quad g_{ij}^{-1} = g_{ji}.} \tag{2.28}$$

Rules of velocity transformation

Having understood the transformation of coordinates, we now study how it affects velocity. We know that the coordinates $\boldsymbol{X}(t)$ of a point $p \in \mathbb{E}^3$ relative to a moving camera are a function of time t:

$$\boldsymbol{X}(t) = g_{cw}(t)\boldsymbol{X}_0.$$

Then the velocity of the point p relative to the (instantaneous) camera frame is

$$\dot{\boldsymbol{X}}(t) = \dot{g}_{cw}(t)\boldsymbol{X}_0. \tag{2.29}$$

In order to express $\dot{\boldsymbol{X}}(t)$ in terms of quantities in the moving frame, we substitute $\boldsymbol{X}_0$ by $g_{cw}^{-1}(t)\boldsymbol{X}(t)$ and, using the notion of twist, define

$$\widehat{V}_{cw}^{c}(t) = \dot{g}_{cw}(t)g_{cw}^{-1}(t) \quad \in se(3), \tag{2.30}$$

where an expression for $\dot{g}_{cw}(t)g_{cw}^{-1}(t)$ can be found in (2.19). Equation (2.29) can be rewritten as

$$\boxed{\dot{\boldsymbol{X}}(t) = \widehat{V}_{cw}^{c}(t)\boldsymbol{X}(t).} \tag{2.31}$$

Since $\widehat{V}_{cw}^{c}(t)$ is of the form

$$\widehat{V}_{cw}^{c}(t) = \begin{bmatrix} \widehat{\omega}(t) & v(t) \\ 0 & 0 \end{bmatrix},$$

we can also write the velocity of the point in 3-D coordinates (instead of homogeneous coordinates) as

$$\boxed{\dot{\boldsymbol{X}}(t) = \widehat{\omega}(t)\boldsymbol{X}(t) + v(t).} \tag{2.32}$$

The physical interpretation of the symbol $\widehat{V}_{cw}^{c}$ is the velocity of the world frame moving relative to the camera frame, as viewed in the camera frame, as indicated by the subscript and superscript of $\widehat{V}_{cw}^{c}$. Usually, to clearly specify the physical meaning of a velocity, we need to specify the velocity of which frame is moving relative to which frame, and which frame it is viewed from. If we change the location from which we view the velocity, the expression will change accordingly. For example, suppose that a viewer is in another coordinate frame displaced relative to the camera frame by a rigid-body transformation $g \in SE(3)$. Then the coordinates of the same point p relative to this frame are $\boldsymbol{Y}(t) = g\boldsymbol{X}(t)$. We compute the velocity in the new frame, and obtain

$$\dot{\boldsymbol{Y}}(t) = g\dot{g}_{cw}(t)g_{cw}^{-1}(t)g^{-1}\boldsymbol{Y}(t) = g\widehat{V}_{cw}^{c}g^{-1}\boldsymbol{Y}(t).$$

So the new velocity (or twist) is

$$\widehat{V} = g\widehat{V}^c_{cw}g^{-1}.$$

This is the same physical quantity but viewed from a different vantage point. We see that the two velocities are related through a mapping defined by the relative motion g; in particular,

$$ad_g : se(3) \to se(3); \quad \widehat{\xi} \mapsto g\widehat{\xi}g^{-1}.$$

This is the so-called *adjoint map* on the space $se(3)$. Using this notation in the previous example we have $\widehat{V} = ad_g(\widehat{V}^c_{cw})$. Note that the adjoint map transforms velocity from one frame to another. Using the fact that $g_{cw}(t)g_{wc}(t) = I$, it is straightforward to verify that

$$\widehat{V}^c_{cw} = \dot{g}_{cw}g^{-1}_{cw} = -g^{-1}_{wc}\dot{g}_{wc} = -g_{cw}(\dot{g}_{wc}g^{-1}_{wc})g^{-1}_{cw} = ad_{g_{cw}}(-\widehat{V}^w_{wc}).$$

Hence $\widehat{V}^c_{cw}$ can also be interpreted as the *negated* velocity of the camera moving relative to the world frame, viewed in the (instantaneous) camera frame.

2.6 Summary

We summarize the properties of 3-D rotations and rigid-body motions introduced in this chapter in Table 2.1:

	Rotation $SO(3)$	Rigid-body motion $SE(3)$
Matrix representation	$R : \begin{cases} R^TR = I \\ \det(R) = 1 \end{cases}$	$g = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}$
Coordinates (3-D)	$\boldsymbol{X} = R\boldsymbol{X}_0$	$\boldsymbol{X} = R\boldsymbol{X}_0 + T$
Inverse	$R^{-1} = R^T$	$g^{-1} = \begin{bmatrix} R^T & -R^TT \\ 0 & 1 \end{bmatrix}$
Composition	$R_{ik} = R_{ij}R_{jk}$	$g_{ik} = g_{ij}g_{jk}$
Exp. representation	$R = \exp(\widehat{\omega})$	$g = \exp(\widehat{\xi})$
Velocity	$\dot{\boldsymbol{X}} = \widehat{\omega}\boldsymbol{X}$	$\dot{\boldsymbol{X}} = \widehat{\omega}\boldsymbol{X} + v$
Adjoint map	$\widehat{\omega} \mapsto R\widehat{\omega}R^T$	$\widehat{\xi} \mapsto g\widehat{\xi}g^{-1}$

Table 2.1. Rotation and rigid-body motion in 3-D space.

2.7 Exercises

Exercise 2.1 (Linear vs. nonlinear maps). Suppose $A, B, C, X \in \mathbb{R}^{n\times n}$. Consider the following maps from $\mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ and determine whether they are linear or not. Give a brief proof if true and a counterexample if false:

$$\begin{array}{llll} \text{(a)} & X & \mapsto & AX + XB, \\ \text{(b)} & X & \mapsto & AX + BXC, \\ \text{(c)} & X & \mapsto & AXA - B, \\ \text{(d)} & X & \mapsto & AX + XBX. \end{array}$$

Note: A map $f : \mathbb{R}^n \to \mathbb{R}^m$, $x \mapsto f(x)$, is called linear if $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$ for all $\alpha, \beta \in \mathbb{R}$ and $x, y \in \mathbb{R}^n$.

Exercise 2.2 (Inner product). Show that for any positive definite symmetric matrix $S \in \mathbb{R}^{3\times 3}$, the map $\langle\cdot,\cdot\rangle_S : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ defined as

$$\langle u, v\rangle_S = u^T S v, \quad \forall u, v \in \mathbb{R}^3,$$

is a valid inner product on $\mathbb{R}^3$, according to the definition given in Appendix A.

Exercise 2.3 (Group structure of $SO(3)$). Prove that the space $SO(3)$ satisfies all four axioms in the definition of group (in Appendix A).

Exercise 2.4 (Skew-symmetric matrices). Given any vector $\omega = [\omega_1, \omega_2, \omega_3]^T \in \mathbb{R}^3$, we know that the matrix $\widehat{\omega}$ is skew-symmetric; i.e. $\widehat{\omega}^T = -\widehat{\omega}$. Now for any matrix $A \in \mathbb{R}^{3\times 3}$ with determinant $\det(A) = 1$, show that the following equation holds:

$$A^T \widehat{\omega} A = \widehat{A^{-1}\omega}. \tag{2.33}$$

Then, in particular, if A is a rotation matrix, the above equation holds.
Hint: Both $A^T\widehat{(\cdot)}A$ and $\widehat{A^{-1}(\cdot)}$ are linear maps with ω as the variable. What do you need in order to prove that two linear maps are the same?

Exercise 2.5 Show that a matrix $M \in \mathbb{R}^{3\times 3}$ is skew-symmetric if and only if $u^T M u = 0$ for every $u \in \mathbb{R}^3$.

Exercise 2.6 Consider a 2×2 matrix

$$R_1 = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

What is the determinant of the matrix? Consider another transformation matrix

$$R_2 = \begin{bmatrix} \sin\theta & \cos\theta \\ \cos\theta & -\sin\theta \end{bmatrix}.$$

Is the matrix orthogonal? What is the determinant of the matrix? Is R_2 a 2-D rigid-body transformation? What is the difference between R_1 and R_2 ?

Exercise 2.7 (Rotation as a rigid-body motion). Given a rotation matrix $R \in SO(3)$, its action on a vector v is defined as Rv. Prove that any rotation matrix must preserve both the inner product and cross product of vectors. Hence, a rotation is indeed a rigid-body motion.

Exercise 2.8 Show that for any nonzero vector $u \in \mathbb{R}^3$, the rank of the matrix $\widehat{u}$ is always two. That is, the three row (or column) vectors span a two-dimensional subspace of $\mathbb{R}^3$.

Exercise 2.9 (Range and null space). Recall that given a matrix $A \in \mathbb{R}^{m \times n}$, its *null space* is defined as a subspace of $\mathbb{R}^n$ consisting of all vectors $x \in \mathbb{R}^n$ such that $Ax = 0$. It is usually denoted by null(A). The *range* of the matrix A is defined as a subspace of $\mathbb{R}^m$ consisting of all vectors $y \in \mathbb{R}^m$ such that there exists some $x \in \mathbb{R}^n$ such that $y = Ax$. It is denoted by range(A). In mathematical terms,

$$\begin{aligned} \text{null}(A) &\doteq \{x \in \mathbb{R}^n \mid Ax = 0\}, \\ \text{range}(A) &\doteq \{y \in \mathbb{R}^m \mid \exists x \in \mathbb{R}^n, y = Ax\}. \end{aligned}$$

1. Recall that a set of vectors V is a subspace if for all vectors $x, y \in V$ and scalars $\alpha, \beta \in \mathbb{R}$, $\alpha x + \beta y$ is also a vector in V. Show that both null(A) and range(A) are indeed subspaces.
2. What are null$(\widehat{\omega})$ and range$(\widehat{\omega})$ for a nonzero vector $\omega \in \mathbb{R}^3$? Can you describe intuitively the geometric relationship between these two subspaces in $\mathbb{R}^3$? (Drawing a picture might help.)

Exercise 2.10 (Noncommutativity of rotation matrices). What is the matrix that represents a rotation about the X-axis or the Y-axis by an angle θ? In addition to that

1. Compute the matrix R_1 that is the combination of a rotation about the X-axis by $\pi/3$ followed by a rotation about the Z-axis by $\pi/6$. Verify that the resulting matrix is also a rotation matrix.
2. Compute the matrix R_2 that is the combination of a rotation about the Z-axis by $\pi/6$ followed by a rotation about the X-axis by $\pi/3$. Are R_1 and R_2 the same? Explain why.

Exercise 2.11 Let $R \in SO(3)$ be a rotation matrix generated by rotating about a unit vector ω by θ radians that satisfies $R = \exp(\widehat{\omega}\theta)$. Suppose R is given as

$$R = \begin{bmatrix} 0.1729 & -0.1468 & 0.9739 \\ 0.9739 & 0.1729 & -0.1468 \\ -0.1468 & 0.9739 & 0.1729 \end{bmatrix}.$$

- Use the formulae given in this chapter to compute the rotation axis and the associated angle.
- Use Matlab's function `eig` to compute the eigenvalues and eigenvectors of the above rotation matrix R. What is the eigenvector associated with the unit eigenvalue? Give its form and explain its meaning.

Exercise 2.12 (Properties of rotation matrices). Let $R \in SO(3)$ be a rotation matrix generated by rotating about a unit vector $\omega \in \mathbb{R}^3$ by θ radians. That is, $R = e^{\widehat{\omega}\theta}$.

1. What are the eigenvalues and eigenvectors of $\widehat{\omega}$? You may use a computer software (e.g., Matlab) and try some examples first. If you cannot find a brute-force way to do it, can you use results from Exercise 2.4 to simplify the problem first (hint: use the relationship between trace, determinant and eigenvalues).
2. Show that the eigenvalues of R are $1, e^{i\theta}, e^{-i\theta}$, where $i = \sqrt{-1}$ is the imaginary unit. What is the eigenvector that corresponds to the eigenvalue 1? This actually gives another proof for $\det(e^{\widehat{\omega}\theta}) = 1 \cdot e^{i\theta} \cdot e^{-i\theta} = +1$, not -1.

Exercise 2.13 (Adjoint transformation on twist). Given a rigid-body motion g and a twist $\widehat{\xi}$,

$$g = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \in SE(3), \quad \widehat{\xi} = \begin{bmatrix} \widehat{\omega} & v \\ 0 & 0 \end{bmatrix} \in se(3),$$

show that $g\widehat{\xi}g^{-1}$ is still a twist. Describe what the corresponding ω and v terms have become in the new twist. The adjoint map is sort of a generalization to $R\widehat{\omega}R^T = \widehat{R\omega}$.

Exercise 2.14 Suppose that there are three camera frames C_0, C_1, C_2 and the coordinate transformation from frame C_0 to frame C_1 is (R_1, T_1) and from C_0 to C_2 is (R_2, T_2). What is the relative coordinate transformation from C_1 to C_2 then? What about from C_2 to C_1? (Express these transformations in terms of R_1, T_1 and R_2, T_2 only.)

2.A Quaternions and Euler angles for rotations

For the sake of completeness, we introduce a few conventional schemes to parameterize rotation matrices, either globally or locally, that are often used in numerical computations for rotation matrices. However, we encourage the reader to use the exponential parameterizations described in this chapter.

Quaternions

We know that the set of complex numbers $\mathbb{C}$ can be simply defined as $\mathbb{C} = \mathbb{R} + \mathbb{R}i$ with $i^2 = -1$. Quaternions generalize complex numbers in a similar fashion. The set of quaternions, denoted by $\mathbb{H}$, is defined as

$$\mathbb{H} = \mathbb{C} + \mathbb{C}j, \quad \text{with } j^2 = -1 \text{ and } i \cdot j = -j \cdot i. \tag{2.34}$$

So, an element of $\mathbb{H}$ is of the form

$$q = q_0 + q_1 i + (q_2 + iq_3)j = q_0 + q_1 i + q_2 j + q_3 ij, \quad q_0, q_1, q_2, q_3 \in \mathbb{R}. \tag{2.35}$$

For simplicity of notation, in the literature ij is sometimes denoted by k. In general, the *multiplication* of any two quaternions is similar to the multiplication of two complex numbers, except that the multiplication of i and j is *anticommutative*: $ij = -ji$. We can also similarly define the concept of *conjugation* for a quaternion:

$$q = q_0 + q_1 i + q_2 j + q_3 ij \quad \Rightarrow \quad \bar{q} = q_0 - q_1 i - q_2 j - q_3 ij. \tag{2.36}$$

It is immediate to check that

$$q\bar{q} = q_0^2 + q_1^2 + q_2^2 + q_3^2. \tag{2.37}$$

Thus, $q\bar{q}$ is simply the square of the norm $\|q\|$ of q as a four-dimensional vector in $\mathbb{R}^4$. For a nonzero $q \in \mathbb{H}$, i.e. $\|q\| \neq 0$, we can further define its *inverse* to be

$$q^{-1} = \frac{\bar{q}}{\|q\|^2}. \tag{2.38}$$

The multiplication and inverse rules defined above in fact endow the space $\mathbb{R}^4$ with an algebraic structure of a *skew field*. In fact $\mathbb{H}$ is called a *Hamiltonian field*, or *quaternion field*.

One important usage of the quaternion field $\mathbb{H}$ is that we can in fact embed the rotation group $SO(3)$ into it. To see this, let us focus on a special subgroup of $\mathbb{H}$, the *unit quaternions*

$$\mathbb{S}^3 = \left\{ q \in \mathbb{H} \mid \|q\|^2 = q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1 \right\}. \tag{2.39}$$

The set of all unit quaternions is simply the unit sphere in $\mathbb{R}^4$. To show that $\mathbb{S}^3$ is indeed a group, we simply need to prove that it is closed under the multiplication and inverse of quaternions; i.e. the multiplication of two unit quaternions is still a unit quaternion, and so is the inverse of a unit quaternion. We leave this simple fact as an exercise to the reader.

Given a rotation matrix $R = e^{\widehat{\omega} t}$ with $\|\omega\| = 1$ and $t \in \mathbb{R}$, we can associate with it a unit quaternion as follows:

$$q(R) = \cos(t/2) + \sin(t/2)(\omega_1 i + \omega_2 j + \omega_3 ij) \in \mathbb{S}^3. \tag{2.40}$$

One may verify that this association preserves the group structure between $SO(3)$ and $\mathbb{S}^3$:

$$q(R^{-1}) = q^{-1}(R), \; q(R_1 R_2) = q(R_1)q(R_2), \; \forall R, R_1, R_2 \in SO(3). \tag{2.41}$$

Further study can show that this association is also *genuine*; i.e. for different rotation matrices, the associated unit quaternions are also different. In the opposite direction, given a unit quaternion $q = q_0 + q_1 i + q_2 j + q_3 ij \in \mathbb{S}^3$, we can use the following formulae to find the corresponding rotation matrix $R(q) = e^{\widehat{\omega} t}$:

$$t = 2\arccos(q_0), \quad \omega_m = \begin{cases} q_m / \sin(t/2), & t \neq 0, \\ 0, & t = 0, \end{cases} \quad m = 1, 2, 3. \tag{2.42}$$

However, one must notice that according to the above formula, there are two unit quaternions that correspond to the same rotation matrix: $R(q) = R(-q)$, as shown in Figure 2.7. Therefore, topologically, $\mathbb{S}^3$ is a double covering of $SO(3)$. So $SO(3)$ is topologically the same as a three-dimensional projective plane $\mathbb{RP}^3$.

Compared to the exponential coordinates for rotation matrices that we studied in this chapter, in using unit quaternions $\mathbb{S}^3$ to represent rotation matrices $SO(3)$, we have less redundancy: there are only two unit quaternions that corresponding to the same rotation matrix, while there are infinitely many for exponential coordinates (all related by periodicity). Furthermore, such a representation for rotation matrices is smooth, and there is no singularity, as opposed to the representation by *Euler angles*, which we will now introduce.

Euler angles

Unit quaternions can be viewed as a way to *globally* parameterize rotation matrices: the parameterization works for every rotation matrix practically the same way. On the other hand, the Euler angles to be introduced below fall into the cat-

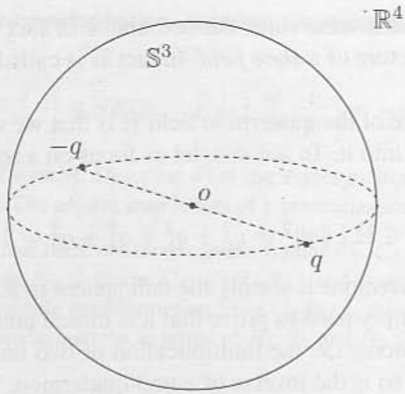

Figure 2.7. Antipodal unit quaternions q and $-q$ on the unit sphere $\mathbb{S}^3 \subset \mathbb{R}^4$ correspond to the same rotation matrix.

egory of *local* parameterizations. This kind of parameterization is good for only a portion of $SO(3)$, but not for the entire space.

In the space of skew-symmetric matrices $so(3)$, pick a *basis* $(\widehat{\omega}_1, \widehat{\omega}_2, \widehat{\omega}_3)$, i.e. the three vectors $\omega_1, \omega_2, \omega_3$ are linearly independent. Define a mapping (a parameterization) from $\mathbb{R}^3$ to $SO(3)$ as

$$\alpha : (\alpha_1, \alpha_2, \alpha_3) \mapsto \exp(\alpha \widehat{\omega}_1 + \alpha_2 \widehat{\omega}_2 + \alpha_3 \widehat{\omega}_3).$$

The coordinates $(\alpha_1, \alpha_2, \alpha_3)$ are called the *Lie-Cartan coordinates of the first kind* relative to the basis $(\widehat{\omega}_1, \widehat{\omega}_2, \widehat{\omega}_3)$. Another way to parameterize the group $SO(3)$ using the same basis is to define another mapping from $\mathbb{R}^3$ to $SO(3)$ by

$$\beta : (\beta_1, \beta_2, \beta_3) \mapsto \exp(\beta_1 \widehat{\omega}_1) \exp(\beta_2 \widehat{\omega}_2) \exp(\beta_3 \widehat{\omega}_3).$$

The coordinates $(\beta_1, \beta_2, \beta_3)$ are called the *Lie-Cartan coordinates of the second kind.*

In the special case in which we choose $\omega_1, \omega_2, \omega_3$ to be the principal axes Z, Y, X, respectively, i.e.

$$\omega_1 = [0, 0, 1]^T \doteq \boldsymbol{z}, \quad \omega_2 = [0, 1, 0]^T \doteq \boldsymbol{y}, \quad \omega_3 = [1, 0, 0]^T \doteq \boldsymbol{x},$$

the Lie-Cartan coordinates of the second kind then coincide with the well-known *ZYX Euler angles* parameterization, and $(\beta_1, \beta_2, \beta_3)$ are the corresponding Euler angles, called "yaw," "pitch," and "roll." The rotation matrix is defined by

$$R(\beta_1, \beta_2, \beta_3) = \exp(\beta_1 \widehat{\boldsymbol{z}}) \exp(\beta_2 \widehat{\boldsymbol{y}}) \exp(\beta_3 \widehat{\boldsymbol{x}}). \tag{2.43}$$

More precisely, $R(\beta_1, \beta_2, \beta_3)$ is the multiplication of the three rotation matrices

$$\begin{bmatrix} \cos(\beta_1) & -\sin(\beta_1) & 0 \\ \sin(\beta_1) & \cos(\beta_1) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} \cos(\beta_2) & 0 & \sin(\beta_2) \\ 0 & 1 & 0 \\ -\sin(\beta_2) & 0 & \cos(\beta_2) \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\beta_3) & -\sin(\beta_3) \\ 0 & \sin(\beta_3) & \cos(\beta_3) \end{bmatrix}.$$

Similarly, we can define the YZX Euler angles and the ZYZ Euler angles. There are instances for which this representation becomes singular, and for certain rotation matrices, their corresponding Euler angles cannot be uniquely determined. For example, when $\beta_2 = -\pi/2$ the ZYX Euler angles become singular. The presence of such singularities is expected because of the topology of the space $SO(3)$. Globally, $SO(3)$ is like a sphere in $\mathbb{R}^4$, as we know from the quaternions, and therefore any attempt to find a global (three-dimensional) coordinate chart is doomed to failure.

Historical notes

The study of rigid-body motion mostly relies on the tools of linear algebra. Elements of screw theory can be tracked back to the early 1800s in the work of Chasles and Poinsot. The use of the exponential coordinates for rigid-body motions was introduced by [Brockett, 1984], and related formulations can be found in the classical work of [Ball, 1900] and others. The use of quaternions in robot vision was introduced by [Broida and Chellappa, 1986b, Horn, 1987]. The presentation of the material in this chapter follows the development in [Murray et al., 1993]. More details on the study of rigid-body motions as well as further references can also be found there.