Robotics Lecture Notes **Spring 2018**

Lecture 3: Rigid Body Transformation and Homogeneous Coordinates

Lecturer: Nilanjan Chakraborty *Scribes:*

Note: *LaTeX template courtesy of UC Berkeley EECS dept.*

3.1 Introduction

Thus far, we have studied in detail the group of rotations and its representation. As we know, pure rotation is only one type of rigid body transformation. A rigid body may also have pure translation or combined translation and rotation. In this lecture we will look at general rigid body motion which consist of both translation and rotation. In particular, we will introduce the notion of homogeneous coordinates and the set of transformation matrices, $SE(3)$, called the Special Euclidean group of dimension 3. We will also look at the exponential coordinates representation of $SE(3)$.

3.2 Rigid Body Transformation

Rigid Body: A rigid body is a collection of particles such that the distance between any two particles remain fixed, regardless of any motion of the body or forces exerted by the body.

Let $\mathbf{p}$, $\mathbf{q}$ be any two points on the body O. Then, as O moves

$$\|\mathbf{p}(t) - \mathbf{q}(t)\| = \|\mathbf{p}(\mathbf{0}) - \mathbf{q}(\mathbf{0})\| = \text{constant}$$

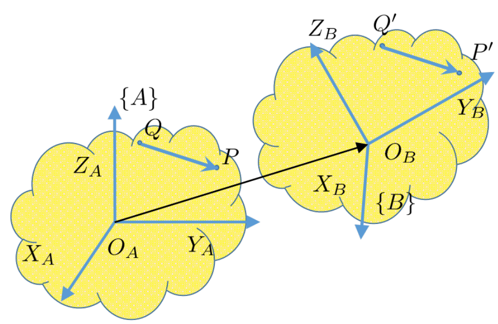


Figure 3.1: Figure showing rigid body motion (transformation) of an object. The points P and Q are transformed to points P' and Q' respectively.

Rigid Motion: A rigid motion of an object is a continuous movement of the particles in the object such that the distance between any two particles remains fixed at all times.

Rigid Displacement Net movement of a rigid body from one location to another. In general a rigid displacement consists of both translation and rotation.

Rigid Transformation: Rigid motion can be expressed as a function called rigid transformation. Let $g(t)$ be a function that maps initial co-ordinates of a point on the body to the co-ordinates at time t, i.e.,

$$g : O \to \mathbb{R}^3 \quad \text{(ignoring the time t)},$$

where O represents the object. Let P and Q be two points in the object O with position vectors $\mathbf{p}$ and $\mathbf{q}$ respectively (see Figure 3.1). In Figure 3.1, the original configuration of the rigid body is determined by the reference frame $\{A\}$, with respect to a global frame (assume that it coincides with the frame $\{A\}$ for simplicity). After the rigid body motion, the configuration of the object is given by frame $\{B\}$. The points P' and Q' corresponds to the points P and Q after the rigid body motion (or transformation). Let $\mathbf{p}'$ and $\mathbf{q}'$ be the position of points P' and Q' respectively, in the frame $\{A\}$. Define the vector $\mathbf{v} = \mathbf{p} - \mathbf{q}$. For every rigid body transformation g, there is a transformation g_* that transforms vectors on the rigid body, and g_* is defined by

$$g_*(\mathbf{v}) = g(\mathbf{p}) - g(\mathbf{q})$$

The above formula can be easily derived by noticing that $\mathbf{v}' = g_*\ (\mathbf{v}) = \mathbf{p}' - \mathbf{q}' = g(\mathbf{p}) - g(\mathbf{q})$. We will now define a rigid body transformation more formally.

A mapping $g : \mathbb{R}^3 \to \mathbb{R}^3$ is a rigid body transformation if it satisfies the following properties:

1. Length is preserved: $\|g(\mathbf{p}) - g(\mathbf{q})\| = \|\mathbf{p} - \mathbf{q}\|, \forall, \mathbf{p}, \mathbf{q} \in \mathbb{R}^3$.
2. The cross product is preserved: $g_*(\mathbf{v} \times \mathbf{w}) = g_*(\mathbf{v}) \times g_*(\mathbf{w})\ \forall \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$.

Important consequences of rigid body transformations are as follows:

1. Inner product of vectors is preserved, i.e.,

$$\mathbf{v}_1^T \mathbf{v}_2 = g_*(\mathbf{v}_1)^T g_*(\mathbf{v}_2)$$

2. Orthogonal vectors are transformed into orthogonal vectors.
3. Orthonormal coordinate frames are transformed to orthonormal coordinate frames (in particular, right handed frames are transformed to right handed frames)

In the last class, we saw that the configuration of a rigid body can be described by a tuple $(\mathbf{p}, \mathbf{R})$ where $\mathbf{p}$ is the position vector of the origin of a frame $\{B\}$ attached to the body (that moves with the body) and $\mathbf{R}$ is the rotation matrix of $\{B\}$ with respect to the frame $\{A\}$ (see Figure 3.1). Thus, we can see that rigid body transform can also be represented by the tuple $(\mathbf{p}, \mathbf{R})$. We will write the rotation matrix using the two notations that mean the same thing.

${}^A_B\mathbf{R} := \mathbf{R}_{ab} \Longleftarrow$ Rotation matrix of frame $\{B\}$ in frame $\{A\}$.

The set of all configurations of a rigid body (or the set of all rigid body transformations) can be defined as,

$$SE(3) = \{(\mathbf{p}, \mathbf{R}) : \mathbf{p}\ \in\ \mathbb{R}^3, \mathbf{R} \in SO(3)\}$$

Note that SE(3) serves as both a specification of configuration of a rigid body as well as transformation for taking points from one co-ordinate frame to another. Thus for any point $\mathbf{q}$, written as $\mathbf{q}_b$ in frame $\{B\}$ and $\mathbf{q}_a$ in frame $\{A\}$, the following holds (see 3.2,

$$\mathbf{q}_a = \mathbf{p}_{ab} + \mathbf{R}_{ab}\mathbf{q}_b \tag{3.1}$$

For a rigid body transformation one can write the function g that transforms a point as,

$$g(\mathbf{q}) := \mathbf{p} + \mathbf{R}\mathbf{q}$$

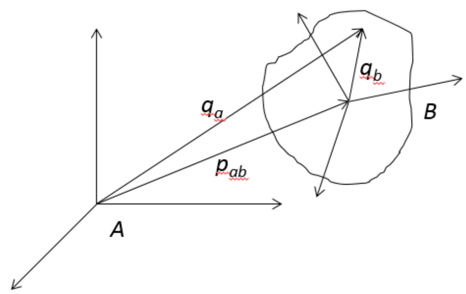


Figure 3.2: Figure depicting the subscript notation

where we drop the reference frames from the notation for convenience. If we want to specify the frames in the notation we will write the above as,

$$\mathbf{q}_a = g_{ab}(\mathbf{q}_b) = \mathbf{p}_{ab} + \mathbf{R}\mathbf{q}_b$$

The mapping for free vectors induced by g, i.e., g_* can be defined as

$$g_*(\mathbf{v}) = \mathbf{R}\mathbf{v}$$

The equation above can be derived as follows: Let $\mathbf{v}_b = \mathbf{s}_b - \mathbf{r}_b$

$$g_*(\mathbf{v}_b) = g_{ab}(\mathbf{s}_b) - g_{ab}(\mathbf{r}_b) \tag{3.2}$$
$$= \mathbf{p}_{ab} + \mathbf{R}_{ab}\mathbf{s}_b - \mathbf{p}_{ab} - \mathbf{R}_{ab}\mathbf{r}_b \tag{3.3}$$
$$= \mathbf{R}_{ab}(\mathbf{s}_b - \mathbf{r}_b) = \mathbf{R}_{ab}\mathbf{v}_b \tag{3.4}$$

Thus, a rigid body transform of a *free vector is equivalent to rotating a vector.* Please note that it is of utmost importance to distinguish between a point and a vector although they are given by 3 coordinates. Conceptually they are different and the *rigid body transformation on them is different.*

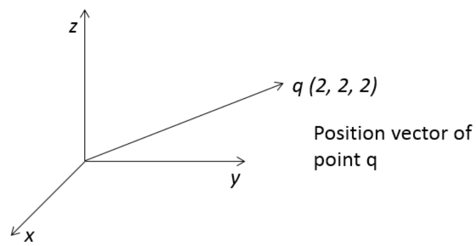
1. Although points and vectors are represented by 3 types of numbers they are conceptually different entities.
2. Vectors are sometimes called free vectors since they are not attached to a body.
3. Vectors have magnitude as well as direction.

We will now introduce the formal techniques where these distinction between points and vectors are captured explicitly.

3.3 Homogeneous Representation

The difference between points and vectors are made explicit in homogeneous representations. The homogeneous coordinates of a point $\mathbf{q}$ are $\bar{\mathbf{q}} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ 1 \end{bmatrix}$, and for a vector $\mathbf{v}$ are $\bar{\mathbf{v}} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ 0 \end{bmatrix}$.

We will use an 'overhead bar 'for homogeneous representation. Thus, for a point $\mathbf{q}$ its homogeneous representation is $\bar{\mathbf{q}}$. For a vector $\mathbf{v}$ its homogeneous representation is $\bar{\mathbf{v}}$. For a rigid body transformation g its homogeneous representation is $\bar{g}$.



Example: Let us consider a point Q with coordinates (2,2,2). In our conventional rotation the point Q is represented by a vector $\begin{bmatrix}2\\2\\2\end{bmatrix}$ The position vector of point Q is also represented by $\begin{bmatrix}2\\2\\2\end{bmatrix}$. However, in homogeneous representation the point Q is represented by $\begin{bmatrix}2\\2\\2\\1\end{bmatrix}$ and its position vector is represented by $\begin{bmatrix}2\\2\\2\\0\end{bmatrix}$

Note: One fundamental reason for using homogeneous coordinates is that it captures the difference between points and vectors. Any field of study that studies transformation of objects use homogeneous representation (since point and vectors associated with a body may transform differently).

Computer Graphics : Homogeneous representation is used since objects needs to be moved rigidly as well as expanded and controlled.

Computer Vision : The objects are transformed to an image by a transformation (or function) called perspective transformation. So here also homogeneous transformation is used.

Rules of syntax in Homogeneous representation :

(1) Sum and difference of vectors are vectors.

(2) Sum of two points is meaningless.

(3) Difference of two points is a vector.

(4) Sum of a vector and a point is a point.

3.3.1 Homogeneous representation of Rigid Body Transformation

We will now rewrite the rigid body transformation given by Equation (3.1) such that the points $\mathbf{q}_a$ and $\mathbf{q}_b$ are written in homogeneous coordinates $\bar{\mathbf{q}}_a$ and $\bar{\mathbf{q}}_b$ respectively.

$$\bar{\mathbf{q}}_a = \begin{bmatrix}\mathbf{q_a}\\1\end{bmatrix} = \begin{bmatrix}\mathbf{R}_{ab} & \mathbf{p}_{ab}\\0 & 1\end{bmatrix} = \begin{bmatrix}q_b\\1\end{bmatrix} = \bar{g}_{ab}\bar{\mathbf{q}}_b$$

Note that the above system of equations is obtained by a very simple algebraic manipulation. We need to add an equation $1 = 1$ to Equation (3.1) and then write the system of equations in matrix form. Thus, the transformation between two reference frames $\{A\}$ and $\{B\}$ can be written as,

$$\bar{g}_{ab} = {}^{A}_{B}\,\mathbf{T} = \begin{bmatrix}\mathbf{R}_{ab} & \mathbf{p}_{ab}\\0 & 1\end{bmatrix}.$$

This is a 4 x 4 matrix representation of a rigid body transformation which is also known as homogeneous representation of a rigid body transformation. Here p_{ab} is read as the position of origin of $\{B\}$ in $\{A\}$, R_{ab} is read as rotation matrix of frame $\{B\}$ in frame $\{A\}$. Note that, throughout this class, we will use g, $\bar{g}$, and $\mathbf{T}$ be represent a rigid body transformation.

__Note__: For a general transformation of a body (not necessarily a rigid transformation) that involves scaling and perspective transformation, we can write the 4×4 transformation matrix as

$$\bar{\mathbf{q}}_a = \begin{bmatrix} \mathbf{R}_{ab} & \mathbf{p}_{ab} \\ \mathbf{s} & c \end{bmatrix} \tag{3.5}$$

where c is a scalar with $c > 1$ representing dilation of the body and $c < 1$ representing contraction and $\mathbf{s}$ is a vector that is required for perspective transformation.

The set of rigid body transformations form a group called SE(3), special Euclidean group of dimension 3. Homogeneous transformation of a vector, $\mathbf{v}$ is as follows:

$$\bar{g}_* \bar{\mathbf{v}} = \bar{g}(\bar{\mathbf{s}}) - \bar{g}(\bar{\mathbf{r}}) = \begin{bmatrix} \mathbf{R} & \mathbf{p} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ 0 \end{bmatrix} = \mathbf{R}\mathbf{v}$$

Thus homogeneous transformation of vectors can be done by multiplying with the same transform matrix (as that used for points). So in this representation we do not need to worry (when writing a program) whether we are transforming points or vectors. Transformation is equivalent to multiplying by a matrix.

Remark 3.1 *Technically, with the 4×4 transformation matrix, we are representing the rigid body transformation as a linear mapping instead of an affine mapping as given in Equation* (3.1).

3.3.2 Composition of Transformation Matrices

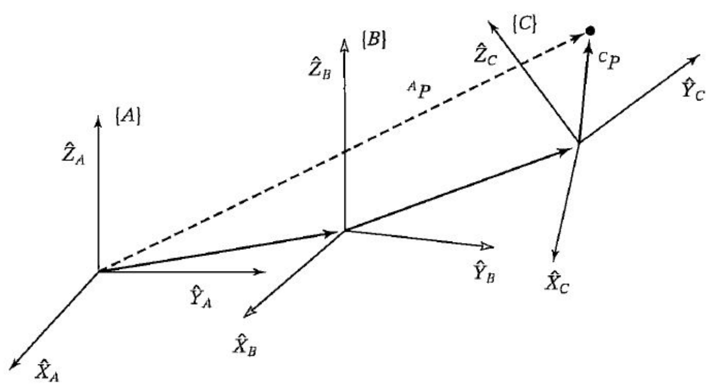


Figure 3.3: Compound transformations

Assume that there are three frames. Frame $\{C\}$ is known relative to frame $\{B\}$, and frame $\{B\}$ is known relative to frame $\{A\}$. We can transform $^C\mathbf{P}$ into $^B\mathbf{P}$ as[1]

$$^B\mathbf{P} = {}^B_C\mathbf{T} \cdot {}^C\mathbf{P}$$

then we can transform $^B\mathbf{P}$ into $^W\mathbf{P}$ as

$$^W\mathbf{P} = {}^W_B\mathbf{T} \cdot {}^B\mathbf{P}$$

[1]For brevity, we are omitting the Q from the subscript, i.e., we are writing $^C\mathbf{P}$ as the position vector of point Q in reference frame $\{C\}$ instead of $^C\mathbf{P}_Q$.

Combining two equations, we get the result

$$^{W}\mathbf{P} = {}^{W}_{B}\mathbf{T} \cdot {}^{B}_{C}\mathbf{T} \cdot {}^{C}\mathbf{P}$$

from which we could define

$$^{W}_{C}\mathbf{T} = {}^{W}_{B}\mathbf{T} \cdot {}^{B}_{C}\mathbf{T}$$

We can obtain the expression for ${}^{W}_{C}\mathbf{T}$ as

$$^{W}_{C}\mathbf{T} = \left[\begin{array}{c|c} {}^{W}_{B}\mathbf{R} \cdot {}^{B}_{C}\mathbf{R} & {}^{W}_{B}\mathbf{R} \cdot {}^{B}\mathbf{P}_{C} + {}^{W}\mathbf{P}_{B} \\ \hline 0 \quad 0 \quad 0 & 1 \end{array}\right] \tag{3.6}$$

If there are N frames $1, 2, 3, \ldots, N$ and we are given ${}^{N}\mathbf{P}_{Q}$. We can compute

$$^{1}\mathbf{P}_{Q} = {}^{1}_{2}\mathbf{T} \cdot {}^{2}_{3}\mathbf{T} \cdots {}^{N-1}_{N}\mathbf{T} \cdot {}^{N}\mathbf{P}_{Q}$$

$$^{1}_{N}\mathbf{T} = {}^{1}_{2}\mathbf{T} \cdot {}^{2}_{3}\mathbf{T} \cdots {}^{N-1}_{N}\mathbf{T}$$

Just by matrix multiplication we can perform any number of transformation.
Inverse of a transform

$$^{W}_{B}\mathbf{T}^{-1} = {}^{B}_{W}\mathbf{T} = \left[\begin{array}{c|c} {}^{W}_{B}\mathbf{R}^{T} & -{}^{W}_{B}\mathbf{R}^{T} \cdot {}^{W}\mathbf{P}_{BO} \\ \hline 0 \quad 0 \quad 0 & 1 \end{array}\right] \tag{3.7}$$

3.3.3 Example Use of Transformation Matrices

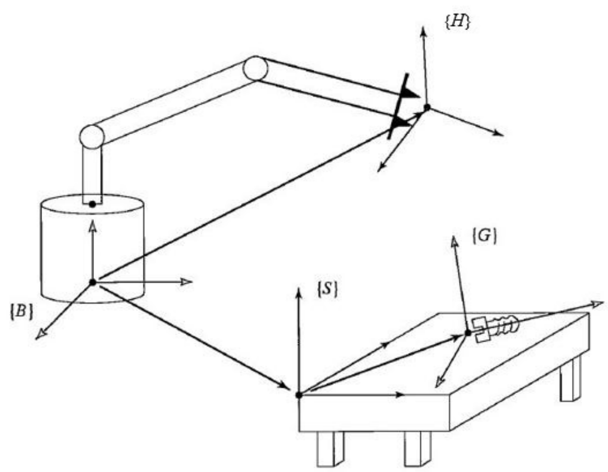


Figure 3.4: Manipulator reaching for a bolt

Figure 3.4 shows a manipulator that has to pick up a bolt lying on a table. The transformation matrix ${}^{B}_{S}\mathbf{T}$ of the table frame with respect to the manipulator base frame is known. The transformation matrix ${}^{S}_{G}\mathbf{T}$ of the bolt frame with respect to the table frame is known. The transformation matrix of the hand frame $\{H\}$ with respect to the base frame $\{B\}$ can be computed using the joint encoder data and the direct kinematics equations of the manipulator. For controlling the manipulator hand to reach the bolt, we have to compute ${}^{H}_{G}\mathbf{T}$ repeatedly. Thus our goal is to do the following: Given quantities ${}^{B}_{H}\mathbf{T}$, ${}^{S}_{G}\mathbf{T}$, ${}^{B}_{S}\mathbf{T}$, we need to find ${}^{H}_{G}\mathbf{T}$. The solution can be found as follows:

$$^{B}_{G}\mathbf{T} = {}^{B}_{G}\mathbf{T} \cdot {}^{S}_{G}\mathbf{T} = {}^{B}_{H}\mathbf{T} \cdot {}^{H}_{G}\mathbf{T}$$

$$\therefore \quad {}^{H}_{G}\mathbf{T} = {}^{B}_{H}\mathbf{T}^{-1} \cdot {}^{B}_{S}\mathbf{T} \cdot {}^{S}_{G}\mathbf{T}$$

3.4 Exponential Coordinates for the Group of Rigid Body Transformations, $SE(3)$

In this section, we will study the exponential coordinates representation of the group of rigid body transformations. The concepts will be analogous to the concepts used in exponential coordinate representation of rigid body rotation. Analogous to the skew symmetric matrix $\hat{\omega}$, we can define a 4×4 matrix $\hat{\xi}$, where

$$\begin{aligned} \hat{\xi} &= \begin{bmatrix} \hat{\omega} & v \\ \mathbf{0} & 0 \end{bmatrix} \quad \text{for combined translation and rotation,} \\ &= \begin{bmatrix} \mathbf{0} & v \\ \mathbf{0} & 0 \end{bmatrix} \quad \text{for pure translation.} \end{aligned} \tag{3.8}$$

where $\omega \in \mathbb{R}^3$ and $v \in \mathbb{R}^3$. The set of all 4×4 matrices of the above form is defined as the space $se(3)$. More formally,

$$se(3) = \left\{(v, \hat{\omega}) : v \in \mathbb{R}^3, \hat{\omega} \in so(3)\right\} \tag{3.9}$$

Each element of $se(3)$ is known as a *twist*. The *twist coordinates* of a twist $\hat{\xi} = \begin{bmatrix} \hat{\omega} & v \\ \mathbf{0} & 0 \end{bmatrix}$ is $\xi = \begin{bmatrix} v \\ \omega \end{bmatrix} \in \mathbb{R}^6$.

Remark 3.2 *Although the notation of $\hat{\xi}$ is analogous to $\hat{\omega}$, $\hat{\xi}$ is not a skew-symmetric matrix.*

Analogous to the the cassse of pure rotations, there is an exponential map that takes an element of $se(3)$ and converts it into an element of $SE(3)$.

Lemma 3.3 *For any $\hat{\xi} \in se(3)$ and $\theta \in \mathbb{R}$, we have*

$$g(v, \omega, \theta) = e^{\hat{\xi}\theta} \equiv exp(\hat{\xi}, \theta) \tag{3.10}$$

More precisely, for $\hat{\xi} = \begin{bmatrix} \hat{\omega} & v \\ \mathbf{0} & 0 \end{bmatrix}$ and $\xi = \begin{bmatrix} v \\ \omega \end{bmatrix}$

$$\begin{aligned} e^{\hat{\xi}\theta} &= \begin{bmatrix} e^{\hat{\omega}\theta} & (I - e^{\hat{\omega}\theta})(\omega \times v) + \omega\omega^T v\theta \\ \mathbf{0} & 1 \end{bmatrix} \in SE(3) \quad \text{for } \omega \neq \mathbf{0} \\ &= \begin{bmatrix} I & v\theta \\ \mathbf{0} & 1 \end{bmatrix} \in SE(3) \quad \text{for } \omega = \mathbf{0} \end{aligned} \tag{3.11}$$

Furthermore, given an element of $SE(3)$, we can obtain the corrersponding twist and angle θ.

Lemma 3.4 *For any $g \in SE(3)$, there exists $\hat{\xi} \in se(3)$ and $\theta \in \mathbb{R}$ such that $g = e^{\hat{\xi}\theta}$. More precisely, let $g = (R, p)$ with $R \in so(3)$ and $p \in \mathbb{R}^3$ be given.*
Then, if the motion is pure translation

$$\omega = 0; \quad \theta = \|p\|, \quad v = \frac{p}{\|p\|}; \tag{3.12}$$

If the motion is combined translation and rotation, we can compute ω and θ from $e^{\hat{\omega}\theta} = R$ and we can compute v by solving the following equation:

$$((I - e^{\hat{\omega}\theta})\hat{\omega} + \theta\omega\omega^T)v = p \tag{3.13}$$

Note that the matrix $(I - e^{\hat{\omega}\theta})\hat{\omega} + \theta\omega\omega^T$ is always invertible.