Annual Reviews in Control 36 (2012) 198–219

Contents lists available at SciVerse ScienceDirect

Annual Reviews in Control

journal homepage: www.elsevier.com/locate/arcontrol

Spacecraft attitude determination and control: Quaternion based method

Y. Yang *

The Office of Research, US Nuclear Regulatory Commission, 21 Church Street, Rockville, MD 20850, United States

ARTICLE INFO

Article history:
Received 25 December 2011
Accepted 27 December 2011
Available online 18 October 2012

Keywords:
Reduced quaternion model
Spacecraft
Attitude determination
Control

ABSTRACT

In this review, we discuss in detail the quaternion based methods for spacecraft attitude determination and control. We summarize some recent developments on this research area. We start with some brief but complete discussions on the theory of quaternion which will be sufficient for the discussion in the remaining part of the review. We review the progress of quaternion based attitude determination which has been well recognized and achieved great success by using Newton's method. We also present a different and more elegant treatment on an analytic solution to Wahba's problem. For quaternion based control system design, we focus on some recently developed reduced quaternion models which use only vector component of the quaternion in the state space models. We discuss some new design method that has the following features: (a) it has an analytic solution of LQR, and (b) the designed system reduces disturbance effect, global stabilizes the nonlinear spacecraft system, and is robust to the modeling uncertainty. The presentation of the review is self-complete. It includes all the background information that is needed to understand the development involving the system modeling, the attitude determination, and the attitude control system design methods.

1. Introduction

Spacecraft attitude determination and control is an important part for a spacecraft to achieve its designed mission. As for today, many spacecrafts have been successfully launched, and most of them have performed well as they were designed. Many research papers have been published to address the attitude determination and control design problems. Several text books are available for students to learn the technology and for engineers to use as references. The most popular spacecraft models for attitude determination algorithms and control design methods are the Euler angle models and the quaternion models. The Euler angle models have been proved very efficient because the linearized Eular angle models are controllable, and all standard linear control system design methods are directly applicable. The drawbacks related to the Euler angle methods are: (a) the designs based on linearized models may not globally stabilize the original nonlinear spacecraft, i.e., the design may not work when the attitude of the spacecraft is far away from the point where the linearization is performed; (b) the models depend on the rotational sequences; and (c) for any rotational sequence, there is a singular point where the model is not applicable. On the other hand, for quaternion models, people have found controllers that can globally stabilize nonlinear spacecraft systems, the models do not depend on rotational sequences, and there is no singular point. The main problem with the quaternion model based control system design is that the linearized quaternion model is not controllable, therefore most published design methods heavily rely on Lyapunov functions for the nonlinear system, and there is not a systematic way to get the Lyapunov functions. Moreover, the Lyapunov function based designs focus on the closed-loop system stability but pay little attention to the closed-loop system performance. Recently, in a series of papers, we proposed some reduced quaternion models which lead to some controllable linearized spacecraft models. Therefore, all standard linear system theory can be directly applied to the analysis and design of the spacecraft control systems. We show that the designed control system is not only optimal for the linearized system, but also globally stabilize the original nonlinear system. Clearly, the reduced quaternion models do not depend on rotational sequences. Due to the special structure of the linearized spacecraft model, some most important design methods, such as LQR design and robust pole assignment design are very simple, enjoy the analytical solution, have direct connection to the performance indexes such as settling time, rising time, and percentage of overshoot. All these features are attractive for high quality control system designs.

Our main purpose in this review is to discuss the details of these recently developed, purely quaternion based approaches for the spacecraft attitude determination and control. We start by discussing all the necessary materials needed for the quaternion based approaches to the spacecraft attitude determination and control. We spend some effort to describe the rotational sequences and the quaternion which provide readers some necessary background that will be used in the following sections. We then introduce the

* Tel.: +1 301 251 7480; fax: +1 301 330 8449.
E-mail address: yaguang.yang@verizon.net

1367-5788/$ - see front matter Published by Elsevier Ltd.
http://dx.doi.org/10.1016/j.arcontrol.2012.09.003

controllable spacecraft models for some mostly desired attitudes for spacecrafts, the inertial pointing attitude and the nadir pointing attitude. To construct a feedback control system, we need to extract quaternion information from vector measurements; we therefore discuss the quaternion based attitude determination method. With all these preparations, we will discuss the control system design methods based on quaternion models and quaternion attitude information. We will discuss some most important design methods, such as LQR design and robust pole assignment design for the quaternion based models, using the quaternion attitude and the angular rate information as the feedback. We will close the discussion by giving our views on the future directions.

2. Rotatinal sequences and quaternion

Based on the missions of a spacecraft, the attitude of the spacecraft represented by the body frame should be aligned with some desired frame. Spacecraft attitude determination is to provide the information of the distance between the spacecraft body frame and the desired frame. This distance can be represented by a single rotation if quaternion is used or a series of rotations if Euler angles are used. In the latter case, the sequence of the rotations is very important. Every rotation is around certain rotational axis for some angle. To determine the desired frame which depends on the spacecraft position and the current time, GPS signals may be used to determine the spacecraft current position and the time. The mostly used time in aerospace engineering is the universal time (UT) (Vallado, 2004). The time and position can be used to obtain the ephemeris astronomical direction information, such as some star directions, the Sun direction, the Earth direction, the Earth magnet field direction, observed from the spacecraft position at the current time and represented in the desired frame. The body frame information can be obtained by the measurements about these directions from the spacecraft onboard instruments. When the body frame is perfectly aligned with the desired frame, the calculated ephemeris star directions, the Sun direction, the Earth direction, and the Earth magnet field direction at the given time should be identical or very close to the measurements from spacecraft instruments. When the body frame is significantly different from the desired frame, the measured astronomical directions are significantly different from the ephemeris astronomical directions at the given time. These differences can be characterized by a rotation or a series rotations, each rotates some angle around certain rotational axis, thereby estimating the distance between the spacecraft body frame and the desired frame. Therefore, mathematical definition on rotation and rotational sequences is one of the most important concepts in spacecraft attitude determination and control. There are many ways to characterize the rotation and rotational sequences. We believe that the quaternion representation is one of the best characterizations, and we will focus our attention on this representation. Our presentation in this section follows the style of Kuipers (1998).

2.1. Some notations and identities

Throughout this review, we will use some common notations. For a column vector $\mathbf{x} = [x_1, x_2, \ldots, x_n]^{\mathrm{T}}$, we sometimes write it as $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ to save space. For any two vectors $\mathbf{x}$ and $\mathbf{y}$, we will denote by $\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^{\mathrm{T}}\mathbf{y}$ the dot product of $\mathbf{x}$ and $\mathbf{y}$, by $\mathbf{x} \times \mathbf{y}$ the cross product of $\mathbf{x}$ and $\mathbf{y}$, by $\|\mathbf{x}\|$ the 2-norm of the vector of $\mathbf{x}$. Let $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ be any three dimensional vectors, we will repeatedly use the following identities.

$$(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{b} \cdot \mathbf{c})\mathbf{a}, \tag{1}$$

and

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}, \tag{2}$$

and

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{a} = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{b} = 0. \tag{3}$$

We denote

$$\mathbf{i} = (1, 0, 0), \quad \mathbf{j} = (0, 1, 0), \quad \mathbf{k} = (0, 0, 1) \tag{4}$$

for the standard basis for $\mathbf{R}^3$. Let $\Omega(\mathbf{x})$ be a skew-symmetric matrix function of $\mathbf{x} = [x_1, x_2, x_3]^{\mathrm{T}}$ defined by

$$\Omega(\mathbf{x}) = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix}. \tag{5}$$

The cross product of $\mathbf{x} \times \mathbf{y}$ can then be represented by a matrix multiplication $\Omega(\mathbf{x})\mathbf{y}$, i.e., $\mathbf{x} \times \mathbf{y} = \Omega(\mathbf{x})\mathbf{y}$. We will use $\bar{p}$, $\bar{q}$, and $\bar{r}$ to denote quaternions which will be defined later on.

2.2. Some frequently used frames

Many frames are used in spacecraft related application. This section discusses some most important frames. For more detailed discussion, readers are referred to Vallado (2004).

2.2.1. Body-fixed frame

Probably the first frame comes to our mind is how to represent the axes of flight object, which is often called the body-fixed frame. The origin of the body-fixed frame is the center of mass of an orbital spacecraft. The fixed $\mathbf{X}_b$ (or roll) axis, $\mathbf{Y}_b$ (or pitch) axis, and $\mathbf{Z}_b$ (or yaw) axis define the orientation of the attitude. Let $\mathbf{v}$ be the vector of spacecraft velocity, $\mathbf{r}$ be the vector from the spacecraft to the Earth center. By convention, $\mathbf{X}_b$ is selected along with the velocity direction, $\mathbf{Y}_b$ is in the direction of $\mathbf{r} \times \mathbf{v}$, which is perpendicular to the orbit plane, and $\mathbf{Z}_b$ follows the right-hand rule. However, other choices are possible as long as $\mathbf{X}_b$, $\mathbf{Y}_b$, and $\mathbf{Z}_b$ are perpendicular to each other.

2.2.2. Earth centered inertial frame

The Earth centered inertial (ECI) frame is important because of two reasons. First, the Newton's laws of motion and gravity applied to the spacecraft are defined in inertial frame. Second, many types of satellites are inertial pointing spacecrafts. This frame is defined relative to the rotation axis of the Earth and the plane of the Earth's orbit (the ecliptic plane) about the Sun. The Earth's equator is perpendicular to the rotation axis of the Earth. As the Earth moving along the ecliptic, the equator plane and the ecliptic have two cross points. Two of these cross points are special as the tilt of the Earth's rotational axis is inclined neither away nor towards the Sun (the center of the Sun being in the same plane as the Earth's equator). The ECI frame is defined at one of these equinoxes, the vernal equinox (or March equinox). Because of many less significant (but may not be negligible) factors, such as the precession of the equinoxes, vernal equinox used by aerospace engineers is defined by 2000 coordinates and the true of date (TOD).[1] The $\mathbf{X}_I$ of inertial frame is the direction from the Earth center to the vernal equinox. The $\mathbf{Z}_I$ axis is the Earth rotational axis. The $\mathbf{Y}_I$ follows the right-hand rule.

2.2.3. Local vertical local horizontal frame

The local vertical local horizontal frame is one of the most desired frames for many satellites because its $\mathbf{Z}_{lvlh}$ direction is always pointing to the center of the Earth (nadir pointing), which is a desired feature of many satellites. The origin of the local vertical local horizontal frame is the center of mass of an orbital spacecraft. The

[1] For the rigorous and precise definition, please read (Vallado, 2004).

$\mathbf{X}_{lvlh}$ direction is along the spacecraft velocity direction and perpendicular to $\mathbf{Z}_{lvlh}$, and $\mathbf{Y}_{lvlh}$ is perpendicular to the orbit plan and follows the right-hand rule.

2.2.4. South east zenith (SEZ) frame

The south east zenith frame is useful for ground stations to track a spacecraft. The location of the tracking instrument is the origin. $\mathbf{X}_{SEZ}$ is the direction pointing to the south, $\mathbf{Y}_{SEZ}$ is the direction pointing to the east, and $\mathbf{Z}_{SEZ}$ is the direction pointing to the zenith. In this system, the azimuth is the angle measured form north, clockwise to the location beneath the object of interest. The elevation is measured from local horizon, positive up to the object of interest.

2.2.5. Earth-centered Earth-fixed frame

Like the Earth centered inertial (ECI) frame, the Earth-centered Earth-fixed (ECEF) frame is the Earth-based frame. The ECI frame is independent from the motion and the rotation of the Earth. However, it may not be convenient in some case as observatories on the ground rotate with the Earth. The center of ECEF frame is the center of the Earth. Using the convention adopted at the International Meridian Conference in Washington, DC, 1884, the primary meridian for the Earth is the meridian that the Royal Observatory at Greenwich lies on. The $\mathbf{X}_{ECEF}$ is the direction from the center of the Earth pointing to the cross point of the primary meridian and equator. The $\mathbf{Z}_{ECEF}$ is the direction from the center of the Earth pointing to the north pole. The $\mathbf{Y}_{ECEF}$ is the direction that follows the right-hand rule. The ECEF frame is sometimes called International Terrestrial Reference Frame (ITRF). Because of the plate tectonic motion, the frame may need some adjustment every year for certain applications.

2.3. Rotation sequences and mathematical representations

2.3.1. A fixed point in a rotational frame

As we discussed at the beginning of this section, we determine the spacecraft attitude by locating the astronomical objects in the sky from the spacecraft instruments which gives the directions in the body frame; from the ephemeris information, we know these directions represented in the desired frame. Therefore, we have the information on some fixed (astronomical object) point in a rotational frame when the spacecraft body frame is different from the desired frame. This is equivalent to represent a fixed point in a rotational frame.

Let $(\mathbf{X},\mathbf{Y},\mathbf{Z})$ be the axes of a frame (see Fig. 1 where $\mathbf{Z}$-axis points out of the paper), and $(\mathbf{x},\mathbf{y},\mathbf{z})$ be the axes of another frame which rotates an angle of θ about $\mathbf{Z}$ axis. Let P be a fixed point in $(\mathbf{X},\mathbf{Y})$ plane. Assume that the distance of P from the origin is r, then we can express P in the first frame coordinate as (x_1, y_1, z_1)

$$x_1 = r\cos(\alpha), \quad y_1 = r\sin(\alpha), \quad z_1 = 0; \tag{6}$$

and in the second frame coordinate as (x_2, y_2, z_2)

$$x_2 = r\cos(\alpha-\theta), \quad y_2 = r\sin(\alpha-\theta), \quad z_2 = 0.$$

Thus, in view of (6), we have

$$\begin{aligned} x_2 &= x_1\cos(\theta) + y_1\sin(\theta),\\ y_2 &= y_1\cos(\theta) - x_1\sin(\theta),\\ z_2 &= 0. \end{aligned}$$

We can write this transformation in a matrix form

$$\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}. \tag{7}$$

Similarly, for a fixed point, if the frame rotates about $\mathbf{Y}$ axis for an angle θ, then the transformation can be expressed as

$$\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} \cos(\theta) & 0 & -\sin(\theta) \\ 0 & 1 & 0 \\ \sin(\theta) & 0 & \cos(\theta) \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}. \tag{8}$$

For a fixed point, if the frame rotates about $\mathbf{X}$ axis for an angle θ, then the transformation can be expressed as

$$\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}. \tag{9}$$

Rotational matrices of (7)–(9) are all unitary matrices. By definition, the length of each column of a unitary matrix is one, each column is orthogonal to other columns. Unitary matrices have many useful properties. Let C_1 and C_2 be two unitary matrices and $\mathbf{v}$ be a vector. Some most important properties of the unitary matrix are (see Golub & Van Loan, 1989):

- $\|C_1\mathbf{v}\| = \|C_2\mathbf{v}\| = \|\mathbf{v}\|$, i.e., unitary matrix does not change the vector length.
- C_2C_1 is a unitary matrix. For rotational matrices, it means that the consecutive rotations can be expressed by the product of the rotational matrices, where C_1 is the first rotation and C_2 is the second rotation.
- $C_1^{-1} = C_1^{\mathrm{T}}$, i.e., the inverse of a rotational matrix is simply a transpose of the rotational matrix.

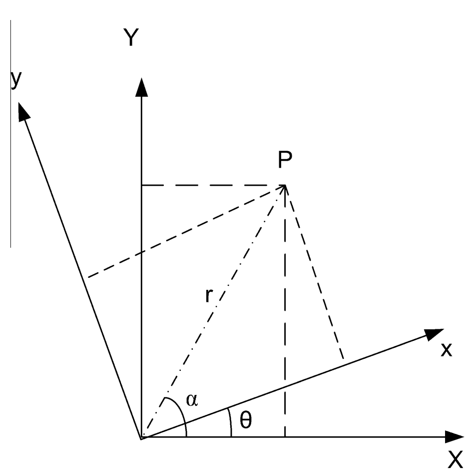

Fig. 1. A fixed point in a rotational frame.

2.3.2. A rotational point in a fixed frame

When analyzing relationship of frames, we sometimes need to represent a rotational point in a fixed frame. Let P_1 be a point obtained by rotating P an angle of θ around $\mathbf{Z}$ axis (see Fig. 2 where $\mathbf{Z}$-axis points out of the paper). Then P_1 can be expressed as

$$x_2 = r\cos(\alpha+\theta), \quad y_2 = r\sin(\alpha+\theta), \quad z_2 = 0.$$

Thus, in view of (6), we have

$$\begin{aligned} x_2 &= x_1\cos(\theta) - y_1\sin(\theta),\\ y_2 &= y_1\cos(\theta) + x_1\sin(\theta),\\ z_2 &= 0. \end{aligned}$$

We can write this transformation in a matrix form

$$\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}. \tag{10}$$

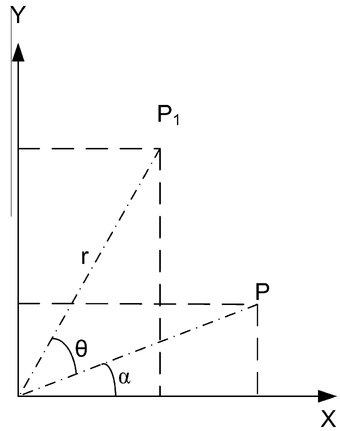

Fig. 2. A rotational point in a fixed frame.

Similarly, for a fixed point, if P rotates about **Y** axis for an angle θ, then the transformation can be expressed

$$\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}. \tag{11}$$

For a fixed point, if P rotates about **X** axis for an angle θ, then the transformation can be expressed

$$\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}. \tag{12}$$

2.3.3. Rotations in three dimensional space

The rotations discussed above are simple rotations in two dimensional space. They are special cases in that the rotational axis is one of the coordinates which is perpendicular to the plane spanned by vectors before and after the rotation. Spacecraft attitude determination and control involve general rotations in three dimensional space. Considering the rotation described in Fig. 3 where we rotate the axis **X** to the axis **x**. A popular method to represent this rotation is to use a series of rotations about coordinate described in the previous subsections, i.e., first we rotate the frame an α angle around $-$**Y** axis, then we rotate the intermediate **x**′ a β angle around the new **Z** axis (**z**′ axis). The α and β angles are the so-called Euler angles. Therefore, the rotational matrix is given by

$$C = \begin{bmatrix} \cos(\beta) & -\sin(\beta) & 0 \\ \sin(\beta) & \cos(\beta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(\alpha) & 0 & -\sin(\alpha) \\ 0 & 1 & 0 \\ \sin(\alpha) & 0 & \cos(\alpha) \end{bmatrix}$$
$$= \begin{bmatrix} \cos(\beta)\cos(\alpha) & -\sin(\beta) & -\cos(\beta)\sin(\alpha) \\ \sin(\beta)\cos(\alpha) & \cos(\beta) & -\sin(\beta)\sin(\alpha) \\ \sin(\alpha) & 0 & \cos(\alpha) \end{bmatrix} \tag{13a}$$

$$= \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix} \tag{13b}$$

which provides a different explanation of the rotation from **X** axis to **x** axis, i.e., the series of rotations can also be represented by a general rotational matrix (13b). Let

$$\cos(\theta) = \frac{1}{2}(C_{11} + C_{22} + C_{33} - 1), \tag{14}$$

$$\hat{e} = \frac{1}{2\sin(\theta)} \begin{bmatrix} C_{23} - C_{32} \\ C_{31} - C_{13} \\ C_{12} - C_{21} \end{bmatrix} = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix}, \tag{15}$$

$$E = \frac{1}{2\sin(\theta)}(C^{\mathrm{T}} - C) = \begin{bmatrix} 0 & -e_3 & e_2 \\ e_3 & 0 & -e_1 \\ -e_2 & e_1 & 0 \end{bmatrix}, \quad \theta \neq \pm k\pi, \quad k = 0, 1, 2, \ldots \tag{16}$$

the general rotational matrix (13b) can be expressed as

$$C = \cos(\theta)I + (1 - \cos(\theta))\hat{e}\hat{e}^{\mathrm{T}} - \sin(\theta)E. \tag{17}$$

It can be verified that C is a rotational matrix, $\hat{e}$ is the rotational axis, and θ is the rotational angle (Hughes, 1986). C is called the direction cosine matrix.

Actually, there may be infinitely many combinations of rotational axes and rotational angles that can rotate **X** to **x**. Moreover, Fig. 4 and the following analysis show that in general case, the

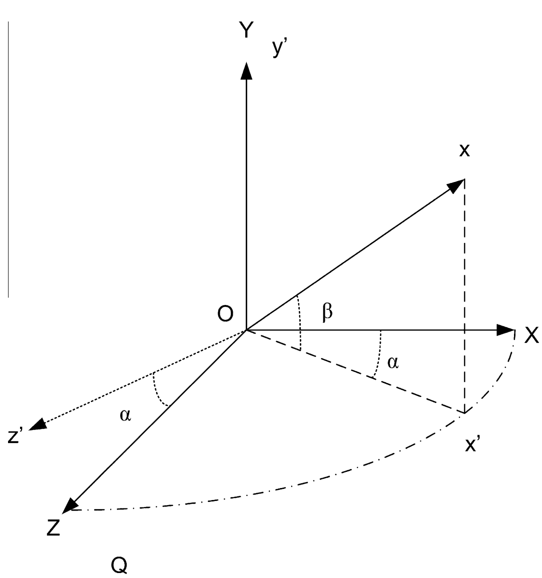

Fig. 3. Rotation in three dimensional space.

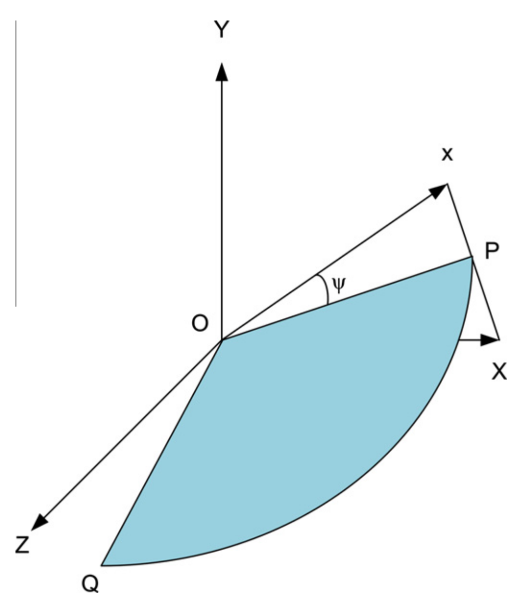

Fig. 4. Plane of all possible rotations axes.

rotational axis of the direction cosine matrix may not be one of the coordinates. Let P be the middle point between $\mathbf{X}$ and $\mathbf{x}$ and ψ be the angle between $O\mathbf{x}$ and OP. Let OQ be the unit vector that is perpendicular to the plane spanned by $\mathbf{X}$ and $\mathbf{x}$ vectors. Obviously, the rotation can be achieved by rotating 2ψ around OQ. Alternatively, another rotation with rotational axis OP and rotational angle π can also rotates $\mathbf{X}$ to $\mathbf{x}$. In fact, we can use any vector on the plane spanned by OP and OQ as the rotational axis and find an appropriate rotational angle which will rotate $\mathbf{X}$ to $\mathbf{x}$. The first rotation we described is sometimes called the minimum-angle rotation, and the second rotation we described is called the maximum-angle rotation.

2.3.4. Rotation from one frame to another frame

In spacecraft attitude determination, we are oftentimes required to find a rotation that brings one frame to another one. This means that we need to find a rotational axis and appropriate rotational angle that rotates one given frame to another given frame. Let S be the middle point of $\mathbf{Y}$ and $\mathbf{y}$, OR be the unit length vector that is perpendicular to the plane spanned by $\mathbf{Y}$ and $\mathbf{y}$. The rotation that brings the frame ($\mathbf{X},\mathbf{Y},\mathbf{Z}$) to ($\mathbf{x},\mathbf{y},\mathbf{z}$) is described in Fig. 5, where the plane OPQ spanned by OP and OQ defines all the rotational axes that can rotate $\mathbf{X}$ to $\mathbf{x}$; the plane OSR spanned by OR and OS defines all the rotational axes that can rotate $\mathbf{Y}$ to $\mathbf{y}$. Therefore, the intersection of these two planes defines the unique rotational axis that can rotate $\mathbf{X}$ to $\mathbf{x}$ and $\mathbf{Y}$ to $\mathbf{y}$ simultaneously. We will provide a rigorous derivation in the next section.

2.3.5. Rate of change of the direction cosine matrix

In spacecraft dynamics modeling and controls, we need to know not only the attitude of the spacecraft, which is represented by the rotation from one frame to another frame, but also the rate of this rotation. The time dependence of the direction cosine matrix A at time t can be expressed by $A(t)$. The time dependence of the direction cosine matrix A at time $t + \Delta t$ can be expressed by

$$A(t+\Delta t) = CA(t),$$

where C is a rotation around $\hat{e}$ with rotational angle $\theta = \Omega\Delta t$, and Ω is the rate of the rotation around the rotational axis. From (17),

$$C = \cos(\Omega\Delta t)I + (1-\cos(\Omega\Delta t))\hat{e}\hat{e}^{\mathrm{T}} - \sin(\Omega\Delta t)E.$$

As $\Delta t \to 0$,

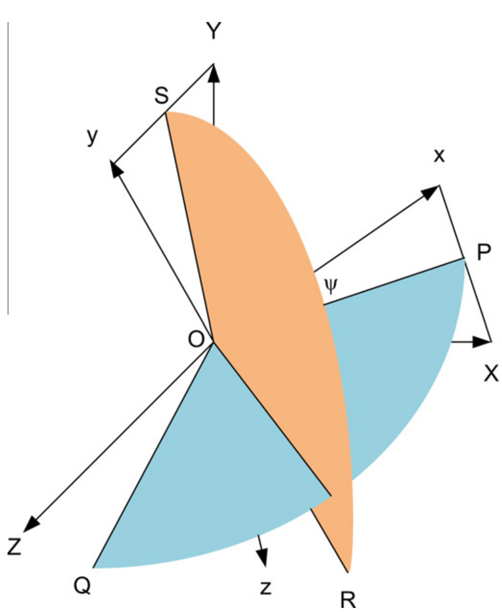

Fig. 5. Rotation from one frame to another.

$$\begin{aligned} C \;\to\; & I - E\Omega\Delta t = I - \Omega(\omega)\Delta t \\ & = I - \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}\Delta t, \end{aligned} \tag{18}$$

where

$$E\Omega = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} = \Omega(\omega).$$

This gives

$$A(t+\Delta t) = (I - \Omega(\omega)\Delta t)A(t),$$

or

$$A(t+\Delta t) - A(t) = -\Omega(\omega)A(t)\Delta t,$$

therefore, we get

$$\frac{dA}{dt} = -\Omega(\omega)A(t). \tag{19}$$

2.3.6. Rate of change of vectors in rotational frame

In spacecraft dynamics modeling and controls, vectors and their rates of changes are oftentimes represented in different frames. For the modeling and control purpose, we need to convert the vectors and their rates of changes represented in different frames into a single frame. Here, we derive the relationship between the time derivatives of an arbitrary vector resolved along a coordinate axes of one system and the derivatives in a different system. Let a' be the vector represented in a reference system and a be the same vector represented in body frame. Then there is a rotational matrix C expressed in (17) such that

$$a = Ca'.$$

The product rule for differentiation gives

$$\left(\frac{da}{dt}\right)\bigg|_b = \frac{dC}{dt}a' + C\left(\frac{da'}{dt}\right)\bigg|_r,$$

where the derivative $\left(\frac{da}{dt}\right)\big|_b$ is represented in the body frame, and the derivative $\left(\frac{da'}{dt}\right)\big|_r$ is represented in the reference frame. Since C is the rotation from reference frame to body frame, $C\left(\frac{da'}{dt}\right)\big|_r = \left(\frac{da'}{dt}\right)\big|_b$. From (19),

$$\begin{aligned} \left(\frac{da}{dt}\right)\bigg|_b &= -\Omega(\omega)Ca' + C\frac{da'}{dt}\bigg|_r = -\Omega(\omega)a + \left(\frac{da'}{dt}\right)\bigg|_b \\ &= -\omega \times a + \left(\frac{da'}{dt}\right)\bigg|_b, \end{aligned} \tag{20}$$

where ω is the rate of the rotation between the reference frame and the body frame.

2.4. Quaternion and its properties

Unlike the Euler angles which represent a rotation by a series of rotations rotating around $\mathbf{X}$, or $\mathbf{Y}$ or $\mathbf{Z}$ axes, quaternion represents a rotation by a rotational angle around a rotational axis, which is not necessarily around $\mathbf{X}$, or $\mathbf{Y}$, or $\mathbf{Z}$ axes. Quaternion was first introduced by the Irish mathematician William Rowan Hamilton in 1843 and applied to mechanics in three-dimensional space. A striking feature of quaternion is that the product of two quaternion is noncommutative, meaning that the product of two quaternions depends on which factor is to the left of the multiplication sign and which factor is to the right. Let the standard basis $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ for the $\mathbf{R}^3$ satisfy the following condition

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1. \tag{21}$$

Let a 4-tuple of real numbers

$$\bar{q} = (q_0, q_1, q_2, q_3), \tag{22}$$

we define a quaternion as the sum of a scalar and a vector

$$\bar{q} = q_0 + \mathbf{i}q_1 + \mathbf{j}q_2 + \mathbf{k}q_3 = q_0 + q, \tag{23}$$

where

$$q = \mathbf{i}q_1 + \mathbf{j}q_2 + \mathbf{k}q_3$$

is called the vector part of the quaternion, and q_0 is called the scalar part of the quaternion. People use (22) and (23) interchangeably if no confusion is introduced. Though in aerospace engineering, we always use a special normalized quaternion $q_0 = \cos\left(\frac{\alpha}{2}\right)$, and $q = \hat{e}\sin\left(\frac{\alpha}{2}\right)$, where $\hat{e}$ is rotational axis, and α is the rotational angle, we will derive some useful properties for the general form of quaternion.

2.4.1. Equality and addition

Let

$$\bar{p} = p_0 + \mathbf{i}p_1 + \mathbf{j}p_2 + \mathbf{k}p_3$$

and

$$\bar{q} = q_0 + \mathbf{i}q_1 + \mathbf{j}q_2 + \mathbf{k}q_3$$

be two quaternions, then the two quaternions are equal if and only if

$$p_0 = q_0, \quad p_1 = q_1, \quad p_2 = q_2, \quad p_3 = q_3.$$

For the special normalized quaternion used in the aerospace engineering, if two quaternions are equal, they have the same rotational angle and the same rotational axis. The sum of the two quaternions is defined as

$$\bar{p} + \bar{q} = (p_0 + q_0) + \mathbf{i}(p_1 + q_1) + \mathbf{j}(p_2 + q_2) + \mathbf{k}(p_3 + q_3).$$

The zero quaternion has scalar part 0 and vector part (0, 0, 0). The negative or an additive inverse of $\bar{q}$ is $-\bar{q}$.

2.4.2. Multiplication and the identity

From (21), we have

$$\mathbf{ij} = \mathbf{k} = -\mathbf{ji}, \quad \mathbf{jk} = \mathbf{i} = -\mathbf{kj}, \quad \mathbf{ki} = \mathbf{j} == -\mathbf{ik}. \tag{24}$$

Let $\bar{p}$ and $\bar{q}$ be defined as before, use (21) and (24), we define the multiplication of two quaternions $\bar{p}$ and $\bar{q}$ by

$$\bar{p} \otimes \bar{q} = p_0 q_0 - p \cdot q + p_0 q + q_0 p + p \times q, \tag{25}$$

with the scalar part $p_0 q_0 - p \cdot q$ and vector part $p_0 q + q_0 p + p \times q$. The quaternion multiplicative identity has scalar part 1 and vector part (0,0,0).

The quaternion multiplication can be used to represent two consecutive rotations. Let $\bar{p}$ and $\bar{q}$ be the two consecutive rotations ($\bar{p}$ represent the first rotation and $\bar{q}$ represent the second rotation). The composed rotation is given by $\bar{r} = \bar{p} \otimes \bar{q}$. The derivation can be found in Wie (1998, pp. 319–320).

2.4.3. Complex conjugate, norm, and inverse

The complex conjugate of quaternion $\bar{q}$ is denoted by

$$\bar{q}^* = q_0 - q = q_0 - \mathbf{i}q_1 - \mathbf{j}q_2 - \mathbf{k}q_3. \tag{26}$$

It is easy to see

$$\bar{q} + \bar{q}^* = (q_0 + q) + (q_0 - q) = 2q_0. \tag{27}$$

Given two quaternions $\bar{p}$ and $\bar{q}$, we have

$$(\bar{p} \otimes \bar{q})^* = \bar{q}^* \otimes \bar{p}^*. \tag{28}$$

The norm of a quaternion is defined as $\|\bar{q}\| = \sqrt{\bar{q}^* \otimes \bar{q}}$. It is also easy to verify that the norm satisfies

$$\|\bar{q}\| = \sqrt{q_0^2 + q_1^2 + q_2^2 + q_3^2}. \tag{29}$$

We define the inverse of a quaternion by

$$\bar{q}^{-1} \otimes \bar{q} = \bar{q} \otimes \bar{q}^{-1} = 1.$$

Pre- and post-multiplying by $\bar{q}^*$ gives

$$\bar{q}^{-1} \otimes \bar{q} \otimes \bar{q}^* = \bar{q}^* \otimes \bar{q} \otimes \bar{q}^{-1} = \bar{q}^*.$$

Since $\bar{q}^* \otimes \bar{q} = \bar{q} \otimes \bar{q}^* = \|\bar{q}\|^2$, we have

$$\bar{q}^{-1} = \frac{\bar{q}^*}{\|\bar{q}\|^2}. \tag{30}$$

For normalized quaternion which satisfies $\|\bar{q}\| = \sqrt{q_0^2 + q_1^2 + q_2^2 + q_3^2} = 1$,

$$\bar{q}^{-1} = \bar{q}^*. \tag{31}$$

Finally, the norm of the product of two quaternions $\bar{p}$ and $\bar{q}$ is the product of the individual norms because

$$\begin{aligned}\|\bar{p} \otimes \bar{q}\|^2 &= (\bar{p} \otimes \bar{q}) \otimes (\bar{p} \otimes \bar{q})^* = \bar{p} \otimes \bar{q} \otimes \bar{q}^* \otimes \bar{p}^* \\ &= \bar{p} \otimes \|q\|^2 \otimes \bar{p}^* = \bar{p} \otimes \bar{p}^* \|q\|^2 = \|p\|^2 \|q\|^2.\end{aligned} \tag{32}$$

2.4.4. Rotation by quaternion operator

Now we are ready to show how to rotate a vector using quaternion operator. For this purpose, we will consider only the normalized quaternion $\bar{q} = q_0 + q = \cos\left(\frac{\alpha}{2}\right) + \hat{e}\sin\left(\frac{\alpha}{2}\right)$, where $\hat{e}$ is the unit length rotational axis and α is the rotational angle. Clearly, quaternion does have the information about the rotational angle and the rotational axis. Similar to rotational matrices, we need the product of quaternions to be able to represent consecutive rotations. Let $\bar{p} = \cos\left(\frac{\alpha}{2}\right) + \hat{e}\sin\left(\frac{\alpha}{2}\right)$ and $\bar{q} = \cos\left(\frac{\beta}{2}\right) + \hat{e}\sin\left(\frac{\beta}{2}\right)$, from (25), we have

$$\begin{aligned}\bar{r} = \bar{p} \otimes \bar{q} &= \left(\cos\left(\frac{\alpha}{2}\right) + \hat{e}\sin\left(\frac{\alpha}{2}\right)\right) \otimes \left(\cos\left(\frac{\beta}{2}\right) + \hat{e}\sin\left(\frac{\beta}{2}\right)\right) \\ &= \cos\left(\frac{\alpha}{2}\right)\cos\left(\frac{\beta}{2}\right) - \hat{e}\sin\left(\frac{\alpha}{2}\right) \cdot \hat{e}\sin\left(\frac{\beta}{2}\right) + \cos\left(\frac{\alpha}{2}\right)\hat{e}\sin\left(\frac{\beta}{2}\right) \\ &\quad + \hat{e}\sin\left(\frac{\alpha}{2}\right)\cos\left(\frac{\beta}{2}\right) + \hat{e}\sin\left(\frac{\alpha}{2}\right) \times \hat{e}\sin\left(\frac{\beta}{2}\right) \\ &= \cos\left(\frac{\alpha}{2}\right)\cos\left(\frac{\beta}{2}\right) - \sin\left(\frac{\alpha}{2}\right)\sin\left(\frac{\beta}{2}\right) \\ &\quad + \hat{e}\left(\sin\left(\frac{\alpha}{2}\right)\cos\left(\frac{\beta}{2}\right) + \cos\left(\frac{\alpha}{2}\right)\sin\left(\frac{\beta}{2}\right)\right) \\ &= \cos\left(\frac{\alpha+\beta}{2}\right) + \hat{e}\sin\left(\frac{\alpha+\beta}{2}\right) = \cos(\gamma) + \hat{e}\sin(\gamma)\end{aligned} \tag{33}$$

This means that the product of two quaternions indeed represents two consecutive rotations. Parallel to the vector rotation using rotational matrix, we expect that a quaternion rotation operator involves multiplication of a quaternion and a vector. Therefore, the multiplication of a quaternion and a vector should be defined. To this end, we consider a vector $\mathbf{v}$ as a pure quaternion in which the scalar part is zero and the vector part is $\mathbf{v}$, i.e., $\bar{v} = 0 + \mathbf{v}$. For the sake of notational simplicity, we use $\bar{v}$ and $\mathbf{v}$ interchangeably for both vector and pure quaternion. From (25), the multiplication of a vector and a quaternion is defined as

$$\bar{q} \otimes \mathbf{v} = (q_0 + q) \otimes (0 + \mathbf{v}) = -q \cdot \mathbf{v} + q_0 \mathbf{v} + q \times \mathbf{v}. \tag{34}$$

We also expect that the quaternion operator will rotate a vector into another vector, or a pure quaternion. Simple evaluation shows that neither $\mathbf{w} = \bar{q} \otimes \mathbf{v}$ nor $\mathbf{w} = \mathbf{v} \otimes \bar{q}$ is necessarily a pure vector. However, using (34) and (1), we have

$$\begin{aligned}\mathbf{w} &= \bar{q}\otimes\mathbf{v}\otimes\bar{q}^* = (q_0+q)\otimes(0+\mathbf{v})\otimes(q_0-q)\\ &= (-q\cdot\mathbf{v}+q_0\mathbf{v}+q\times\mathbf{v})\otimes(q_0-q)\\ &= -q_0(q\cdot\mathbf{v})+q_0(\mathbf{v}\cdot q)+(q\times\mathbf{v})\cdot q+(q\cdot\mathbf{v})q+q_0^2\mathbf{v}\\ &\quad +q_0(q\times\mathbf{v})-q_0(\mathbf{v}\times q)-(q\times\mathbf{v})\times q\\ &= (q\cdot\mathbf{v})q+q_0^2\mathbf{v}+2q_0(q\times\mathbf{v})-(q\cdot q)\mathbf{v}+(\mathbf{v}\cdot q)q\\ &= (2q_0^2-1)\mathbf{v}+2(q\cdot\mathbf{v})q+2q_0(q\times\mathbf{v})\\ &= \left(\cos^2\left(\frac{\alpha}{2}\right)-\sin^2\left(\frac{\alpha}{2}\right)\right)\mathbf{v}+2(q\cdot\mathbf{v})q+2q_0(q\times\mathbf{v}),\end{aligned} \tag{35}$$

which is a vector. In fact, the quaternion operator can be expressed by direction cosine matrix that may be more convenient in some cases. From (35), since

$$2(q_0^2-1)\mathbf{v} = \begin{bmatrix}(2q_0^2-1) & 0 & 0\\ 0 & (2q_0^2-1) & 0\\ 0 & 0 & (2q_0^2-1)\end{bmatrix}\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix},$$

$$2(\mathbf{v}\cdot q)q = \begin{bmatrix}2q_1^2 & 2q_1q_2 & 2q_1q_3\\ 2q_1q_2 & 2q_2^2 & 2q_2q_3\\ 2q_1q_3 & 2q_2q_3 & 2q_3^2\end{bmatrix}\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix},$$

$$2q_0(q\times\mathbf{v}) = \begin{bmatrix}0 & -2q_0q_3 & 2q_0q_2\\ 2q_0q_3 & 0 & -2q_0q_1\\ -2q_0q_2 & 2q_0q_1 & 0\end{bmatrix}\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix},$$

we have

$$\begin{bmatrix}w_1\\w_2\\w_3\end{bmatrix} = \begin{bmatrix}2q_0^2-1+2q_1^2 & 2q_1q_2-2q_0q_3 & 2q_1q_3+2q_0q_2\\ 2q_1q_2+2q_0q_3 & 2q_2^2+2q_0^2-1 & 2q_2q_3-2q_0q_1\\ 2q_1q_3-2q_0q_2 & 2q_2q_3+2q_0q_1 & 2q_3^2+2q_0^2-1\end{bmatrix}\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}. \tag{36}$$

This means that we can use either (35) or (36) for quaternion rotation. We will use them in different applications in the rest of the review. It is worthwhile to note, in view of (35), that (36) defines a general rotational matrix as

$$C = (q_0^2-q^{\mathrm{T}}q)I+2qq^{\mathrm{T}}+2q_0\Omega(q). \tag{37}$$

We now show that $\bar{q}\otimes\mathbf{v}\otimes\bar{q}^*$ is indeed the quaternion operator that rotates $\mathbf{v}$ an α angle around $\hat{e}$. First, it is easy to verify that $\bar{q}\otimes\mathbf{v}\otimes\bar{q}^*$ is linear operator, i.e., for two vectors $\mathbf{a}$, $\mathbf{b}$, and a scalar k, the following relation holds.

$$\bar{q}\otimes(k\mathbf{a}+\mathbf{b})\otimes\bar{q}^* = k\bar{q}\otimes\mathbf{a}\otimes\bar{q}^*+\bar{q}\otimes\mathbf{b}\otimes\bar{q}^*. \tag{38}$$

Then, we decompose vector $\mathbf{v}$ into two components, $\mathbf{v}=\mathbf{v}_q+\mathbf{v}_n$, where $\mathbf{v}_q$ is parallel to q and $\mathbf{v}_n$ is perpendicular to q. We show (a) under quaternion operator $\bar{q}\otimes\mathbf{v}\otimes\bar{q}^*$, the first component $\mathbf{v}_q$ is invariant, and (b) the second component $\mathbf{v}_n$ rotates an angle of α. Since $\mathbf{v}_q=kq$, where $k\leqslant 1$ is a constant, from (38) (34) and (25), using the fact that $\bar{q}$ is a normalized quaternion, we have

$$\begin{aligned}\bar{q}\otimes\mathbf{v}_q\otimes\bar{q}^* &= \bar{q}\otimes(kq)\otimes\bar{q}^* = k\bar{q}\otimes(q)\otimes\bar{q}^*\\ &= k(-q\cdot q+q_0q)\otimes(q_0-q) = kq.\end{aligned}$$

This proves (a). Using the facts that

$$q\cdot\mathbf{v}_n = 0,$$

$$\cos(\alpha) = \cos^2\left(\frac{\alpha}{2}\right)-\sin^2\left(\frac{\alpha}{2}\right),$$

$$\sin(\alpha) = 2\cos\left(\frac{\alpha}{2}\right)\sin\left(\frac{\alpha}{2}\right),$$

$$q_0 = \cos\left(\frac{\alpha}{2}\right),$$

$$\|q\| = \sin\left(\frac{\alpha}{2}\right),$$

$$q\times\mathbf{v}_n = \|q\|\|\mathbf{v}_n\|\sin\left(\frac{\pi}{2}\right)\mathbf{v}_\perp = \|q\|\|\mathbf{v}_n\|\mathbf{v}_\perp,$$

where $\mathbf{v}_\perp$ is a unit length vector perpendicular to both q and $\mathbf{v}_n$, and from (35), we have

$$\begin{aligned}\bar{q}\otimes(\mathbf{v}_n)\otimes\bar{q}^* &= \left(\cos^2\left(\frac{\alpha}{2}\right)-\sin^2\left(\frac{\alpha}{2}\right)\right)\mathbf{v}_n+2(q\cdot\mathbf{v}_n)q+2q_0(q\times\mathbf{v}_n)\\ &= \cos(\alpha)\mathbf{v}_n+2q_0(q\times\mathbf{v}_n)\\ &= \cos(\alpha)\mathbf{v}_n+2\cos\left(\frac{\alpha}{2}\right)(q\times\mathbf{v}_n)\\ &= \cos(\alpha)\mathbf{v}_n+2\cos\left(\frac{\alpha}{2}\right)\sin\left(\frac{\alpha}{2}\right)\|\mathbf{v}_n\|\mathbf{v}_\perp\\ &= \cos(\alpha)\mathbf{v}_n+\sin(\alpha)\|\mathbf{v}_n\|\mathbf{v}_\perp.\end{aligned} \tag{39}$$

Since $\mathbf{v}_n$ and $\|\mathbf{v}_n\|\mathbf{v}_\perp$ have the same length, and they both perpendicular to $\mathbf{v}_q$, (39) indicates that $\bar{q}\otimes(\mathbf{v}_n)\otimes\bar{q}^*$ rotates $\mathbf{v}_n$ an angle of α around axis q. This proves (b).

A fact parallel to the rotational matrix is that $\bar{q}\otimes(\mathbf{v})\otimes\bar{q}^*$ does not change the length of $\mathbf{v}$, which is a direct result of (32) and the fact that $\bar{q}$ is a normalized quaternion.

$$\|\bar{q}\otimes\mathbf{v}\otimes\bar{q}^*\| = \|\bar{q}\|\|\mathbf{v}\|\|\bar{q}\|^* = \|\mathbf{v}\|. \tag{40}$$

Similar to the rotational matrix, the inverse of the operator $\mathbf{w}=\bar{q}\otimes(\mathbf{v})\otimes\bar{q}^*$ on $\mathbf{v}$ is simple and it is given by

$$\begin{aligned}\bar{q}^*\otimes(\mathbf{w})\otimes\bar{q} &= \bar{q}^*\otimes(\bar{q}\otimes(\mathbf{v})\otimes\bar{q}^*)\otimes\bar{q}\\ &= (\bar{q}^*\otimes\bar{q})\otimes\mathbf{v}\otimes(\bar{q}^*\otimes\bar{q}) = \mathbf{v}\end{aligned}$$

which rotates $\mathbf{w}$ an angle of α around $-q$ and brings $\mathbf{w}$ back to $\mathbf{v}$. It is easy to verify that

$$\mathbf{v} = \bar{q}^*\otimes\mathbf{w}\otimes\bar{q} = (2q_0^2-1)\mathbf{w}+2(q\cdot\mathbf{w})q-2q_0(q\times\mathbf{w}). \tag{41}$$

This gives

$$\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix} = \begin{bmatrix}2q_0^2-1+2q_1^2 & 2q_1q_2+2q_0q_3 & 2q_1q_3-2q_0q_2\\ 2q_1q_2-2q_0q_3 & 2q_0^2-1+2q_2^2 & 2q_2q_3+2q_0q_1\\ 2q_1q_3+2q_0q_2 & 2q_2q_3-2q_0q_1 & 2q_0^2-1+2q_3^2\end{bmatrix}\begin{bmatrix}w_1\\w_2\\w_3\end{bmatrix}. \tag{42}$$

It is worthwhile to note, in view of (41), that (42) defines a general rotational matrix as

$$A = (q_0^2-q^{\mathrm{T}}q)I+2qq^{\mathrm{T}}-2q_0\Omega(q). \tag{43}$$

(43) is another form of the rotational matrix (17).

2.4.5. Matrix form of quaternion production

We also find that in some applications, a matrix form of quaternion production is more convenient than the form of (25). Let $\bar{r}=(r_0,r_1,r_2,r_3)$ be the composed quaternion of two consecutive quaternions of $\bar{p}$ and $\bar{q}$, i.e., $\bar{r}=\bar{p}\otimes\bar{q}$. Expanding (25) gives

$$r_0 = p_0q_0-p_1q_1-p_2q_2-p_3q_3 \tag{44a}$$

$$r_1 = p_0q_1+p_1q_0+p_2q_3-p_3q_2 \tag{44b}$$

$$r_2 = p_0q_2-p_1q_3+p_2q_0+p_3q_1 \tag{44c}$$

$$r_3 = p_0q_3+p_1q_2-p_2q_1+p_3q_0. \tag{44d}$$

(44) can be written in matrix form

$$\begin{bmatrix}r_0\\r_1\\r_2\\r_3\end{bmatrix} = \begin{bmatrix}p_0 & -p_1 & -p_2 & -p_3\\ p_1 & p_0 & -p_3 & p_2\\ p_2 & p_3 & p_0 & -p_1\\ p_3 & -p_2 & p_1 & p_0\end{bmatrix}\begin{bmatrix}q_0\\q_1\\q_2\\q_3\end{bmatrix} \tag{45a}$$

$$= \begin{bmatrix}q_0 & -q_1 & -q_2 & -q_3\\ q_1 & q_0 & q_3 & -q_2\\ q_2 & -q_3 & q_0 & q_1\\ q_3 & q_2 & -q_1 & q_0\end{bmatrix}\begin{bmatrix}p_0\\p_1\\p_2\\p_3\end{bmatrix} \tag{45b}$$

2.4.6. Derivative of the quaternion

Let $\bar{q}(t)$ be the quaternion to a reference frame at time $t, \bar{q}(t+\Delta t)$ be the quaternion to the reference frame at $t + \Delta t$, and $\bar{p}(t) = \cos\left(\frac{\Delta\alpha}{2}\right) + \hat{e}(t)\sin\left(\frac{\Delta\alpha}{2}\right)$ be the quaternion that brings $\bar{q}(t)$ to $\bar{q}(t+\Delta t)$, i.e., $\bar{p}(t)$ is an incremental quaternion with rotational axis $\hat{e}(t)$ and rotational angle $\Delta\alpha$. For $\Delta t \to 0, \cos\left(\frac{\Delta\alpha}{2}\right) \to 1$ and $\sin\left(\frac{\Delta\alpha}{2}\right) \to \frac{\Delta\alpha}{2}$, therefore, $\bar{p}(t) \approx 1 + \hat{e}(t)\frac{\Delta\alpha}{2}$. This gives

$$\bar{q}(t+\Delta t) = \bar{q}(t) \otimes \left(1 + \hat{e}(t)\frac{\Delta\alpha}{2}\right),$$

or

$$\bar{q}(t+\Delta t) - \bar{q}(t) = \bar{q}(t) \otimes \left(0 + \hat{e}(t)\frac{\Delta\alpha}{2}\right).$$

Divide Δt at both sides and let $\Delta t \to 0$, we obtain

$$\frac{d\bar{q}}{dt} = \bar{q}(t) \otimes \left(0 + \frac{1}{2}\hat{e}(t)\Omega(t)\right) = \bar{q}(t) \otimes \left(0 + \frac{1}{2}\omega(t)\right),$$

where $\Omega(t) = \lim_{\Delta t \to 0}\frac{\Delta\alpha}{\Delta t}$ is a scalar, and $\omega(t) = \hat{e}(t)\Omega(t)$ is a vector, and $\left(0 + \frac{1}{2}\omega(t)\right) = \frac{1}{2}(0, \omega_1, \omega_2, \omega_3)$ is a quaternion. Using matrix expression (45) for the quaternion product, we obtain

$$\begin{bmatrix} \dot{q}_0 \\ \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 0 & -\omega_1 & -\omega_2 & -\omega_3 \\ \omega_1 & 0 & \omega_3 & -\omega_2 \\ \omega_2 & -\omega_3 & 0 & \omega_1 \\ \omega_3 & \omega_2 & -\omega_1 & 0 \end{bmatrix}\begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & -q_3 & q_2 \\ q_2 & q_3 & q_0 & -q_1 \\ q_3 & -q_2 & q_1 & q_0 \end{bmatrix}\begin{bmatrix} 0 \\ \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix}. \tag{46a}$$

3. Spacecraft dynamics and modeling

The quaternion based models have several advantages over the Euler angle based models. For example, the quaternion based model is uniquely defined because it does not depend on rotational sequence, while a Euler angle based model can be different for different rotational sequences. Therefore, Euler angle based models may be error-prone if different groups of people are working on the same project but use different rotational sequences. In engineering design practice, an agreement has to be reached among different design groups working on the same project. Another attractive feature of quaternion based model is that a full quaternion model does not have any singular point in any rotational sequence. Therefore, quaternion model-based control design methods have been discussed in a number of papers. In Wen and Kreutz-Delgado (1991), Lyapunov function was used to design model-independent control law, model dependant control law, and adaptive control law. In Boskovic, Li, and Mehra (2001) and Wallsgrove and Akella (2005), Lyapunov Functions were used to design control systems under the restriction of control input saturation. Though Lyapunov function is a powerful tool in global stability analysis, obtaining a control law and the associated Lyapunov function for the nonlinear systems is postulated by intuition, as noted in Paielli and Bach (1993). Moreover, most of these designs focus on the global stability and do not pay much attention on the performance of the control system. In Paielli and Bach (1993) and Wie (1998), quaternion based linear error dynamics are adapted to get desired performance for the attitude control system using classical frequency domain methods. However, state space time domain design methods, such as optimal control and pole assignment, are more attractive than the classical frequency domain design methods. In Zhou and Colgren (2005), a linearized state space quaternion model is derived. Unfortunately, the analysis shows that the linearized state space representation of the full quaternion model using all four components of quaternion is uncontrollable. Therefore, pole assignment can only be achieved in some controllable subspace in the linearized state space quaternion model using all four components of quaternion. In addition, the stability of the linearized closed loop system is unknown because an uncontrollable eigenvalue is at the origin of the complex plane. Another restriction in the existing quaternion modeling and controller design methods is that most researches focus on inertial pointing spacecraft without using reaction wheel while many low Earth orbit spacecraft is nadir pointing and use momentum wheel.

In this section, we first describe a controllable quaternion model for inertial pointing spacecraft, the simplest one in many applications. To obtain a controllable quaternion model, only vector component of the quaternion is used in the model. The cost of using only three components of the quaternion in the model is that, similar to the Euler angle representation, the reduced model has a singular point at $\alpha = \pm\pi$, where α is the rotation angle around the rotation axis. However, this singular point is the farthest point to the point where the linearization is carried out. Therefore, the model and designed controller will work well in practice.

We also present a controllable quaternion model for nadir pointing spacecraft with momentum wheel(s). This is a different model from the inertial pointing spacecraft without a momentum wheel discussed in many literatures. This model includes three important features of many low orbit nadir pointing spacecrafts: (a) an additional term for the momentum wheels is incorporated to the nonlinear dynamic equations, (b) the local vertical local horizontal frame is used as the reference frame and the rotation between local vertical local horizontal frame and inertial frame is considered in the model similar to the treatment in Sidi (1997) for the Euler angle based models, and (c) gravity gradient torque, a dominant and predictable disturbance for low orbit spacecrafts, is included to improve the model accuracy.

We will show that by using only vector component of the quaternion, these linearized spacecraft models are fully controllable. Therefore, it is easier to use these reduced models than the full quaternion models in controller design because all modern state space control system design methods can be applied directly. The stability of the designed closed-loop spacecraft system is guaranteed because the linearized control system is fully controllable.

3.1. The general spacecraft system equations

3.1.1. The dynamics equation

Let J be the inertia matrix of a spacecraft defined by

$$J = \begin{bmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & J_{22} & J_{23} \\ J_{31} & J_{32} & J_{33} \end{bmatrix}, \tag{47}$$

$\omega_I = [\omega_{I1}, \omega_{I2}, \omega_{I3}]^T$ be the angular velocity vector of the spacecraft body with respect to the inertial frame, represented in the spacecraft body frame, h_I be the angular momentum vector of the spacecraft about its center of mass represented in the inertial frame, $h = J\omega_I$ be the same vector of h_I but represented in the body frame, M be the external moment acting on the body about its center of mass. Then, from Serway and Jewett (2004), we have

$$M = \left(\frac{dh_I}{dt}\right)\bigg|_b.$$

In view of (20), we have

$$M = \left(\frac{dh_I}{dt}\right)\bigg|_b = \left(\frac{dh}{dt}\right) + \omega_I \times h.$$

This gives

$$\left(\frac{dh}{dt}\right) = J\dot{\omega}_I = -\omega_I \times J\omega_I + M.$$

The external moment M are normally composed of (a) disturbance torques T_d due to gravitational, aerodynamic, solar radiation, and other environmental torques in body frame, and is expressed by

$$T_d = [T_{d1}, T_{d2}, T_{d3}]^{\mathrm{T}}, \tag{48}$$

and (b) the control torque u expressed by

$$u = [u_1, u_2, u_3]^{\mathrm{T}}. \tag{49}$$

Therefore,

$$J\dot{\omega}_I = -\omega_I \times (J\omega_I) + T_d + u = -\Omega(\omega_I)(J\omega_I) + T_d + u, \tag{50}$$

3.1.2. The kinematics equation

Denote the rotational axis of a body frame relative to a reference frame by a unit length vector $\hat{e}$, the rotational angle around the rotational axis by α, the scalar component of the quaternion by $q_0 = \cos\left(\frac{\alpha}{2}\right)$, the vector component of the quaternion by $q = [q_1, q_2, q_3]^{\mathrm{T}} = \hat{e}\sin\left(\frac{\alpha}{2}\right)$, then, the quaternion that represents the rotation of the body frame relative to the reference frame is given by

$$\bar{q} = [q_0, q^{\mathrm{T}}]^{\mathrm{T}} = \left[\cos\left(\frac{\alpha}{2}\right), \hat{e}^{\mathrm{T}}\sin\left(\frac{\alpha}{2}\right)\right]^{\mathrm{T}}. \tag{51}$$

From (46a), the nonlinear spacecraft kinematics equations of motion can be represented by the quaternion (see also Wie, 1998; Wertz, 1978; Sidi, 1997) as follows

$$\begin{cases} \dot{q} = -\frac{1}{2}\omega \times q + \frac{1}{2}q_0\omega, \\ \dot{q}_0 = -\frac{1}{2}\omega^{\mathrm{T}} q. \end{cases} \tag{52}$$

In view of (46a), we have

$$\begin{aligned}\begin{bmatrix} \dot{q}_0 \\ \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{bmatrix} &= \frac{1}{2}\begin{bmatrix} 0 & -\omega_1 & -\omega_2 & -\omega_3 \\ \omega_1 & 0 & \omega_3 & -\omega_2 \\ \omega_2 & -\omega_3 & 0 & \omega_1 \\ \omega_3 & \omega_2 & -\omega_1 & 0 \end{bmatrix}\begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & -q_3 & q_2 \\ q_2 & q_3 & q_0 & -q_1 \\ q_3 & -q_2 & q_1 & q_0 \end{bmatrix}\begin{bmatrix} 0 \\ \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix}. \end{aligned} \tag{53}$$

Using the fact that $q_0 = \sqrt{1 - q_1^2 - q_2^2 - q_3^2}$, we have,

$$\begin{aligned}\begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{bmatrix} &= \frac{1}{2}\begin{bmatrix} q_0 & -q_3 & q_2 \\ q_3 & q_0 & -q_1 \\ -q_2 & q_1 & q_0 \end{bmatrix}\begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} \\ &= \frac{1}{2}Q(q_1, q_2, q_3)\omega = g(q_1, q_2, q_3, \omega). \end{aligned} \tag{54}$$

It is easy to verify

$$\det\begin{pmatrix} q_0 & -q_3 & q_2 \\ q_3 & q_0 & -q_1 \\ -q_2 & q_1 & q_0 \end{pmatrix} = \det(Q(q_1, q_2, q_3)) = \frac{1}{\sqrt{1 - q_1^2 - q_2^2 - q_3^2}}, \tag{55}$$

hence $Q(q_1, q_2, q_3)$ is always a full rank matrix except for $\alpha = \pm\pi$. This means that unless $\alpha = \pm\pi$, the kinematics equation of motion using reduced quaternion representation can be simplified from (53) and (54).

The main advantages of using (54) instead of (53) is as follows: (a) the system dimension is reduced from 7 to 6, yielding a simpler model, (b) the linearized system is controllable, (c) the stability analysis can be directly conducted based on the linearized system (there is no uncontrollable unstable pole, see Zhou & Colgren, 2005), and (d) all closed loop eigenvalues can be assigned to any position by appropriate feedback control law because the linearized system is controllable. The results presented in this section are based on Yang (2010) and Yang (2012).

3.2. The inertial pointing spacecraft model

3.2.1. The nonlinear inertial pointing spacecraft model

The inertial pointing spacecraft is desired in many applications. The inertial pointing spacecraft model is one of the simplest spacecraft models. In this section, we assume that the spacecraft does not have a momentum wheel, therefore, the control torques are either thrusters or magnet torque bars. We will not discuss the actuators in this review. To simplify the model further, we assume that the disturbance torque is negligible. In this case, (50) is reduced to

$$J\dot{\omega}_I = -\omega_I \times (J\omega_I) + u = -\Omega(\omega_I)(J\omega_I) + u. \tag{56}$$

Let $\bar{q}$ be the quaternion that represents the rotation of the body frame relative to the inertial frame, the reduced kinematics equation becomes

$$\begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} q_0 & -q_3 & q_2 \\ q_3 & q_0 & -q_1 \\ -q_2 & q_1 & q_0 \end{bmatrix}\begin{bmatrix} \omega_{I1} \\ \omega_{I2} \\ \omega_{I3} \end{bmatrix} = g(q_1, q_2, q_3, \omega_I). \tag{57}$$

3.2.2. The linearized inertial pointing spacecraft models

We can derive the linearized spacecraft system from (56) and (57) by using the first order Taylor expansion around the stationary point $q_1 = q_2 = q_3 = 0$ and $\omega_I = 0$ as follows:

$$\dot{\omega}_I \approx J^{-1}u,$$

$$\left.\frac{\partial g}{\partial \omega_I}\right|_{\substack{\omega_I \approx 0 \\ q_1 = q_2 = q_3 \approx 0}} \approx \frac{1}{2}I_3,$$

$$\left.\frac{\partial g}{\partial q}\right|_{\substack{\omega_I \approx 0 \\ q_1 = q_2 = q_3 \approx 0}} \approx \frac{1}{2}0_3.$$

Therefore,

$$\begin{bmatrix} \dot{\omega}_I \\ \dot{q} \end{bmatrix} = \begin{bmatrix} 0_3 & 0_3 \\ \frac{1}{2}I_3 & 0_3 \end{bmatrix}\begin{bmatrix} \omega_I \\ q \end{bmatrix} + \begin{bmatrix} J^{-1} \\ 0_3 \end{bmatrix}u = Ax + Bu, \tag{58}$$

where

$$A = \begin{bmatrix} 0_3 & 0_3 \\ \frac{1}{2}I_3 & 0_3 \end{bmatrix}, \quad x = \begin{bmatrix} \omega_I \\ q \end{bmatrix}, \quad B = \begin{bmatrix} J^{-1} \\ 0_3 \end{bmatrix} \tag{59}$$

It is easy to verify that this linearized spacecraft system equation is controllable.

3.3. Nadir pointing momentum biased spacecraft model

3.3.1. The nonlinear nadir pointing spacecraft model

Momentum biased spacecraft is of particularly interesting in practice, and is discussed extensively in Sidi (1997, Chapter 8). For momentum biased spacecraft, a momentum wheel is installed in $\mathbf{Y}_b$ axis which is perpendicular to the orbit plane. Normally, the momentum wheel spins in a constant speed, but it may also be used to generate control torque by changing the speed. Let

$$H = [h_1, h_2, h_3]^{\mathrm{T}} = [0, h_2, 0]^{\mathrm{T}} \tag{60}$$

be the angular momentum of the momentum wheel in the body frame. The spacecraft model (50) is therefore becomes

$$J\dot{\omega}_I = -\omega_I \times (J\omega_I + H) + T_d + u = -\Omega(\omega_I)(J\omega_I + H) + T_d + u, \tag{61}$$

For a nadir pointing spacecraft, the attitude of the spacecraft is represented by the rotation of the spacecraft body frame relative to the local vertical and local horizontal (LVLH) frame. Therefore, we will represent the quaternion and spacecraft body rate in terms of the rotations of the spacecraft body frame relative to the LVLH frame. Let $\omega = [\omega_1, \omega_2, \omega_3]^{\mathrm{T}}$ be the body rate with respect to the LVLH frame represented in the body frame, $\omega_{lvlh} = [0, \omega_0, 0]^{\mathrm{T}}$ be the orbit (and LVLH frame) rate with respect to the inertial frame, represented in the LVLH frame. Let v be the speed of the spacecraft, r be the distance from the spacecraft to the center of the Earth, p be the orbit period, then for circular orbit spacecraft, we have

$$\omega_0 = \frac{v}{r} = \frac{2\pi}{p}. \tag{62}$$

Let A_l^b represent the transformation matrix from the LVLH frame to the spacecraft body frame. Then, ω_I can be expressed by

$$\omega_I = \omega + A_l^b \omega_{lvlh} = \omega + \omega_{lvlh}^b \tag{63}$$

where ω_{lvlh}^b is the rate of the LVLH frame with respect to the inertial frame, represented in the body frame. From (19), $\dot{A}_l^b = -\omega \times A_l^b$, therefore, $\dot{\omega}_I$ is given by

$$\begin{aligned}\dot{\omega}_I &= \dot{\omega} + \dot{A}_l^b \omega_{lvlh} + A_l^b \dot{\omega}_{lvlh} = \dot{\omega} - \omega \times A_l^b \omega_{lvlh} \\ &= \dot{\omega} - \omega \times \omega_{lvlh}^b\end{aligned} \tag{64}$$

where we assumed that $\dot{\omega}_{lvlh}$ is small and can be neglected.[2] Using Eqs. (63) and (64), we can rewrite Eq. (50) as

$$\begin{aligned}J\dot{\omega} &= J(\omega \times \omega_{lvlh}^b) - \omega \times (J\omega) - \omega \times (J\omega_{lvlh}^b) - \omega_{lvlh}^b \times (J\omega) \\ &\quad - \omega_{lvlh}^b \times (J\omega_{lvlh}^b) - \omega \times H - \omega_{lvlh}^b \times H + T_d + u \\ &= f(\omega, \omega_{lvlh}^b, H) + T_d + u.\end{aligned} \tag{65}$$

Let $\bar{q} = [q_0, q_1, q_2, q_3]^{\mathrm{T}} = [q_0, q^{\mathrm{T}}]^{\mathrm{T}} = \left[\cos\left(\frac{\alpha}{2}\right), \hat{e}^{\mathrm{T}} \sin\left(\frac{\alpha}{2}\right)\right]^{\mathrm{T}}$ be the quaternion representing the rotation of the body frame relative to the LVLH frame, where $\hat{e}$ is the unit length rotational axis and α is the rotation angle about $\hat{e}$. Therefore, the reduced kinematics equation becomes

$$\begin{bmatrix}\dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3\end{bmatrix} = \frac{1}{2}\begin{bmatrix} q_0 & -q_3 & q_2 \\ q_3 & q_0 & -q_1 \\ -q_2 & q_1 & q_0\end{bmatrix}\begin{bmatrix}\omega_1 \\ \omega_2 \\ \omega_3\end{bmatrix} = g(q_1, q_2, q_3, \omega). \tag{66}$$

From (42),

$$A_l^b = \begin{bmatrix} 2q_0^2 - 1 + 2q_1^2 & 2q_1q_2 + 2q_0q_3 & 2q_1q_3 - 2q_0q_2 \\ 2q_1q_2 - 2q_0q_3 & 2q_0^2 - 1 + 2q_2^2 & 2q_2q_3 + 2q_0q_1 \\ 2q_1q_3 + 2q_0q_2 & 2q_2q_3 - 2q_0q_1 & 2q_0^2 - 1 + 2q_3^2\end{bmatrix}.$$

3.3.2. The linearized nadir pointing spacecraft model

It is difficult to design a controller with specified performance using the nonlinear spacecraft system model described by (65) and (66). The common practice is to design the controller using a linearized system and then check if the designed controller works for the original nonlinear system using simulation and/or test. For a nadir pointing spacecraft system, we need the closed loop spacecraft system to have the following features: (a) the spacecraft body rate with respect to the LVLH frame is as small as possible, ideally, $\omega = 0$; and (b) the spacecraft body frame is aligned with the LVLH frame, i.e., the error is as small as possible, ideally, $q_1 = q_2 = q_3 = 0$. Since the rotation axis length is always 1, this implies that the rotation angle $\alpha = 0$. Therefore the linearized model is the first order model of Taylor expansion of the nonlinear system (65) and (66) about $\omega = 0$ and $q_1 = q_2 = q_3 = 0$. By using quaternion representation of A_l^b, assuming J is almost diagonal (which is almost always true in real spacecraft designs), and neglecting high order terms of q_1, q_2, and q_3, we have the following relations.

$$\omega_{lvlh}^b = A_l^b \omega_{lvlh} = \begin{bmatrix} 2q_1q_2 + 2q_0q_3 \\ 2q_0^2 - 1 + 2q_2^2 \\ 2q_2q_3 - 2q_0q_1\end{bmatrix}\omega_0 \Bigg|_{\substack{\omega \approx 0 \\ q_1 = q_2 = q_3 \approx 0}} \approx \begin{bmatrix} 2q_3 \\ 1 \\ -2q_1\end{bmatrix}\omega_0. \tag{67}$$

Using (5), we have

$$\begin{aligned}&\omega_{lvlh}^b \times (J\omega_{lvlh}^b)\Big|_{\substack{\omega \approx 0 \\ q_1 = q_2 = q_3 \approx 0}} \\ &= \begin{bmatrix} 0 & 2q_1\omega_0 & \omega_0 \\ -2q_1\omega_0 & 0 & -2q_3\omega_0 \\ -\omega_0 & 2q_3\omega_0 & 0\end{bmatrix}\begin{bmatrix} 2J_{11}q_3\omega_0 \\ J_{22}\omega_0 \\ -2J_{33}q_1\omega_0\end{bmatrix}\Bigg|_{\substack{\omega \approx 0 \\ q_1 = q_2 = q_3 \approx 0}} \\ &\approx \omega_0^2\begin{bmatrix} 2(J_{22} - J_{33})q_1 \\ 0 \\ 2(J_{22} - J_{11})q_3\end{bmatrix},\end{aligned} \tag{68}$$

$$\begin{aligned}\omega_{lvlh}^b \times H\Big|_{\substack{\omega \approx 0 \\ q_1 = q_2 = q_3 \approx 0}} &= \begin{bmatrix} 0 & 2q_1\omega_0 & \omega_0 \\ -2q_1\omega_0 & 0 & -2q_3\omega_0 \\ -\omega_0 & 2q_3\omega_0 & 0\end{bmatrix}\begin{bmatrix} 0 \\ h_2 \\ 0\end{bmatrix}\Bigg|_{\substack{\omega \approx 0 \\ q_1 = q_2 = q_3 \approx 0}} \\ &\approx \omega_0\begin{bmatrix} 2h_2q_1 \\ 0 \\ 2h_2q_3\end{bmatrix}.\end{aligned} \tag{69}$$

Using (67)–(69), we have

$$\begin{aligned}\frac{\partial f}{\partial \omega}\Big|_{\substack{\omega \approx 0 \\ q_1 = q_2 = q_3 \approx 0}} &\approx -J\Omega(\omega_{lvlh}^b) + \Omega(J\omega_{lvlh}^b) - \Omega(\omega_{lvlh}^b)J \\ &\quad + \Omega(H),\end{aligned} \tag{70}$$

$$\begin{aligned}\frac{\partial f}{\partial q}\Big|_{\substack{\omega \approx 0 \\ q_1 = q_2 = q_3 \approx 0}} &= \frac{\partial(-\omega_{lvlh}^b \times (J\omega_{lvlh}^b) - \omega_{lvlh}^b \times H)}{\partial q}\Bigg|_{\substack{\omega \approx 0 \\ q_1 = q_2 = q_3 \approx 0}} \\ &\approx \begin{bmatrix} 2\omega_0^2(J_{33} - J_{22}) - 2h_2\omega_0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2\omega_0^2(J_{11} - J_{22}) - 2h_2\omega_0\end{bmatrix},\end{aligned} \tag{71}$$

$$\frac{\partial g}{\partial \omega}\Big|_{\substack{\omega \approx 0 \\ q_1 = q_2 = q_3 \approx 0}} \approx \frac{1}{2}I_3, \tag{72}$$

[2] This assumption is true for most satellites as long as the orbit eccentricity is small, i.e., the orbit is close to a circle.

$$\left.\frac{\partial g}{\partial q}\right|_{\substack{\omega\approx 0\\ q_1=q_2=q_3\approx 0}} \approx \frac{1}{2}0_3, \tag{73}$$

where I_3 is a 3×3 dimensional identity matrix, 0_3 is a 3×3 dimensional zero matrix. (70) can be simplified further as follows.

$$J\Omega(\omega^b_{lvlh}) = \begin{bmatrix} -J_{13}\omega_0 & 0 & J_{11}\omega_0 \\ -J_{23}\omega_0 & 0 & J_{21}\omega_0 \\ -J_{33}\omega_0 & 0 & J_{31}\omega_0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & J_{11}\omega_0 \\ 0 & 0 & 0 \\ -J_{33}\omega_0 & 0 & J0 \end{bmatrix}. \tag{74}$$

$$\Omega(J\omega^b_{lvlh}) = \begin{bmatrix} 0 & -J_{32}\omega_0 & J_{22}\omega_0 \\ J_{32}\omega_0 & 0 & -J_{12}\omega_0 \\ -J_{22}\omega_0 & J_{12}\omega_0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & J_{22}\omega_0 \\ 0 & 0 & 0 \\ -J_{22}\omega_0 & 0 & 0 \end{bmatrix}. \tag{75}$$

$$\Omega(\omega^b_{lvlh})J = \begin{bmatrix} J_{31}\omega_0 & J_{32}\omega_0 & J_{33}\omega_0 \\ 0 & 0 & 0 \\ -J_{11}\omega_0 & -J_{12}\omega_0 & -J_{13}\omega_0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & J_{33}\omega_0 \\ 0 & 0 & 0 \\ -J_{11}\omega_0 & 0 & 0 \end{bmatrix}. \tag{76}$$

$$\Omega(H) = \begin{bmatrix} 0 & 0 & h_2 \\ 0 & 0 & 0 \\ -h_2 & 0 & 0 \end{bmatrix}. \tag{77}$$

Therefore

$$\left.\frac{\partial f}{\partial \omega}\right|_{\substack{\omega\approx 0\\ q_1=q_2=q_3\approx 0}} = \begin{bmatrix} 0 & 0 & -(J_{11}-J_{22}+J_{33})\omega_0+h_2 \\ 0 & 0 & 0 \\ (J_{11}-J_{22}+J_{33})\omega_0-h_2 & 0 & 0 \end{bmatrix}. \tag{78}$$

For many nadir pointing satellites, we need to model disturbance torque in the linearized model. For low Earth orbit spacecraft, aerodynamic torque and gravity gradient torque are the dominant disturbance torques. It is difficult to model the aero dynamic torque because it is related to solar activity, geomagnetic index, spacecraft geometry, spacecraft attitude, spacecraft altitude, and many other factors, but it is known that the gravity gradient torque can be modeled by (see Sidi, 1997)

$$T_{gg} = \begin{bmatrix} 3\omega_0^2(J_{33}-J_{22})\phi \\ 3\omega_0^2(J_{33}-J_{11})\theta \\ 0 \end{bmatrix}. \tag{79}$$

where ϕ and θ are the Euler angles for the roll and the pitch. For small Euler angles (see Wertz, 1978), $\phi = 2q_1$ and $\theta = 2q_2$, this gives

$$T_{gg} = \begin{bmatrix} 6\omega_0^2(J_{33}-J_{22})q_1 \\ 6\omega_0^2(J_{33}-J_{11})q_2 \\ 0 \end{bmatrix}. \tag{80}$$

From (65),

$$J\dot{\omega} \approx \frac{\partial f}{\partial \omega}\omega + \frac{\partial f}{\partial q}q + T_d + u. \tag{81}$$

Assuming $T_d = T_{gg}$, and combining Eqs. (81), (71)–(73), (78), and (80), we have the quaternion based linearized spacecraft system described by

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & J_{11} & J_{12} & J_{13} \\ 0 & 0 & 0 & J_{21} & J_{22} & J_{23} \\ 0 & 0 & 0 & J_{31} & J_{32} & J_{33} \end{bmatrix} \begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \\ \dot{\omega}_1 \\ \dot{\omega}_2 \\ \dot{\omega}_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & .5 & 0 & 0 \\ 0 & 0 & 0 & 0 & .5 & 0 \\ 0 & 0 & 0 & 0 & 0 & .5 \\ f_{41} & 0 & 0 & 0 & 0 & f_{46} \\ 0 & f_{52} & 0 & 0 & 0 & 0 \\ 0 & 0 & f_{63} & f_{64} & 0 & 0 \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ u_x \\ u_y \\ u_z \end{bmatrix} \tag{82}$$

where $f_{41} = 8(J_{33}-J_{22})\omega_0^2 - 2h_2\omega_0, f_{46} = (-J_{11}+J_{22}-J_{33})\omega_0 + h_2, f_{64} = -f_{46}, f_{52} = 6(J_{33}-J_{11})\omega_0^2$, and $f_{63} = 2(J_{11}-J_{22})\omega_0^2 - 2h_2\omega_0$. It is straightforward to check that the linearized spacecraft model is fully controllable. Therefore, all modern control design methods in linear system theory can be applied directly, and the designed linear system is guaranteed to be stable.

4. Attitude determination

Attitude determination is very important for two reasons. First, control engineers need to know if the spacecraft attitude is in the right attitude. Scond, if the spacecraft attitude is not in the perfect position, the attitude information will be compared automatically with the desired attitude, and the error information is then used to calculate how much action is needed for each actuator to bring the spacecraft to the desired attitude.

From the Section 2.3.4, we have seen that to determine the frame rotation, we need to know the coordinates of two vector pairs before and after the rotation. Given this coordinate information, we can determine the rotational axis and the rotational angle, which determine the quaternion describing the attitude deviation of the body frame from the desired frame. This intuition has been used by many researchers to develop their attitude determination methods, such as Black (1964), Wahba (1965), Shuster and Oh (1981), Bar-Itzhack and Oshman (1985), Shuster (1993), Reynolds (1998), and Markley (2002). In this section, we will first introduce Wahba's problem (Wahba, 1965), then Davenport's formula (Davenport, 1965), then a well-known method QUEST (Shuster & Oh, 1981), an analytic solution for a special case of Wahba's problem developed in Reynolds (1998) and Markley (2002), and finally our analytic solution to the general Wahba's problem.

4.1. Wahba's problem

Suppose that we have observed two astronomical objects such as stars, or the Sun, or the Earth, or measured some ambient vector field such as the Earth's magnetic field; and we represent the measurement as two unit vectors $\mathbf{b}_1$ and $\mathbf{b}_2$ in the spacecraft body frame. We use unit vectors because the length of the vectors has no information relevant to the attitude determination and unit length makes expression simpler. To determine the spacecraft attitude, it is also necessary to know the same two unit vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ represented in some reference frame. The reference frame is usually taken to be an inertial frame or local vertical local horizontal frame. The attitude to be determined is then the rotational matrix or the quaternion that rotates these vectors from the reference frame to the spacecraft body frame. Therefore one can find an attitude matrix such that

$$A\mathbf{r}_1 = \mathbf{b}_1, \tag{83a}$$

$$A\mathbf{r}_2 = \mathbf{b}_2. \tag{83b}$$

(83) implies

$$\mathbf{b}_1 \cdot \mathbf{b}_2 = (A\mathbf{r}_1) \cdot (A\mathbf{r}_2) = \mathbf{r}_1^{\mathrm{T}} A^{\mathrm{T}} A \mathbf{r}_2 = \mathbf{r}_1 \cdot \mathbf{r}_2. \tag{84}$$

In general, given two sets of m known reference vectors $\{\mathbf{r}_1, \ldots, \mathbf{r}_m\}$ and m observation vectors $\{\mathbf{b}_1, \ldots, \mathbf{b}_m\}$, $m \geqslant 2$, find the proper rotational matrix A which brings the first set into the best least squares coincidence with the second, i.e.,

$$\min_A \frac{1}{2} \sum_{i=1}^{m} \|\mathbf{b}_i - A\mathbf{r}_i\|^2. \tag{85}$$

This problem was first defined by Wahba and is called Wahba's problem (Wahba, 1965) which is the base of the most attitude determination methods.

4.2. Davenport's formula

All popular methods, such as QUEST (Shuster & Oh, 1981), ESOQ (Mortari, 1997), and FOMA (Markley, 1993), use Davenport's q-method (Davenport, 1965) (K-matrix derivation is accessible in Keat (1977)). Rewrite (85) and use the facts that $\mathbf{b}_i$ and $\mathbf{r}_i$ are unit vectors and A is orthogonal matrix, we have

$$\frac{1}{2} \sum_{i=1}^{m} \|\mathbf{b}_i - A\mathbf{r}_i\|^2 = m - \sum_{i=1}^{m} \mathbf{b}_i^{\mathrm{T}} A \mathbf{r}_i = m - Tr(W^{\mathrm{T}} A V), \tag{86}$$

where $W = [\mathbf{b}_1, \ldots, \mathbf{b}_m]$, $V = [\mathbf{r}_1, \ldots, \mathbf{r}_m]$, and $Tr(\cdot)$ represents the trace of the matrix in the argument. Using (43), we have

$$\begin{aligned} Tr(W^{\mathrm{T}} A V) &= Tr(W^{\mathrm{T}}((q_0^2 - q^{\mathrm{T}} q) I + 2qq^{\mathrm{T}} - 2q_0 \Omega(q)) V) \\ &= (q_0^2 - q^{\mathrm{T}} q) Tr(W^{\mathrm{T}} V) + 2Tr(qq^{\mathrm{T}} V W^{\mathrm{T}}) \\ &\quad - 2q_0 Tr(W^{\mathrm{T}} \Omega(q) V). \end{aligned} \tag{87}$$

Let $B = WV^{\mathrm{T}}$, $\sigma = Tr(B)$, $S = B + B^{\mathrm{T}}$, and $Z^{\mathrm{T}} = [B_{23} - B_{32},\ B_{31} - B_{13},\ B_{12} - B_{21}]$. The second term of (87) can be rewritten as

$$2Tr(qq^{\mathrm{T}} V W^{\mathrm{T}}) = q^{\mathrm{T}} S q. \tag{88}$$

Since $\Omega(Z) = B^{\mathrm{T}} - B$, $\Omega(q)^{\mathrm{T}} = -\Omega(q)$, and $Tr(\Omega(q)\Omega(Z)) = -2q^{\mathrm{T}} Z$, the third term of (87) can be rewritten as

$$\begin{aligned} 2q_0 Tr(\Omega(q) V W^{\mathrm{T}}) &= q_0 Tr(\Omega(q) B^{\mathrm{T}} + B\Omega(q)^{\mathrm{T}}) = q_0 Tr(\Omega(q) B^{\mathrm{T}} + \Omega(q)^{\mathrm{T}} B) \\ &= q_0 Tr(\Omega(q)(B^{\mathrm{T}} - B)) = q_0 Tr(\Omega(q)\Omega(Z)) \\ &= -2q_0 q^{\mathrm{T}} Z. \end{aligned} \tag{89}$$

Substituting (88) and (89) into (87) produces

$$\begin{aligned} Tr(W^{\mathrm{T}} A V) &= (q_0^2 - q^{\mathrm{T}} q)\sigma + q^{\mathrm{T}} S q + 2q_0 q^{\mathrm{T}} Z \\ &= [q_0\ q^{\mathrm{T}}] \begin{bmatrix} \sigma & Z^{\mathrm{T}} \\ Z & S - \sigma I \end{bmatrix} \begin{bmatrix} q_0 \\ q \end{bmatrix} := \bar{q}^{\mathrm{T}} K \bar{q}. \end{aligned} \tag{90}$$

Therefore,

$$\min_A \frac{1}{2} \sum_{i=1}^{m} \|\mathbf{b}_i - A\mathbf{r}_i\|^2 = m - \max_A Tr(W^{\mathrm{T}} A V) = m - \max_{\bar{q}=1} \bar{q}^{\mathrm{T}} K \bar{q}. \tag{91}$$

By introducing the Lagrange multiplier λ for the unit length constraint of $\|\bar{q}\| = 1$, we reduce Wahba's problem to Davenport's problem

$$\max_{\lambda, \bar{q}} \bar{q}^{\mathrm{T}} K \bar{q} - \lambda \bar{q}^{\mathrm{T}} \bar{q}. \tag{92}$$

Taking the derivative of (92) gives the optimal solution which satisfies

$$K\bar{q} = \lambda \bar{q}. \tag{93}$$

The optimization problem is reduced to finding the largest eigenvalue of K and its corresponding eigenvector which is Davenport's formula.

4.3. Quaternion estimation using QUEST, FOMA, and ESOQ

In the early of 1980s, the computation of the largest eigenvalue and its corresponding eigenvector of the K-matrix in an on-board computer was a burden. Shuster and Oh (1981) developed QUEST algorithm to approximately solve (93). By using the Cayley-Hamilton theorem (cf. Rugh, 1993), Shuster and Oh (1981) derived the first analytic formula of the characteristic polynomial of the K-matrix which is a polynomial of degree of 4, given as

$$f(\lambda) = \lambda^4 - (a + b)\lambda^2 - c\lambda + (ab + c\sigma - d) = 0, \tag{94}$$

where $\sigma = 0.5Tr(S) = Tr(B)$, $\kappa = Tr(adj(S))$, $\Delta = \det(S)$, $a = \sigma^2 - \kappa$, $b = \sigma^2 + Z^{\mathrm{T}} Z$, $c = \Delta + Z^{\mathrm{T}} S Z$, $d = Z^{\mathrm{T}} S^2 Z$.

For many applications, the largest eigenvalue may be approximated by $\lambda \approx m$. Shuster and Oh (1981) suggested using Newton-Raphson iteration to find the λ using the initial guess $\lambda^0 = m$. To calculate the eigenvector using λ, Shuster used the Rodriguez parameters defined as follows

$$p = \frac{q}{q_0} = q \tan\left(\frac{\alpha}{2}\right).$$

Since $K\bar{q} = \lambda \bar{q}$, from the K-matrix, it is easy to see that

$$[(\lambda + \sigma) I - S] p = Z.$$

p can be obtained by solving linear system equations. Once p is available, the quaternion is given by

$$\bar{q} = \frac{1}{\sqrt{1 + p^{\mathrm{T}} p}} \begin{bmatrix} p \\ 1 \end{bmatrix}.$$

To avoid the possible singularity in Rodriguez parameter, Shuster and Oh developed a method of sequential rotations which avoids the singularity. This method is widely recognized and is refereed to as the QUEST method.

In 1993, Markley (1993) derived an equivalent characteristic polynomial for the K-matrix and also used Newton's method for his expression of the polynomial to find the largest eigenvalue λ iteratively. Markley's method is now refereed to as the FOMA algorithm. By using a different derivation, Mortari (1997) found another equivalent characteristic polynomial for the K-matrix. He proposed an algorithm based on Newton's method for this expression of the polynomial and refereed to it as the ESOQ algorithm.

4.4. Attitude determination using two vector measurements

Though QUEST is very efficient, if the attitude determination is based on only two vector measurements, there is a simpler method which is an analytic solution (Markley, 2002).

4.4.1. The minimum-angle rotation quaternion

First, it is worthwhile to notice that for the quaternion which maps the reference vector $\mathbf{r}_1$ to the body frame vector $\mathbf{b}_1$, the minimal rotational angle α is determined by $\cos(\alpha) = \mathbf{b}_1 \cdot \mathbf{r}_1$; using the minimum-angle rotation quaternion, the rotational axis must be perpendicular to $\mathbf{r}_1$ and $\mathbf{b}_1$ and satisfy the right-hand rule, which means that the unit length rotational axis is given by $\hat{e} = \frac{\mathbf{b}_1 \times \mathbf{r}_1}{\sin(\alpha)}$. Using the following identities of the trigonometry (Polyanin & Manzhirov, 2007)

$$\frac{1 - \cos(\alpha)}{2} = \sin^2\left(\frac{\alpha}{2}\right),$$

$$\cot\left(\frac{\alpha}{2}\right) = \frac{1 + \cos(\alpha)}{\sin(\alpha)},$$

we can verify that the minimum-angle rotation quaternion is given by

$$\begin{aligned}
&\frac{1}{\sqrt{2(1+\mathbf{b}_1\cdot\mathbf{r}_1)}}(1+\mathbf{b}_1\cdot\mathbf{r}_1,\mathbf{b}_1\times\mathbf{r}_1)\\
&\quad=(1+\cos(\alpha),\mathbf{b}_1\times\mathbf{r}_1)\sqrt{\frac{1}{2(1+\cos(\alpha))}}\\
&\quad=(1+\cos(\alpha),\mathbf{b}_1\times\mathbf{r}_1)\sqrt{\frac{1-\cos(\alpha)}{2(1-\cos^2(\alpha))}}\\
&\quad=(1+\cos(\alpha),\mathbf{b}_1\times\mathbf{r}_1)\frac{\sin\left(\frac{\alpha}{2}\right)}{\sin(\alpha)}\\
&\quad=\left(\frac{1+\cos(\alpha)}{\sin(\alpha)},\frac{\mathbf{b}_1\times\mathbf{r}_1}{\sin(\alpha)}\right)\sin\left(\frac{\alpha}{2}\right)=\left(\cot\left(\frac{\alpha}{2}\right),\hat{e}\right)\sin\left(\frac{\alpha}{2}\right)\\
&\quad=\left(\cos\left(\frac{\alpha}{2}\right),\hat{e}\sin\left(\frac{\alpha}{2}\right)\right)=\bar{q}_{min}. \qquad (95)
\end{aligned}$$

4.4.2. The general rotation quaternion

Denote $\bar{q}(\hat{e},\alpha)$ as the quaternion that has rotational axis $\hat{e}$ and rotational angle α. The most general rotation that maps $\mathbf{r}_1$ to $\mathbf{b}_1$ is given by

$$\bar{q}_1=\bar{q}(\mathbf{r}_1,\phi_r)\otimes\bar{q}_{min}\otimes\bar{q}(\mathbf{b}_1,\phi_b), \qquad (96)$$

where ϕ_b and ϕ_r are arbitrary angles of rotation about $\mathbf{b}_1$ and $\mathbf{r}_1$, respectively. Using (1)–(3), (25), and the facts that

$$\sin(\alpha+\beta)=\sin(\alpha)\cos(\beta)+\cos(\alpha)\sin(\beta), \qquad (97)$$

and

$$\cos(\alpha+\beta)=\cos(\alpha)\cos(\beta)-\sin(\alpha)\sin(\beta), \qquad (98)$$

we can reduce (96) as follows.

$$\begin{aligned}
&(1+\mathbf{b}_1\cdot\mathbf{r}_1,\mathbf{b}_1\times\mathbf{r}_1)\otimes\left(\cos\left(\frac{\phi_b}{2}\right),\mathbf{b}_1\sin\left(\frac{\phi_b}{2}\right)\right)\\
&\quad=\left((1+\mathbf{b}_1\cdot\mathbf{r}_1)\cos\left(\frac{\phi_b}{2}\right)-(\mathbf{b}_1\times\mathbf{r}_1)\cdot\mathbf{b}_1\sin\left(\frac{\phi_b}{2}\right),\right.\\
&\qquad(1+\mathbf{b}_1\cdot\mathbf{r}_1)\mathbf{b}_1\sin\left(\frac{\phi_b}{2}\right)+(\mathbf{b}_1\times\mathbf{r}_1)\cos\left(\frac{\phi_b}{2}\right)\\
&\qquad\left.+(\mathbf{b}_1\times\mathbf{r}_1)\times\mathbf{b}_1\sin\left(\frac{\phi_b}{2}\right)\right)\\
&\quad=\left((1+\mathbf{b}_1\cdot\mathbf{r}_1)\cos\left(\frac{\phi_b}{2}\right),(1+\mathbf{b}_1\cdot\mathbf{r}_1)\mathbf{b}_1\sin\left(\frac{\phi_b}{2}\right)\right.\\
&\qquad\left.+(\mathbf{b}_1\times\mathbf{r}_1)\cos\left(\frac{\phi_b}{2}\right)+(\mathbf{r}_1-(\mathbf{b}_1\cdot\mathbf{r}_1)\mathbf{b}_1)\sin\left(\frac{\phi_b}{2}\right)\right)\\
&\quad=\left((1+\mathbf{b}_1\cdot\mathbf{r}_1)\cos\left(\frac{\phi_b}{2}\right),(\mathbf{b}_1+\mathbf{r}_1)\sin\left(\frac{\phi_b}{2}\right)\right.\\
&\qquad\left.+(\mathbf{b}_1\times\mathbf{r}_1)\cos\left(\frac{\phi_b}{2}\right)\right). \qquad (99)
\end{aligned}$$

Thus,

$$\begin{aligned}
&\left(\cos\left(\frac{\phi_r}{2}\right),\mathbf{r}_1\sin\left(\frac{\phi_r}{2}\right)\right)\\
&\quad\otimes(1+\mathbf{b}_1\cdot\mathbf{r}_1,\mathbf{b}_1\times\mathbf{r}_1)\otimes\left(\cos\left(\frac{\phi_b}{2}\right),\mathbf{b}_1\sin\left(\frac{\phi_b}{2}\right)\right)\\
&\quad=\left(\cos\left(\frac{\phi_r}{2}\right),\mathbf{r}_1\sin\left(\frac{\phi_r}{2}\right)\right)\otimes\left((1+\mathbf{b}_1\cdot\mathbf{r}_1)\cos\left(\frac{\phi_b}{2}\right),\right.\\
&\qquad\left.(\mathbf{b}_1+\mathbf{r}_1)\sin\left(\frac{\phi_b}{2}\right)+(\mathbf{b}_1\times\mathbf{r}_1)\cos\left(\frac{\phi_b}{2}\right)\right). \qquad (100)
\end{aligned}$$

Let q_0 and q be the scalar part and vector part of the quaternion defined by (100). using (3) and (98), we have

$$\begin{aligned}
q_0&=(1+\mathbf{b}_1\cdot\mathbf{r}_1)\cos\left(\frac{\phi_r}{2}\right)\cos\left(\frac{\phi_b}{2}\right)\\
&\quad-\mathbf{r}_1\cdot(\mathbf{b}_1+\mathbf{r}_1)\sin\left(\frac{\phi_r}{2}\right)\sin\left(\frac{\phi_b}{2}\right)\\
&\quad-\mathbf{r}_1\cdot(\mathbf{b}_1\times\mathbf{r}_1)\sin\left(\frac{\phi_r}{2}\right)\cos\left(\frac{\phi_b}{2}\right)\\
&=(1+\mathbf{b}_1\cdot\mathbf{r}_1)\cos\left(\frac{\phi_r}{2}\right)\cos\left(\frac{\phi_b}{2}\right)\\
&\quad-(1+\mathbf{b}_1\cdot\mathbf{r}_1)\sin\left(\frac{\phi_r}{2}\right)\sin\left(\frac{\phi_b}{2}\right)\\
&=(1+\mathbf{b}_1\cdot\mathbf{r}_1)\left(\cos\left(\frac{\phi_r}{2}\right)\cos\left(\frac{\phi_b}{2}\right)-\sin\left(\frac{\phi_r}{2}\right)\sin\left(\frac{\phi_b}{2}\right)\right)\\
&=(1+\mathbf{b}_1\cdot\mathbf{r}_1)\cos\left(\frac{\phi_r+\phi_b}{2}\right). \qquad (101)
\end{aligned}$$

From (100), using (2), (97), and (98), we have

$$\begin{aligned}
q&=(1+\mathbf{b}_1\cdot\mathbf{r}_1)\mathbf{r}_1\cos\left(\frac{\phi_b}{2}\right)\sin\left(\frac{\phi_r}{2}\right)+(\mathbf{b}_1+\mathbf{r}_1)\sin\left(\frac{\phi_b}{2}\right)\cos\left(\frac{\phi_r}{2}\right)\\
&\quad+(\mathbf{b}_1\times\mathbf{r}_1)\cos\left(\frac{\phi_b}{2}\right)\cos\left(\frac{\phi_r}{2}\right)\\
&\quad+\mathbf{r}_1\times(\mathbf{b}_1+\mathbf{r}_1)\sin\left(\frac{\phi_r}{2}\right)\sin\left(\frac{\phi_b}{2}\right)\\
&\quad+\mathbf{r}_1\times(\mathbf{b}_1\times\mathbf{r}_1)\sin\left(\frac{\phi_r}{2}\right)\cos\left(\frac{\phi_b}{2}\right)\\
&=\mathbf{r}_1\cos\left(\frac{\phi_b}{2}\right)\sin\left(\frac{\phi_r}{2}\right)+(\mathbf{b}_1\cdot\mathbf{r}_1)\mathbf{r}_1\cos\left(\frac{\phi_b}{2}\right)\sin\left(\frac{\phi_r}{2}\right)\\
&\quad+(\mathbf{b}_1+\mathbf{r}_1)\sin\left(\frac{\phi_b}{2}\right)\cos\left(\frac{\phi_r}{2}\right)+(\mathbf{b}_1\times\mathbf{r}_1)\cos\left(\frac{\phi_r}{2}\right)\cos\left(\frac{\phi_b}{2}\right)\\
&\quad-(\mathbf{b}_1\times\mathbf{r}_1)\sin\left(\frac{\phi_r}{2}\right)\sin\left(\frac{\phi_b}{2}\right)+\mathbf{b}_1\sin\left(\frac{\phi_r}{2}\right)\cos\left(\frac{\phi_b}{2}\right)\\
&\quad-(\mathbf{b}_1\cdot\mathbf{r}_1)\mathbf{r}_1\cos\left(\frac{\phi_b}{2}\right)\sin\left(\frac{\phi_r}{2}\right)\\
&=(\mathbf{b}_1+\mathbf{r}_1)\sin\left(\frac{\phi_b}{2}\right)\cos\left(\frac{\phi_r}{2}\right)+(\mathbf{b}_1+\mathbf{r}_1)\sin\left(\frac{\phi_r}{2}\right)\cos\left(\frac{\phi_b}{2}\right)\\
&\quad+(\mathbf{b}_1\times\mathbf{r}_1)\left(\cos\left(\frac{\phi_r}{2}\right)\cos\left(\frac{\phi_b}{2}\right)-\sin\left(\frac{\phi_r}{2}\right)\sin\left(\frac{\phi_b}{2}\right)\right)\\
&=(\mathbf{b}_1+\mathbf{r}_1)\sin\left(\frac{\phi_r+\phi_b}{2}\right)+(\mathbf{b}_1\times\mathbf{r}_1)\cos\left(\frac{\phi_r+\phi_b}{2}\right)\\
&=(\mathbf{b}_1+\mathbf{r}_1)\sin\left(\frac{\phi}{2}\right)+(\mathbf{b}_1\times\mathbf{r}_1)\cos\left(\frac{\phi}{2}\right), \qquad (102)
\end{aligned}$$

where $\phi = \phi_r + \phi_b$. Combining (96), (95), (100), (101), and (102) yields

$$\begin{aligned}
\bar{q}_1&=\frac{1}{\sqrt{2(1+\mathbf{b}_1\cdot\mathbf{r}_1)}}\left((1+\mathbf{b}_1\cdot\mathbf{r}_1)\cos\left(\frac{\phi}{2}\right),(\mathbf{b}_1\times\mathbf{r}_1)\cos\left(\frac{\phi}{2}\right)\right.\\
&\qquad\left.+(\mathbf{b}_1+\mathbf{r}_1)\sin\left(\frac{\phi}{2}\right)\right). \qquad (103)
\end{aligned}$$

Similarly, the most general rotation that maps $\mathbf{r}_2$ to $\mathbf{b}_2$ is given by

$$\begin{aligned}
\bar{q}_2&=\frac{1}{\sqrt{2(1+\mathbf{b}_2\cdot\mathbf{r}_2)}}\left((1+\mathbf{b}_2\cdot\mathbf{r}_2)\cos\left(\frac{\psi}{2}\right),(\mathbf{b}_2\times\mathbf{r}_2)\cos\left(\frac{\psi}{2}\right)\right.\\
&\qquad\left.+(\mathbf{b}_2+\mathbf{r}_2)\sin\left(\frac{\psi}{2}\right)\right) \qquad (104)
\end{aligned}$$

for some angle ψ.

4.4.3. Attitude determination using two vector measurements

As every quaternion in the family of $\bar{q}_1(\phi)$ maps $\mathbf{r}_1$ to $\mathbf{b}_1$ and every quaternion in the family of $\bar{q}_2(\psi)$ maps $\mathbf{r}_2$ to $\mathbf{b}_2$, we need to find a quaternion $\bar{q}$ which is in both families so that it maps $\mathbf{r}_1$ to $\mathbf{b}_1$ and $\mathbf{r}_2$ to $\mathbf{b}_2$ simultaneously. This means that both the scalar part and the vector part of $\bar{q}_1$ and $\bar{q}_2$ are equal for some ϕ and ψ. For the scalar part, we need

$$\frac{(1+\mathbf{r}_1\cdot\mathbf{b}_1)}{\sqrt{2(1+\mathbf{r}_1\cdot\mathbf{b}_1)}}\cos\left(\frac{\phi}{2}\right)=\frac{(1+\mathbf{r}_2\cdot\mathbf{b}_2)}{\sqrt{2(1+\mathbf{r}_2\cdot\mathbf{b}_2)}}\cos\left(\frac{\psi}{2}\right)$$

$$\iff \cos\left(\frac{\psi}{2}\right)=\sqrt{\frac{1+\mathbf{r}_1\cdot\mathbf{b}_1}{1+\mathbf{r}_2\cdot\mathbf{b}_2}}\cos\left(\frac{\phi}{2}\right) \tag{105a}$$

$$\iff \sin\left(\frac{\psi}{2}\right)=\sqrt{\frac{1+\mathbf{r}_2\cdot\mathbf{b}_2-(1+\mathbf{r}_1\cdot\mathbf{b}_1)\cos^2\left(\frac{\phi}{2}\right)}{1+\mathbf{r}_2\cdot\mathbf{b}_2}}. \tag{105b}$$

For vector part, we need

$$\frac{(\mathbf{b}_1\times\mathbf{r}_1)}{\sqrt{(1+\mathbf{b}_1\cdot\mathbf{r}_1)}}\cos\left(\frac{\phi}{2}\right)+\frac{(\mathbf{b}_1+\mathbf{r}_1)}{\sqrt{(1+\mathbf{b}_1\cdot\mathbf{r}_1)}}\sin\left(\frac{\phi}{2}\right)$$
$$=\frac{(\mathbf{b}_2\times\mathbf{r}_2)}{\sqrt{(1+\mathbf{b}_2\cdot\mathbf{r}_2)}}\cos\left(\frac{\psi}{2}\right)+\frac{(\mathbf{b}_2+\mathbf{r}_2)}{\sqrt{(1+\mathbf{b}_2\cdot\mathbf{r}_2)}}\sin\left(\frac{\psi}{2}\right) \tag{106}$$

Substituting (105a) and (105b) into (106) yields

$$(\mathbf{b}_1\times\mathbf{r}_1)\cos\left(\frac{\phi}{2}\right)+(\mathbf{b}_1+\mathbf{r}_1)\sin\left(\frac{\phi}{2}\right)$$
$$=\frac{1+\mathbf{b}_1\cdot\mathbf{r}_1}{1+\mathbf{b}_2\cdot\mathbf{r}_2}\cos\left(\frac{\phi}{2}\right)(\mathbf{b}_2\times\mathbf{r}_2)+(\mathbf{b}_2+\mathbf{r}_2)\frac{\sqrt{(1+\mathbf{b}_1\cdot\mathbf{r}_1)}}{(1+\mathbf{b}_2\cdot\mathbf{r}_2)}$$
$$\times\sqrt{1+\mathbf{b}_2\cdot\mathbf{r}_2-(1+\mathbf{b}_1\cdot\mathbf{r}_1)\cos^2\left(\frac{\phi}{2}\right)}.$$

Apply dot product of $\mathbf{b}_2-\mathbf{r}_2$ on both side, the right-hand side vanishes because $(\mathbf{b}_2+\mathbf{r}_2)\cdot(\mathbf{b}_2-\mathbf{r}_2)=0$ and $(\mathbf{b}_2\times\mathbf{r}_2)\cdot(\mathbf{b}_2-\mathbf{r}_2)=0$. Therefore,

$$(\mathbf{b}_1\times\mathbf{r}_1)\cdot(\mathbf{b}_2-\mathbf{r}_2)\cos\left(\frac{\phi}{2}\right)+(\mathbf{b}_1+\mathbf{r}_1)\cdot(\mathbf{b}_2-\mathbf{r}_2)\sin\left(\frac{\phi}{2}\right)=0,$$

or

$$\frac{\sin\left(\frac{\phi}{2}\right)}{\cos\left(\frac{\phi}{2}\right)}=-\frac{(\mathbf{b}_1\times\mathbf{r}_1)\cdot(\mathbf{b}_2-\mathbf{r}_2)}{(\mathbf{b}_1+\mathbf{r}_1)\cdot(\mathbf{b}_2-\mathbf{r}_2)}. \tag{107}$$

For any two vectors $\mathbf{a}$ and $\mathbf{b}$, if $\mathbf{a}$ is proportional to $\mathbf{b}$, we denote this relation as $\mathbf{a}\propto\mathbf{b}$. Clearly, if $\mathbf{a}\propto\mathbf{b}$, and $\mathbf{b}\propto\mathbf{c}$, then $\mathbf{a}\propto\mathbf{c}$. From (103) and (107), we have

$$\bar{q}\propto\left(1+\mathbf{b}_1\cdot\mathbf{r}_1,\mathbf{b}_1\times\mathbf{r}_1+\frac{\sin\left(\frac{\phi}{2}\right)}{\cos\left(\frac{\phi}{2}\right)}(\mathbf{b}_1+\mathbf{r}_1)\right)$$
$$\propto\left((1+\mathbf{b}_1\cdot\mathbf{r}_1),\mathbf{b}_1\times\mathbf{r}_1-\frac{(\mathbf{b}_1\times\mathbf{r}_1)\cdot(\mathbf{b}_2-\mathbf{r}_2)}{(\mathbf{b}_1+\mathbf{r}_1)\cdot(\mathbf{b}_2-\mathbf{r}_2)}(\mathbf{b}_1+\mathbf{r}_1)\right)$$
$$\propto\left((\mathbf{b}_1+\mathbf{r}_1)\cdot(\mathbf{b}_2-\mathbf{r}_2),\frac{(\mathbf{b}_1\times\mathbf{r}_1)((\mathbf{b}_1+\mathbf{r}_1)\cdot(\mathbf{b}_2-\mathbf{r}_2))-((\mathbf{b}_1\times\mathbf{r}_1)\cdot(\mathbf{b}_2-\mathbf{r}_2))(\mathbf{b}_1+\mathbf{r}_1)}{(1+\mathbf{b}_1\cdot\mathbf{r}_1)}\right) \tag{108}$$

In view of (84), the scalar part of (108) implies

$$(\mathbf{b}_1+\mathbf{r}_1)\cdot(\mathbf{b}_2-\mathbf{r}_2)=\mathbf{b}_2\cdot\mathbf{r}_1-\mathbf{b}_1\cdot\mathbf{r}_2. \tag{109}$$

For the numerator of the vector part of (108), using (2), (1), and the fact that $\mathbf{b}_1$ and $\mathbf{r}_1$ are unit vectors, we have

$$\begin{aligned}&(\mathbf{b}_1\times\mathbf{r}_1)((\mathbf{b}_1+\mathbf{r}_1)\cdot(\mathbf{b}_2-\mathbf{r}_2))-((\mathbf{b}_1\times\mathbf{r}_1)\cdot(\mathbf{b}_2-\mathbf{r}_2))(\mathbf{b}_1+\mathbf{r}_1)\\&=(\mathbf{b}_2-\mathbf{r}_2)\times((\mathbf{b}_1\times\mathbf{r}_1)\times(\mathbf{b}_1+\mathbf{r}_1))\\&=(\mathbf{b}_2-\mathbf{r}_2)\times(\mathbf{r}_1-(\mathbf{r}_1\cdot\mathbf{b}_1)\mathbf{b}_1+(\mathbf{b}_1\cdot\mathbf{r}_1)\mathbf{r}_1-\mathbf{b}_1)\\&=(\mathbf{b}_2-\mathbf{r}_2)\times(\mathbf{r}_1-\mathbf{b}_1+(\mathbf{r}_1-\mathbf{b}_1)(\mathbf{r}_1\cdot\mathbf{b}_1))\\&=((\mathbf{b}_2-\mathbf{r}_2)\times(\mathbf{r}_1-\mathbf{b}_1))(1+\mathbf{r}_1\cdot\mathbf{b}_1)\\&=((\mathbf{b}_1-\mathbf{r}_1)\times(\mathbf{b}_2-\mathbf{r}_2))(1+\mathbf{r}_1\cdot\mathbf{b}_1)\end{aligned} \tag{110}$$

Combining (108)–(110) yields

$$\bar{q}\propto(\mathbf{b}_2\cdot\mathbf{r}_1-\mathbf{b}_1\cdot\mathbf{r}_2,(\mathbf{b}_1-\mathbf{r}_1)\times(\mathbf{b}_2-\mathbf{r}_2)).$$

Normalizing the right-hand side gives

$$\bar{q}=\frac{(\mathbf{b}_2\cdot\mathbf{r}_1-\mathbf{b}_1\cdot\mathbf{r}_2,(\mathbf{b}_1-\mathbf{r}_1)\times(\mathbf{b}_2-\mathbf{r}_2))}{\sqrt{(\mathbf{b}_2\cdot\mathbf{r}_1-\mathbf{b}_1\cdot\mathbf{r}_2)^2+\|(\mathbf{b}_1-\mathbf{r}_1)\times(\mathbf{b}_2-\mathbf{r}_2)\|^2}} \tag{111}$$

Therefore, given known ephemeris $\mathbf{r}_1$ and $\mathbf{r}_2$, observations $\mathbf{b}_1$ and $\mathbf{b}_2$, the attitude quaternion is uniquely defined. The attitude quaternion is extremely simple though the derivation is tedious.

4.5. Analytic formula for quaternion estimation

Although all flight experiences were successful for QUEST method, using a specific example, Markley and Mortari (2000) demonstrated that QUEST may not converge. In fact, it is well known that Newton's method is inadequate for general use since it may fail to converge to a solution. Even if it does converge, its behavior may be erratic in regions where the function is not convex (Nocedal & Wright, 1999). On the other hand, (94) is a polynomial of degree 4 which admits analytic solutions. In this section, we provide the analytic solution based on the characteristic polynomial of the K-matrix presented in Mortari (1997) which is given as follows.

$$p(x)=x^4+ax^3+bx^2+cx+d=0, \tag{112}$$

where $a=0, b=-2(tr[B]+tr[adj(S)]-z^Tz, S=B+B^T, adj(S)$ the adjugate matrix of $S, c=-tr[adj(K)]$, and $d=det(K)$ are all known parameters. Several different methods were proposed in the last several hundred years (Herbison-Evans, 1994) to solve (112). Recently, Shmakov (2011) found a universal method to find the roots of the general quartic polynomial. A special case of this method is simpler than all previous methods and it can be directly adopted to solve (112). We summarize the steps as follows.

First, (112) can be factorized as the product of two quadratic polynomials as

$$\begin{aligned}&(x^2+g_1x+h_1)(x^2+g_2+h_2)\\&=x^4+(g_1+g_2)x^3+(g_1g_2+h_1+h_2)x^2\\&\quad+(g_1h_2+g_2h_1)x+h_1h_2=0.\end{aligned} \tag{113}$$

Moreover, g_1, g_2, h_1, and h_2 are solutions of two quadratic equations defined by

$$g^2-ag+\frac{2}{3}b-y=0 \tag{114a}$$

$$h^2-\left(y+\frac{b}{3}\right)h+d=0 \tag{114b}$$

where y is the real root(s) of the following cubic polynomial

$$y^3+py+q=y^3+\left(ac-\frac{b^2}{3}-4d\right)y+\left(\frac{abc}{3}-a^2d-\frac{2}{27}b^3-c^2+\frac{8}{3}bd\right)=0. \tag{115}$$

The roots of the cubic equation can be obtained by the famous Cardano's formula (Polyanin & Manzhirov, 2007)

$$y_1 = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} \tag{116a}$$

$$y_2 = \omega_1 \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} + \omega_2 \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} \tag{116b}$$

$$y_3 = \omega_2 \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} + \omega_1 \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} \tag{116c}$$

where $\omega_1 = \frac{-1+i\sqrt{3}}{2}$ and $\omega_2 = \frac{-1-i\sqrt{3}}{2}$. Given a real y, from (114b), we have

$$g_{1,2} = \pm\sqrt{y - \frac{2}{3}b}, \tag{117a}$$

$$h_{1,2} = \frac{y + \frac{b}{3} \pm \sqrt{(y+b/3)^2 - 4d}}{2} \tag{117b}$$

and finally, the roots of the quartic (112) are given by

$$x_{1,2} = \frac{-g_1 \pm \sqrt{g_1^2 - 4h_1}}{2}, \tag{118a}$$

$$x_{3,4} = \frac{-g_2 \pm \sqrt{g_2^2 - 4h_2}}{2}. \tag{118b}$$

The rest problem is how to get ephemeris and observation vectors. These vector pairs can be any astronomical vectors, such as the Sun vector pairs, the Earth vector pairs, the Earth's magnet vector pairs, any star vector pair. There are a lot of literature that discusses this. For example, for the sun direction measurement, one can read (Liebe & Mobasser, 2001). For the ephemeris sun direction, the formula is given in Vallado (2004). For geomagnetic vector measurement, a magnetometer can be used (Wikipedia, 2011). For the ephemeris geomagnetic vector, the formula is given in Wertz (1978). For star tracker and algorithms, one can read (Jer-Nan Juang, Hye Young Kim, & Junkins, 2003). The details of these topics are out of the scope of this review.

4.6. Rotation rate determination using vector measurements

The rotation rate of the spacecraft may be needed in the feedback controller design. Many spacecrafts have equipped with on-board three axis rate-gyros to measure the angular rate (Fortescue, Stark, & Swinerd, 2003). But some spacecrafts do not install the rate-gyros for the economical reason. In this case, angular rate can be estimated using vector measurements, for example, the method published in Singla, L Crassida, and Junkins (2003). In this section, we present a very simple method.

Let

$$E = \begin{bmatrix} -q_1 & q_0 & q_3 & -q_2 \\ -q_2 & -q_3 & q_0 & q_1 \\ -q_3 & q_2 & -q_1 & q_0 \end{bmatrix}. \tag{119}$$

Pre-multiplying $2E$ on both sides of (46a) gives,

$$\begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} = 2\begin{bmatrix} -q_1 & q_0 & q_3 & -q_2 \\ -q_2 & -q_3 & q_0 & q_1 \\ -q_3 & q_2 & -q_1 & q_0 \end{bmatrix} \begin{bmatrix} \frac{dq_0}{dt} \\ \frac{dq_1}{dt} \\ \frac{dq_2}{dt} \\ \frac{dq_3}{dt} \end{bmatrix} = 2E\frac{d\bar{q}}{dt}. \tag{120}$$

In theory, after we get the quaternion, then take the difference between the current $\bar{q}(t_i)$ and the previous $\bar{q}(t_{i-1})$ and divided by the $\Delta t = t_i - t_{i-1}$ to approximate $\frac{d\bar{q}}{dt}$, we can get the angular rate. However, in practical application, due to the measurement noise, this angular rate determination based on the differentiation may not be accurate because of high frequency noise. A low pass Butterworth digital filter (Oppenheim, Schafer, & Buck, 1999), whose input is the ω obtained from (120) and the output is a better estimation of the angular rate, will significantly suppress the noise and thereby improving the angular rate determination.

5. Attitude control

Control design methods based on quaternion spacecraft model have been investigated for decades. As we mentioned in Section 3, most quaternion based design methods use Lyapunov functions and focus on the global stability; they pay little attention to the control system performance which is important in practical system design. Only a few researchers considered the performance of the quaternion based control systems. Using classical frequency domain method, Paielli and Bach (1993) adopted quaternion based linear error dynamics to get desired performance for the attitude control system; Wie, Weiss, and Arapostathis (1989) showed that there exists some state feedback that globally stabilizes the nonlinear spacecraft system and the feedback matrix assigns the closed loop poles for the dynamics described by the rotational angle about the rotational axis. These methods are in classical domain and they are not easy to extend to modern designs. Zhou and Colgren (2005) obtained a linearized state space model with all components of the quaternion in the state variables. However, this linearized state space model is not fully controllable. This explains why many powerful design methods in linear control system theory such as pole assignment, LQR control, and H_∞ control cannot be directly applied to the spacecraft control system design if full quaternion based linearized model is used.

On the other hand, although the Euler angle representation has a singular point and the representation is dependent on the rotational sequential, the linearized Euler angle based spacecraft model has been proved to be fully controllable. Therefore, all linear system design methods can be directly applied to spacecraft control system design for the Euler angle model and these methods are described in many standard text books, for example, Wie (1998), Wertz (1978), and Sidi (1997). More importantly, there are many successful applications of using these powerful control design methods, for example, Stoltz, Sivapiragasam, and Anthony (1998) and Won (1999).

In Section 3, we showed that the reduced quaternion model that uses only vector components of the quaternion is fully controllable. Moreover, the linearized reduced quaternion models have some simple and special structure. In this section, we will consider the design methods based on the reduced quaternion models obtained in Section 3. For nadir pointing spacecraft, one can directly use standard linear control system design methods, such as LQR design (Athans & Falb, 1966), robust pole assignment design (Kautsky, Nichols, & Van Dooren, 1985; Tits & Yang, 1996; Yang, 1996), and H_∞ design (Doyle, Glover, Khargonekar, & Francis, 1989), for the linearized system. The designed controller can then be checked by simulation with the original nonlinear spacecraft system. For inertial pointing spacecraft, since the linearized system has a very simple structure, using this linearized reduced quaternion model, one can derive an analytical formula for LQR optimal control that is explicitly related to the cost matrices Q and R. Moreover, it can be shown that under some mild restriction, the LQR feedback controller globally stabilizes the original nonlinear spacecraft. In addition, the LQR controller has a diagonal structure in the state feedback matrices D and K, this structure is actually a robust pole assignment design. The main results presented here are from Athans and Falb (1966), Yang (2012), and Yang (2012).

5.1. LQR design for general linear system

We first consider the general linear system described as follows.

$$\dot{x} = Ax + Bu,\qquad y = Cx. \tag{121}$$

The LQR design is to find a state feedback matrix

$$u = -[D, K]x = -Gx$$

to minimize the following cost function

$$L = \frac{1}{2}\int_0^\infty (x^T Qx + u^T Ru)dt. \tag{122}$$

where Q and R are positive definite matrices, $x^T Qx$ represents the cost of the performance, $u^T Ru$ represents the cost of the fuel consumption. The LQR control problem was first considered by HALL (1943) and WIENER (1949), but Kalman (1960) provided a much better solution and popularized the design. If Kalman filter (Kalman, 1960) is used as part of the feedback loop, then the control design method is the LQG control. Surprisingly, Kalman filter and LQR control law can be designed separately because of the separation theorem obtained by Wonham (1968).

The optimal control of *LQR* design is uniquely given by (see Athans & Falb, 1966)

$$u(t) = -R^{-1}B^T Fx(t) = -Gx, \tag{123}$$

where F is a constant positive definite matrix which is the solution of the Lyapunov matrix algebraic equation

$$-FA - A^T F + FBR^{-1}B^T F - Q = 0. \tag{124}$$

This control law can be directly used for the nadir pointing spacecraft without any modification. For inertial pointing sapcecarft, due to the simple structure of the linearized reduced quaternion model, analytic solution to LQR design can be obtained.

5.2. The LQR design for inertial pointing spacecraft

In this subsection, we consider LQR design for inertial pointing spacecrafts for which A and B are defined in (59). We assume further that the constant inertia matrix of the spacecraft J defined in (47) is diagonal. This assumption is reasonable because in practical spacecraft design, J is always designed close to a diagonal matrix. In the rest of the discussion of this subsection, we assume further that Q, and R are diagonal matrices because Q and R are oftentimes selected to be diagonal in engineering design practice. With these assumptions, the problem can greatly be simplified.

5.2.1. The analytic solution

It is well known that the LQR feedback (123) guarantees the stability of the linearized closed loop system and minimize the cost function of (122) that is a combined cost of cumulative control system error and cumulative fuel consumption.

First we derive the analytical solution for the spacecraft model (58). Let

$$F = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix},\quad Q = \begin{bmatrix} Q_{11} & 0 \\ 0 & Q_{22} \end{bmatrix}, \tag{125}$$

where the elements of F and Q in (125) are all 3 by 3 matrices. Substituting A and B defined in (59), F and Q defined in (125) into (124), after simple manipulations, we get

$$\begin{bmatrix} F_{11}J^{-1}R^{-1}J^{-1}F_{11} & F_{11}J^{-1}R^{-1}J^{-1}F_{12} \\ F_{12}^{-1}J^T R^{-1}J^{-1}F_{11} & F_{12}^T J^{-1}R^{-1}J^{-1}F_{12} \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\left(F_{12}^T + F_{12}\right) + Q_{11} & \frac{1}{2}F_{22} \\ \frac{1}{2}F_{22} & Q_{22} \end{bmatrix}. \tag{126}$$

Since J, Q and R are positive definite, noticing that $F_{21}^T = F_{12}$, comparing the (2,2) block on both sides of (126) yields,

$$F_{12} = JR^{\frac{1}{2}}Q_{22}^{\frac{1}{2}}. \tag{127}$$

Since J, $Q_{11} = \mathrm{diag}(q_{1i})$, $Q_{22} = \mathrm{diag}(q_{2i})$, and $R = \mathrm{diag}(r_i)$ are diagonal, we conclude that F_{12} is diagonal. Substituting (127) into the (1,1) block of (126) gives,

$$F_{11} = JR^{\frac{1}{2}}\left(Q_{11} + \frac{1}{2}\left(JR^{\frac{1}{2}}Q_{22}^{\frac{1}{2}} + Q_{22}^{\frac{1}{2}}R^{\frac{1}{2}}J\right)\right)^{\frac{1}{2}}. \tag{128}$$

Therefore, F_{11} is diagonal. Substituting (127) and (128) into the (2,1) block of (126) gives

$$F_{22} = 2Q_{22}^{\frac{1}{2}}\left(Q_{11} + JR^{\frac{1}{2}}Q_{22}^{\frac{1}{2}}\right)^{\frac{1}{2}}, \tag{129}$$

which is also diagonal. Eqs. (127)–(129) give a complete solution of Lyapunov matrix Eq. (124). Therefore, (123) can be rewritten as

$$u(t) = -R^{-1}B^T Fx(t) = -[R^{-1}J^{-1}F_{11}, R^{-1}J^{-1}F_{12}]x = -[D, K]x. \tag{130}$$

Clearly, matrices D and K are diagonal.

5.2.2. The global stability of the design

Next, we show that under some additional conditions, the LQR optimal control given by (130) globally stabilizes the nonlinear system described by (56) and (52). Let $P = Q_{22}^{-\frac{1}{2}}R^{\frac{1}{2}}J$, and the Lyapunov function be

$$V = \frac{1}{2}\omega_I^T P\omega_I + q_1^2 + q_2^2 + q_3^2 + (1 - q_0)^2. \tag{131}$$

Then, in view of (123), the derivative of the Lyapunov function along the trajectory described by the nonlinear system Eqs. (56) and (52) is given by

$$\begin{aligned} \frac{dV}{dt} &= \omega_I^T P\left(-J^{-1}\omega_I \times J\omega_I - J^{-1}R^{-1}[J^{-1}\ 0]\begin{bmatrix} F_{11} & F_{12} \\ F_{12}^T & F_{22} \end{bmatrix}\begin{bmatrix} \omega_I \\ q \end{bmatrix}\right) + \omega_I^T q \\ &= -\omega_I^T PJ^{-1}\omega_I \times J\omega_I - \omega_I^T PJ^{-1}R^{-1}J^{-1}F_{11}\omega_I - \omega_I^T PJ^{-1}R^{-1}J^{-1}F_{12}q + \omega_I^T q \\ &= -\omega_I^T PJ^{-1}\omega_I \times J\omega_I - \omega_I^T Q_{22}^{-\frac{1}{2}}\left(Q_{11} + \frac{1}{2}\left(JR^{\frac{1}{2}}Q_{22}^{\frac{1}{2}} + Q_{22}^{\frac{1}{2}}R^{\frac{1}{2}}J\right)\right)^{\frac{1}{2}}\omega_I \\ &= -\omega_I^T Q_{22}^{-\frac{1}{2}}R^{\frac{1}{2}}\omega_I \times J\omega_I - \omega_I^T Q_{22}^{-\frac{1}{2}}\left(Q_{11} + \frac{1}{2}\left(JR^{\frac{1}{2}}Q_{22}^{\frac{1}{2}} + Q_{22}^{\frac{1}{2}}R^{\frac{1}{2}}J\right)\right)^{\frac{1}{2}}\omega_I \end{aligned} \tag{132}$$

Since P, Q, R, and J are all diagonal positive definite matrices, the second term of the last expression is negative definite. If $Q_{22}^{-1}R = cI$ i.e.,

$$R = cQ_{22} \tag{133}$$

or $Q_{22}^{-1}R = cJ$, i.e.,

$$R = cQ_{22}J, \tag{134}$$

where c is a constant, then the first term is vanishing, therefore $\frac{dV}{dt}$ is negative semi-definite, and the nonlinear system described by (56) and (52) is globally stable with the optimal controller given by (130). To show that the closed loop nonlinear system is asymptotically stable, we define $S = \{x | \dot{V}(x) = 0\}$. Since D and K are full rank matrices, clearly $S = \{x : x = (\omega_I, q) = (0, q)\}$. From (56), since $u = -D\omega_I - Kq \neq 0$ if $q \neq 0$, no solution can always stay in S except $S = \{x = (\omega_I, q) = (0, 0)\}$. Using Corollary 3.2 in Khalil (1992), the origin is globally asymptotically stable. By definition Khalil (1992, p. 111), the region of attraction of the nonlinear system is the whole space spanned by x.

In system design practice, if the performance and the local stability are the only design considerations, Q and R can be chosen without any restriction; if the global stability is also required for the nonlinear spacecraft system, some restriction, though it is mild, should be placed on Q and R, i.e., either $R = cQ_{22}$ or $R = cQ_{22}J$, where c is any positive constant.

5.2.3. The closed-loop poles

To establish the relationship between the closed loop poles and the design matrices Q, and R, we can simplify (130) further as follows.

$$D = R^{-\frac{1}{2}}\left(Q_{11} + \frac{1}{2}\left(JR^{\frac{1}{2}}Q_{22}^{\frac{1}{2}} + Q_{22}^{\frac{1}{2}}R^{\frac{1}{2}}J\right)\right)^{\frac{1}{2}} = \mathrm{diag}(d_i)$$
$$= \mathrm{diag}\left(\sqrt{\frac{q_{1i}}{r_i} + J_{ii}\sqrt{\frac{q_{2i}}{r_i}}}\right) \tag{135}$$

with

$$d_i = \sqrt{\frac{q_{1i}}{r_i} + J_{ii}\sqrt{\frac{q_{2i}}{r_i}}},$$

and

$$K = R^{-\frac{1}{2}}Q_{22}^{\frac{1}{2}} = \mathrm{diag}(k_i) = \mathrm{diag}\left(\sqrt{\frac{q_{2i}}{r_i}}\right) \tag{136}$$

with

$$k_i = \sqrt{\frac{q_{2i}}{r_i}}.$$

Therefore, (130) becomes

$$u(x) = -[D, K]x$$
$$= -\begin{bmatrix} \sqrt{\frac{q_{11}}{r_1} + J_{11}\sqrt{\frac{q_{21}}{r_1}}} & 0 & 0 & \sqrt{\frac{q_{21}}{r_1}} & 0 & 0 \\ 0 & \sqrt{\frac{q_{12}}{r_2} + J_{22}\sqrt{\frac{q_{22}}{r_2}}} & 0 & 0 & \sqrt{\frac{q_{22}}{r_2}} & 0 \\ 0 & 0 & \sqrt{\frac{q_{13}}{r_3} + J_{33}\sqrt{\frac{q_{23}}{r_3}}} & 0 & 0 & \sqrt{\frac{q_{23}}{r_3}} \end{bmatrix} \begin{bmatrix} \omega_{I1} \\ \omega_{I2} \\ \omega_{I3} \\ q_1 \\ q_2 \\ q_3 \end{bmatrix}. \tag{137}$$

From (58), it is straightforward to write the closed loop system as follows

$$\begin{bmatrix} \frac{d\omega_I}{dt} \\ \frac{dq}{dt} \end{bmatrix} = \begin{bmatrix} -J^{-1}R^{-\frac{1}{2}}\left(Q_{11} + + \frac{1}{2}\left(JR^{\frac{1}{2}}Q_{22}^{\frac{1}{2}} + Q_{22}^{\frac{1}{2}}R^{\frac{1}{2}}J\right)\right)^{\frac{1}{2}} & -J^{-1}R^{-\frac{1}{2}}Q_{22}^{\frac{1}{2}} \\ \frac{1}{2}I_3 & 0_3 \end{bmatrix} \begin{bmatrix} \omega_I \\ q \end{bmatrix}$$
$$= \begin{bmatrix} -\frac{d_1}{J_{11}} & 0 & 0 & -\frac{k_1}{J_{11}} & 0 & 0 \\ 0 & -\frac{d_2}{J_{22}} & 0 & 0 & -\frac{k_2}{J_{22}} & 0 \\ 0 & 0 & -\frac{d_3}{J_{33}} & 0 & 0 & -\frac{k_3}{J_{33}} \\ 0.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \omega_{I1} \\ \omega_{I2} \\ \omega_{I3} \\ q_1 \\ q_2 \\ q_3 \end{bmatrix}$$
$$= \bar{A}x. \tag{138}$$

For $i = 1, 2$, and 3, let $s_i = \frac{d_i}{J_{ii}}, t_i = \frac{k_i}{J_{ii}}$, and

$$C_i = \frac{\frac{d_i}{J_{ii}} + \sqrt{\left(\frac{d_i}{J_{ii}}\right)^2 - 2\frac{k_i}{J_{ii}}}}{2\frac{k_i}{J_{ii}}} = \frac{s_i + \sqrt{s_i^2 - 2t_i}}{2t_i}. \tag{139}$$

Then,

$$\bar{A} = \begin{bmatrix} -s_1 & 0 & 0 & -t_1 & 0 & 0 \\ 0 & -s_2 & 0 & 0 & -t_2 & 0 \\ 0 & 0 & -s_3 & 0 & 0 & -t_3 \\ 0.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0 \end{bmatrix} \tag{140}$$

Let the linear matrix transformation $T_{ij}(C)$ be a matrix with the following properties: (a) the (i,j) element of $T_{ij}(C)$ is C, (b) the diagonal elements are ones, (c) all the remaining elements are zeros. It is well known that the inverse of $T_{ij}(C)$ is $T_{ij}^{-1}(C) = T_{ij}(-C)$. Premultiplying $T_{41}(C_1)$ to $\bar{A}$ is equivalent to multiply the first row of $\bar{A}$ by C_1 and add this result to the 4th row of the matrix. This gives

$$T_{41}(C_1)\bar{A} = \begin{bmatrix} -s_1 & 0 & 0 & -t_1 & 0 & 0 \\ 0 & -s_2 & 0 & 0 & -t_2 & 0 \\ 0 & 0 & -s_3 & 0 & 0 & -t_3 \\ -\frac{s_1^2 + s_1\sqrt{s_1^2 - 2t_1}}{2t_1} + 0.5 & 0 & 0 & -\frac{s_1 + \sqrt{s_1^2 - 2t_1}}{2} & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0 \end{bmatrix}. \tag{141}$$

Post-multiplying $T_{41}(-C_1)$ to this matrix is equivalent to multiply the 4th column by $-C_1$ and add this result to the first column of the matrix. This gives, noticing that $-s_1 + \frac{t_1\left(s_1 + \sqrt{s_1^2 - 2t_1}\right)}{2t_1} = \frac{-s_1 + \sqrt{s_1^2 - 2t_1}}{2}$ and $\frac{s_1 + \sqrt{s_1^2 - 2t_1}}{2}\,\frac{s_1 + \sqrt{s_1^2 - 2t_1}}{2t_1} = \frac{s_1^2 + s_1\sqrt{s_1^2 - 2t_1}}{2t_1} - 0.5$,

$$T_{41}(C_1)\bar{A}T_{41}(-C_1) = \begin{bmatrix} \frac{-s_1 + \sqrt{s_1^2 - 2t_1}}{2} & 0 & 0 & -t_1 & 0 & 0 \\ 0 & -s_2 & 0 & 0 & -t_2 & 0 \\ 0 & 0 & -s_3 & 0 & 0 & -t_3 \\ 0 & 0 & 0 & \frac{-s_1 - \sqrt{s_1^2 - 2t_1}}{2} & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0 \end{bmatrix}. \tag{142}$$

Repeating similar manipulation, we have

$$T_{63}(C_3)T_{52}(C_2)T_{14}(C_1)\bar{A}T_{41}(-C_1)T_{52}(-C_2)T_{63}(-C_3)$$
$$= \begin{bmatrix} \frac{-s_1 + \sqrt{s_1^2 - 2t_1}}{2} & 0 & 0 & -t_1 & 0 & 0 \\ 0 & \frac{-s_i + \sqrt{s_2^2 - 2t_2}}{2} & 0 & 0 & -t_2 & 0 \\ 0 & 0 & \frac{-s_3 + \sqrt{s_3^2 - 2t_3}}{2} & 0 & 0 & -t_3 \\ 0 & 0 & 0 & \frac{-s_1 - \sqrt{s_1^2 - 2t_1}}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{-s_2 - \sqrt{s_2^2 - 2t_2}}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{-s_3 - \sqrt{s_3^2 - 2t_3}}{2} \end{bmatrix}. \tag{143}$$

Therefore, the closed loop eigenvalues of linear system (138) are given by, for $i = 1, 2$, and 3,

$$\lambda_i, \lambda_{i+3} = \frac{-\sqrt{\frac{1}{J_{ii}}\sqrt{\frac{q_{2i}}{r_i}} + \frac{q_{1i}}{J_{ii}^2 r_i}} \pm \sqrt{\frac{q_{1i}}{J_{ii}^2 r_i} - \frac{1}{J_{ii}}\sqrt{\frac{q_{2i}}{r_i}}}}{2}. \tag{144}$$

Eq. (144) provides a lot of useful information for the LQR design. First, as $r_i \to 0$, the corresponding pair of eigenvalues go to minus infinity of the complex plane; as $r_i \to \infty$, the corresponding pair of eigenvalues go to origin of the complex plane. Second, As long as $q_{1i} > \sqrt{q_{2i} r_i} J_{ii}$, the corresponding pair of eigenvalues are real and

Table 1
Required closed-loop poles.

−0.01273212110421 ± 0.01272387326295i;
−0.00798572833825 ± 0.00798369205833i;
−0.00947996395486 ± 0.00947655794419i.

unequal; since $\frac{d_i}{J_{ii}} > \sqrt{\left(\frac{d_i}{J_{ii}}\right)^2 - 2\frac{k_i}{J_{ii}}}$, these two eigenvalues are always negative. Third, if $q_{1i} = \sqrt{q_{2i} r_i} J_{ii}$, there are two equal real negative eigenvalues. Fourth, if $q_{1i} < \sqrt{q_{2i} r_i} J_{ii}$, there is a pair of complex eigenvalues with negative real part. Therefore increasing q_{1i} and decreasing q_{2i} will increase the dumping ratio; otherwise, it will decrease the dumping ratio. Finally, increasing q_{2i} and decreasing r_i will increase the natural frequency; otherwise, it will decrease the natural frequency. This information can be useful in spacecraft system design.

Using the LQR design, we implicitly assign the closed loop poles as defined by (138) and we can balance the requirements on accumulative control error and power consumption (both are important in practical design).

5.2.4. A simulation result

We use an example in Zhou and Colgren (2005) to illustrate the design procedure. The spacecraft inertia matrix is give by

$$J = \begin{bmatrix} 1200 & 100 & -200 \\ 100 & 2200 & 300 \\ -200 & 300 & 3100 \end{bmatrix} \tag{145}$$

It is clear that the diagonal elements of the matrix are significantly larger than off-diagonal elements. Assume that the spacecraft inertia matrix can be approximate by a diagonal matrix whose diagonal elements are equal to these of J, let $Q = \mathrm{diag}(5,5,5,5,5,5)$ and $R = \mathrm{diag}(8,8,8)$, the closed loop poles are then given as in Table 1, and the feedback matrix D and K are as follows

$$D = \begin{bmatrix} 31.0663 & 0 & 0 \\ 0 & 41.7118 & 0 \\ 0 & 0 & 49.5115 \end{bmatrix}, \tag{146}$$

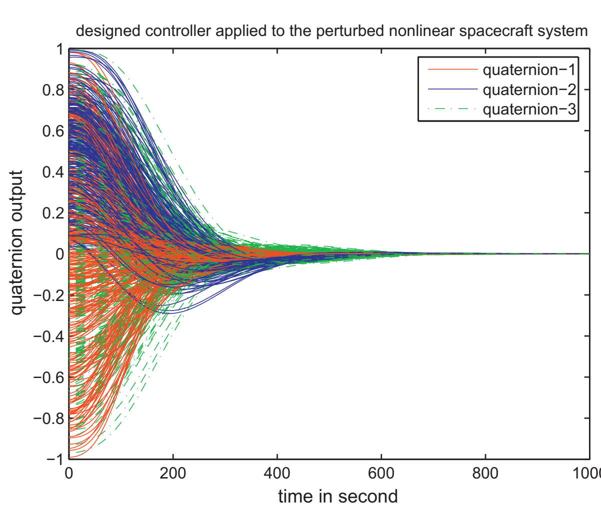

Fig. 6. Monte Carlo simulation for the nonlinear spacecraft model with perturbation.

$$K = \begin{bmatrix} 0.7905 & 0 & 0 \\ 0 & 0.7905 & 0 \\ 0 & 0 & 0.7905 \end{bmatrix}. \tag{147}$$

We apply the designed feedback controller to the nonlinear spacecraft system described by (50) and (52) with the full Monte Carlo perturbation model described as follows: (a) in inertia matrix J, the off-diagonal elements are randomly selected between [0,310], (b) the initial Euler angle errors of the nonlinear spacecraft system are randomly selected between $[0,\pi]$ and these initial Euler angles are converted into quaternion, and (c) the initial angular rates are randomly selected between [0,0.1] deg/second, and we conduct 300 Monte Carlo simulation runs; the simulated runs are all asymptotically stable. This result is shown in Fig. 6.

5.3. Robust pole assignment for inertial pointing spacecraft

5.3.1. Robustness of the closed-loop poles

In the previous subsection, we have derived a simple analytic LQR control design method. The closed loop eigenvalues are explicitly related to the spacecraft inertia matrix, the selected Q and R matrices. In this subsection, we will show that the LQR design is equivalent to the pole assignment design. Furthermore, we will show that the pole assignment design is a robust pole assignment design. Therefore, the LQR design is equivalent to a robust pole assignment design which is insensitive to the modeling error.

First, from (143) and (139), it is easy to see that

$$\lambda_i, \lambda_{i+3} = \frac{-\frac{d_i}{J_{ii}} \pm \sqrt{\left(\frac{d_i}{J_{ii}}\right)^2 - 2\frac{k_i}{J_{ii}}}}{2}. \tag{148}$$

Let the desired spacecraft closed-loop eigenvalues be given by

$$\lambda_i, \lambda_{i+3} = -\zeta_i \omega_{in} \pm j\omega_{in}\sqrt{1-\zeta_i^2} = -\zeta_i \omega_{in} \pm j\omega_{id}. \tag{149}$$

Comparing (148) and (149) yields the analytic feedback controller

$$k_i = 2\omega_{in}^2 J_{ii}, \tag{150}$$
$$d_i = 2\zeta_i \omega_{in} J_{ii}. \tag{151}$$

Therefore, for any LQR design which minimizes (122), there is an implicit set of desired spacecraft closed-loop eigenvalues defined by (144) or (148) or (149), the diagonal feedback matrices D and K with diagonal elements given by (150) and (151) assign the prescribed closed-loop eigenvalues. It is shown in the previous subsection that the closed-loop nonlinear system is globally asymptotically stable if some additional condition holds.

It is well known that for any controllable linear system and for any prescribed closed-loop pole locations, one can always find a state feedback controller such that the closed-loop system has the prescribed pole locations. For multi-input systems, the solution that achieves the closed-loop pole positions is not unique. As an example, let (A,B) be a linear system with

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The open-loop system has two eigenvalues (0,1) and the system is not stable. Assuming that the desired close-loop eigenvalues are $(-1,-1)$, one may select two different feedback matrices

$$G_1 = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix}, G_2 = \begin{bmatrix} 1 & 4*10^{10} \\ -10^{-10} & -4 \end{bmatrix}.$$

such that

$$A + BG_1 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix},$$

$$A + BG_2 = \begin{bmatrix} 1 & 4 * 10^{10} \\ -10^{-10} & -3 \end{bmatrix}.$$

It is easy to verify that $\det(\lambda I - (A + BG_1)) = \det(\lambda I - (A + BG_2)) = (\lambda + 1)^2$. Both feedbacks achieve the desired closed-loop poles. The first system is robust because any small perturbation will not destabilize the system. However, the second system is not robust as a small perturbation of 10^{-10} in the left low corner of the matrix $A + BG_2$ will change the closed-loop eigenvalues to $(1, -3)$. We show that the LQR defined pole assignment is a robust pole assignment.

5.3.2. The robust pole assignment

The robust pole assignment design makes full use of the extra degrees of freedom in a multi-input system to find the most robust controller from indefinitely many solutions of the pole assignment feedback matrices. Since the spacecraft attitude control system is a typical multi-input system that has three control torque inputs (roll, pitch, and yaw), getting a robust pole assignment design that is insensitive to the modeling error is very attractive and desirable. We will show that the controller with diagonal D and K proposed in the previous subsection is a robust pole assignment design.

There are many different robust indices that can be used in robust pole assignment (Wilkinson, 1965; Kautsky et al., 1985). We will adopt the robust measurement proposed in Yang (1989) as the design criterion because the algorithms based on this robust measurement lead to some efficient and effective design (Sima, Tits, & Yang, 2006). Let X be the matrix whose columns are the unit length closed-loop eigenvectors. The robustness of the closed-loop eigenvalues (poles) can be measured by the absolute value of the determinant of X. Geometrically, this determinant measures how close the matrix X is to an orthogonal matrix. The greater the absolute value of the determinant, the more robust the closed-loop eigenvalues will be (see detailed discussions in Yang (1989) and Yang (1996)). By maximizing the absolute value of the determinant under some constraints, we are guaranteed that the closed-loop poles obtained by the robust pole assignment design are insensitive to the modeling errors (Yang, 2012). For a controllable linear system (A,B), where B is full column rank, and any given set of desired closed-loop eigenvalues λ_i, the corresponding closed-loop eigenvectors x_i must be in the subspace (see Yang, 1996)

$$\mathcal{S}_i = \{x : (A - \lambda I)x \in \mathbf{R}_c(B)\}, \tag{152}$$

where

$$\mathbf{R}_c(B) = \{By : y \in \mathbf{C}^m\},$$

m is the rank of B, and $\mathbf{C}^m$ is a m-dimensional complex space. The proof is given in Kautsky et al. (1985). First, using QR decomposition on B, we have

$$B = [U_0\ U_1]\begin{bmatrix} V \\ 0 \end{bmatrix}.$$

Let Λ be the diagonal matrix whose diagonal elements are the desired eigenvalues, and X be the matrix whose columns are composed of the eigenvectors corresponding to the desired eigenvalues. Then,

$$BG = U_0VG = X\Lambda X^{-1} - A.$$

Pre-multiplication of U_0^{T} and U_1^{T} gives

$$VG = U_0^{\mathrm{T}}(X\Lambda X^{-1} - A) \tag{153a}$$

$$0 = U_1^{\mathrm{T}}(X\Lambda X^{-1} - A) \tag{153b}$$

The first relation gives the closed-loop feedback matrix as

$$G = V^{-1}U_0^{\mathrm{T}}(A - X\Lambda X^{-1}). \tag{154}$$

The second relation shows that x_i must be in the subspace $\mathcal{S}_i$, or

$$U_1^{\mathrm{T}}(A - \lambda_i I)x_i = 0.$$

Therefore, x_i must be in the null space of $(A^{\mathrm{T}} - \lambda_i I)U_1$. Using QR decomposition again on $(A^{\mathrm{T}} - \lambda_i I)U_1$ gives

$$(A^{\mathrm{T}} - \lambda_i I)U_1 = [W_{1i}\ W_{2i}]\begin{bmatrix} R \\ 0 \end{bmatrix}.$$

W_{2i} forms the basis of $\mathcal{S}_i$. We now apply the similar procedure to the linearized spacecraft system (58). Since B can be written as

$$B = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}\begin{bmatrix} J^{-1} \\ 0 \end{bmatrix},$$

therefore and

$$U_0 = \begin{bmatrix} I \\ 0 \end{bmatrix}, U_1 = \begin{bmatrix} 0 \\ I \end{bmatrix}, V = J^{-1}. \tag{155}$$

Since A is defined as in (59), we can write a similar decomposition of $(A^{\mathrm{T}} - \lambda_i I)U_1$ as

$$(A^{\mathrm{T}} - \lambda_i I)U_1 = \begin{bmatrix} -\lambda_i I & \frac{1}{2}I \\ 0 & -\lambda_i I \end{bmatrix}\begin{bmatrix} 0 \\ I \end{bmatrix} = \begin{bmatrix} \frac{1}{2}I \\ -\lambda_i I \end{bmatrix} = \begin{bmatrix} 0.5I & -\lambda_i I \\ -\lambda_i I & -0.5I \end{bmatrix}\begin{bmatrix} I \\ 0 \end{bmatrix}, \tag{156}$$

therefore,

$$W_{1i} = \begin{bmatrix} 0.5I \\ -\lambda_i I \end{bmatrix},$$

which is orthogonal to the subspace

$$W_{2i} = \begin{bmatrix} -\lambda_i I \\ -0.5I \end{bmatrix},$$

where W_{2i} forms the base of $\mathcal{S}_i$. Though $[W_{1i}\ W_{2i}]$ may not be a unitary matrix, it is clear that W_{2i} forms the basis of $\mathcal{S}_i$ (and we can always normalize W_{2i} to make it orthonormal). For the sake of simplicity, we prove that the design given by (137) is a robust pole assignment only for the case that all eigenvalues are real. For robust pole assignmnet design, since $x_i \in \mathcal{S}_i = W_{2i}$, the closed-loop eigenvector matrix must have the form,

$$X = \begin{bmatrix} \lambda_1 p_{11} & \lambda_2 p_{21} & \lambda_3 p_{31} & \lambda_4 p_{41} & \lambda_5 p_{51} & \lambda_6 p_{61} \\ \lambda_1 p_{12} & \lambda_2 p_{22} & \lambda_3 p_{32} & \lambda_4 p_{42} & \lambda_5 p_{52} & \lambda_6 p_{62} \\ \lambda_1 p_{13} & \lambda_2 p_{23} & \lambda_3 p_{33} & \lambda_4 p_{43} & \lambda_5 p_{53} & \lambda_6 p_{63} \\ 0.5p_{11} & 0.5p_{21} & 0.5p_{31} & 0.5p_{41} & 0.5p_{51} & 0.5p_{61} \\ 0.5p_{12} & 0.5p_{22} & 0.5p_{32} & 0.5p_{42} & 0.5p_{52} & 0.5p_{62} \\ 0.5p_{13} & 0.5p_{23} & 0.5p_{33} & 0.5p_{43} & 0.5p_{53} & 0.5p_{63} \end{bmatrix},$$

where p_{ij}, $i = 1, 2, 3, 4, 5, 6$, $j = 1, 2, 3$ are the real parameters that will be used to optimize the objective function. Therefore, the robust pole assignment design for linearized spacecraft system (11) becomes[3]

[3] In Yang (1996), $|\det(X)|$ is used as the measurement of the robustness. If the maximum of $|\det(X)|$ is achieved at $-\det(X^*)$, let X^0 be the matrix obtained by changing the sign of some column of X^*, $|\det(X)|$ is also achieved at X^0. Therefore, we can simply use $\det(X)$ here as the objective function in our problem.

$$\max \quad \det(X)$$
$$s.t. \quad \sum_{j=1}^{3}\left(|\lambda_i|^2+0.5^2\right)p_{ij}^2=1, \tag{157}$$
$$i=1,2,3,4,5,6.$$

It is well-known that an optimal solution for a general optimization problem has to satisfy the KKT condition (see Nocedal & Wright, 1999). For (157), let

$$\mathcal{L}=\det(X)-\mu_1\left(\sum_{j=1}^{3}\left(|\lambda_1|^2+0.5^2\right)p_{1j}^2-1\right)-\mu_2\left(\sum_{j=1}^{3}\left(|\lambda_2|^2+0.5^2\right)p_{2j}^2-1\right)$$
$$-\mu_3\left(\sum_{j=1}^{3}\left(|\lambda_3|^2+0.5^2\right)p_{3j}^2-1\right)-\mu_4\left(\sum_{j=1}^{3}\left(|\lambda_4|^2+0.5^2\right)p_{4j}^2-1\right)$$
$$-\mu_5\left(\sum_{j=1}^{3}\left(|\lambda_5|^2+0.5^2\right)p_{5j}^2-1\right)-\mu_6\left(\sum_{j=1}^{3}\left(|\lambda_6|^2+0.5^2\right)p_{6j}^2-1\right),$$

the KKT condition is as follows.

$$\frac{\partial\mathcal{L}}{\partial p_{ij}}=0,\ i=1,2,3,4,5,6,\ j=1,2,3, \tag{158a}$$

$$-\frac{\partial\mathcal{L}}{\partial \mu_1}=\sum_{j=1}^{3}(|\lambda_1|^2+0.5^2)p_{1j}^2-1=0, \tag{158b}$$

$$-\frac{\partial\mathcal{L}}{\partial \mu_2}=\sum_{j=1}^{3}(|\lambda_2|^2+0.5^2)p_{2j}^2-1=0, \tag{158c}$$

$$-\frac{\partial\mathcal{L}}{\partial \mu_3}=\sum_{j=1}^{3}(|\lambda_3|^2+0.5^2)p_{3j}^2-1=0. \tag{158d}$$

$$-\frac{\partial\mathcal{L}}{\partial \mu_4}=\sum_{j=1}^{3}(|\lambda_4|^2+0.5^2)p_{4j}^2-1=0, \tag{158e}$$

$$-\frac{\partial\mathcal{L}}{\partial \mu_5}=\sum_{j=1}^{3}(|\lambda_5|^2+0.5^2)p_{5j}^2-1=0, \tag{158f}$$

$$-\frac{\partial\mathcal{L}}{\partial \mu_6}=\sum_{j=1}^{3}(|\lambda_6|^2+0.5^2)p_{6j}^2-1=0. \tag{158g}$$

It is tedious but straightforward to verify that the following solution satisfies the KKT condition (see Yang, 2012).

$$\begin{cases} p_{i,i}=\sqrt{\frac{1}{|\lambda_i|^2+0.5^2}}, & i=j,\ \ i=1,2,3,\ \ j=1,2,3 \\ p_{i+3,j}=\sqrt{\frac{1}{|\lambda_{i+3}|^2+0.5^2}}, & i=j,\ \ i=1,2,3,\ \ j=1,2,3 \\ p_{i,j}=0, & i\neq j,\ i\neq j+3,\ \ i=1,2,3,4,5,6,\ \ j=1,2,3. \end{cases} \tag{159}$$

Clearly, this set of $p_{i,j}$ meets (158b)–(158g). To show that the set of $p_{i,i}$ satisfies (158a), we use the observation that $\frac{\partial det(X)}{\partial p_{ij}}=0$ for all p_{ij} defined in (159) except p_{11}, p_{22}, p_{33}, p_{41}, p_{52}, p_{63}, therefore, $\frac{\partial\mathcal{L}}{\partial p_{ij}}=0$ for all $p_{ij}\notin\{p_{11}, p_{22}, p_{33}, p_{41}, p_{52}, p_{63}\}$. As an example, let us consider $\frac{\partial\mathcal{L}}{\partial p_{12}}$, since

$$\frac{\partial det(X)}{\partial p_{12}}=\lambda\begin{vmatrix} 0 & 0 & \lambda_4 p_{41} & 0 & 0 \\ 0 & \lambda_3 p_{33} & 0 & 0 & \lambda_6 p_{63} \\ 0 & 0 & 0.5p_{41} & 0 & 0 \\ 0.5p_{22} & 0 & 0 & 0.5p_{52} & 0 \\ 0 & 0.5p_{33} & 0 & 0 & 0.5p_{63} \end{vmatrix} + 0.5\begin{vmatrix} 0 & 0 & \lambda_4 p_{41} & 0 & 0 \\ \lambda_2 p_{22} & 0 & 0 & \lambda_5 p_{52} & 0 \\ 0 & \lambda_3 p_{33} & 0 & 0 & \lambda_6 p_{63} \\ 0 & 0 & 0.5p_{41} & 0 & 0 \\ 0 & 0.5p_{33} & 0 & 0 & 0.5p_{63} \end{vmatrix}=0,$$

because the first row and the third row are proportional in the first determinant and the first row and the fourth row are proportional in the second determinant. Therefore, we have

$$\frac{\partial\mathcal{L}}{\partial p_{12}}=\frac{\partial det(X)}{\partial p_{12}}-2\mu_1 p_{12}(|\lambda_1|^2+0.5^2)|_{p_{12}=0}=0. \tag{160}$$

Similarly, we can use the same way to check all $p_{ij}\notin\{p_{11}, p_{22}, p_{33}, p_{41}, p_{52}, p_{63}\}$. For each of these 6 $p_{ij}\in\{p_{11},p_{22},p_{33},p_{41},p_{52},p_{63}\}$, $\frac{\partial det(X)}{\partial p_{ij}}\neq 0$, one can select one of the multipliers μ_1, μ_2, μ_3, μ_4, μ_5, μ_6 to make $\frac{\partial\mathcal{L}}{\partial p_{ij}}=0$. Therefore, the set of p_{ij} satisfying (159) is a candidate of the optimal solution of (157). This proves that the closed-loop eigenvector matrix have the form as

$$X=\begin{bmatrix} \lambda_1 p_{1,1} & 0 & 0 & \lambda_4 p_{4,1} & 0 & 0 \\ 0 & \lambda_2 p_{2,2} & 0 & 0 & \lambda_5 p_{5,2} & 0 \\ 0 & 0 & \lambda_3 p_{3,3} & 0 & 0 & \lambda_6 p_{6,3} \\ 0.5p_{1,1} & 0 & 0 & 0.5p_{4,1} & 0 & 0 \\ 0 & 0.5p_{2,2} & 0 & 0 & 0.5p_{5,2} & 0 \\ 0 & 0 & 0.5p_{3,3} & 0 & 0 & 0.5p_{6,3} \end{bmatrix}$$
$$=\begin{bmatrix} \mathrm{diag}(\lambda_i p_{i,i}) & \mathrm{diag}(\lambda_{i+3}p_{i+3,i}) \\ \mathrm{diag}(0.5p_{i,i}) & \mathrm{diag}(0.5p_{i+3,i}) \end{bmatrix},\quad i=1,2,3. \tag{161}$$

It is easy to verify that

$$X^{-1}=\begin{bmatrix} \mathrm{diag}\left(\frac{1}{(\lambda_i-\lambda_{i+3})p_{i,i}}\right) & \mathrm{diag}\left(\frac{-\lambda_{i+3}}{0.5(\lambda_i-\lambda_{i+3})p_{i,i}}\right) \\ \mathrm{diag}\left(\frac{-1}{(\lambda_i-\lambda_{i+3})p_{i+3,i}}\right) & \mathrm{diag}\left(\frac{\lambda_i}{0.5(\lambda_i-\lambda_{i+3})p_{i+3,i}}\right) \end{bmatrix}. \tag{162}$$

Substituting (155), (161), and (162) into (154) gives the robust pole assignment state feedback

$$G=J[I\quad 0]\left(\begin{bmatrix} 0 & 0 \\ 0.5I & 0 \end{bmatrix}-\begin{bmatrix} \mathrm{diag}(\lambda_i p_{i,i}) & \mathrm{diag}(\lambda_{i+3}p_{i+3,i}) \\ \mathrm{diag}(0.5p_{i,i}) & \mathrm{diag}(0.5p_{i+3,i}) \end{bmatrix}\begin{bmatrix} \mathrm{diag}(\lambda_i) & 0 \\ 0 & \mathrm{diag}(\lambda_{i+3}) \end{bmatrix}X^{-1}\right)$$
$$=-J\left[\mathrm{diag}(\lambda_i p_{i,i})\quad \mathrm{diag}(\lambda_{i+3}p_{i+3,i})\right]\begin{bmatrix} \mathrm{diag}(\lambda_i) & 0 \\ 0 & \mathrm{diag}(\lambda_{i+3}) \end{bmatrix}X^{-1}$$
$$=-J\left[\mathrm{diag}(\lambda_i^2 p_{i,i})\quad \mathrm{diag}(\lambda_{i+3}^2 p_{i+3,i})\right]$$
$$\times\begin{bmatrix} \mathrm{diag}\left(\frac{1}{(\lambda_i-\lambda_{i+3})p_{i,i}}\right) & \mathrm{diag}\left(\frac{-\lambda_{i+3}}{0.5(\lambda_i-\lambda_{i+3})p_{i,i}}\right) \\ \mathrm{diag}\left(\frac{-1}{(\lambda_i-\lambda_{i+3})p_{i+3,i}}\right) & \mathrm{diag}\left(\frac{\lambda_i}{0.5(\lambda_i-\lambda_{i+3})p_{i+3,i}}\right) \end{bmatrix}$$
$$=-J\left[\mathrm{diag}\left(\frac{\lambda_i^2-\lambda_{i+3}^2}{\lambda_i-\lambda_{i+3}}\right),\quad \mathrm{diag}\left(\frac{\lambda_{i+3}^2\lambda_i-\lambda_i^2\lambda_{i+3}}{0.5(\lambda_i-\lambda_{i+3})}\right)\right]$$
$$=-J[\mathrm{diag}(\lambda_i+\lambda_{i+3}),\quad \mathrm{diag}(-2\lambda_i\lambda_{i+3})], \tag{163}$$

or

$$G=[\mathrm{diag}(-J_{ii}(\lambda_i+\lambda_{i+3})),\quad \mathrm{diag}(2J_{ii}(\lambda_i\lambda_{i+3}))]. \tag{164}$$

Substituting (144) into (164) yields (137). Therefore, we conclude that the LQR design method is actually a robust pole assignment design for the linearized system (58), and the feedback matrix $G = -[D, K]$ is composed of two diagonal matrices D and K. With the same restriction as discussed before, the robust pole assignment controller globally stabilizes the nonlinear spacecraft system. Finally, we would like to point out that robust pole assignment design will also reduce the impact of the disturbance torques on the system output. The detailed discussion can be found in Yang (2012).

6. Future research

We presented some recent development on quaternion based spacecraft modeling, attitude determination and control. We showed that the reduced quaternion models are attractive because the linearized reduced quaternion models are controllable. The controllability of the linearized models allows us to make full use of the linear control system theory and design methods. The limited results on attitude controller design have demonstrated some attractive features in the designed control system, such as global stability of the nonlinear spacecraft, analytic LQR solutions for optimal performance, robustness to the modeling uncertainty, and disturbance rejection to the external disturbance.

We expect that the reduced quaternion model can be extended to other control system designs in aerospace engineering, such as soft landing of a lunar module presented in Zhang and Duan (2012), spacecraft dock and rendezvous studied in Fejzic (2008), missile attitude control discussed in Song, Kim, and Nam (2006), orbit-raising attitude control proposed in Sidi (1997) and Stoltz et al. (1998), and lauch vehicle attitude control reported in Stott and Shtessel (2012).

To effectively use the design method, the quaternion determination algorithm is needed. QUEST method has been widely known as a robust and efficient algorithm (Cheng & Shuster, 2007). But we believe that the analytic solution sketched in this review may be a more attractive method because it provides not only an accurate solution in theory but also a predictable computational cost without using iteration. However, more detailed and extensive tests will be needed.

We have not addressed the quaternion estimation algorithms such as Kalman filter and extended Kalman filter. Currently, nonlinear and extended Kalman filters are the choices for many onboard spacecraft attitude estimation because spacecraft system is a nonlinear system (Crassidis, Markley, & Cheng, 2007). However, many nonlinear and extended Kalman filters are not proved to be convergent and their computational cost is much higher than linear Kalman filter. Since the linearized reduced quaternion models are fully controllable, linear Kalman filter should be applied to the linearized reduced quaternion models directly. Though we believe that using linear Kalman filter with a linearized reduced quaternion model is a better choice, extensive simulation and flight test may be conducted in the future.

The design method presented in this review does not consider the actuator saturation restrictions which always exist in the real spacecraft systems and need to be considered. For discrete systems, the LQR design with the actuator saturation restrictions is equivalent to a convex quadratic programming problem (Wright, 1993). This problem can be solved very efficiently using interior-point methods. This will be another future research direction given the fact that some very efficient interior-point method, such as arc-search method (Yang, 2011), is recently obtained.

References

Vallado, David A. (2004). *Fundamentals of astrodynamics and applications*. El Segundo, CA: Microcosm Press.

Kuipers, JB (1998). *Quaternions and rotation sequences: A primer with applications to orbits, aerospace, and virtual reality*. Princeton and Oxford: Princeton University Press.

Golub, G. H., & Van Loan, C. F. (1989). *Matrix computations*. Baltimore: The Johns Hopkins University press.

Hughes, Peter C. (1986). *Spacecraft attitude dynamics*. New York, USA: Wiley.

Wie, B. (1998). *Vehicle dynamics and control*. AIAA Education Series, Reston, VA.

Wen, J., & Kreutz-Delgado, K. (1991). The attitude control problem. *IEEE Transactions on Automatic Control, 36*, 1148–1161.

Boskovic, J., Li, S., & Mehra, R. (2001). Robust adaptive variable structure control of spacecraft under control input saturation. *Journal of Guidance, Control and Dynamics, 24*, 14–22.

Wallsgrove, R., & Akella, M. (2005). Globally stabilizing saturated control in the presence of bounded unknown disturbances. *Journal of Guidance, Control and Dynamics, 28*, 957–963.

Paielli, R., & Bach, R. (1993). Control with realization of linear error dynamics. *Journal of Guidance, Control and Dynamics, 16*, 182–189.

Zhou, Z., & Colgren, R. (2005). Nonlinear spacecraft attitude tracking controller for large non-constant rate commands. *International Journal of Control, 78*, 311–325.

Sidi, M. (1997). *Spacecraft dynamics and control: A practical engineering approach*. Cambridge, UK: Cambridge University Press.

Serway, R. A., & Jewett, J. W. (2004). *Physics for scientists and engineers*. Belmont, CA: Books/Cole Thomson Learing.

Wertz, J. (1978). *Spacecraft attitude determination and control*. Dordrecht, Holland: Kluwer Academic Publishers.

Yang, Y. (2010). Quaternion based model for momentum biased nadir pointing spacecraft. *Aerospace Science and Technology, 14*, 199–202.

Yang, Y. (2012). Analytic LQR design for spacecraft control system based on quaternion model. *Journal of Aerospace Engineering, 25*, 448–453.

Black, H. D. (1964). A passive system for determining the attitude of a satellite. *AIAA Journal, 2*, 1350–1351.

Wahba, G. (1965). A least squares estimate of spacecraft attitude. *SIAM Review, 7*, 409.

Shuster, M. D., & Oh, S. D. (1981). Three-axis attitude determination form vector observations. *Journal of Guidance and Control, 4*, 70–77.

Bar-Itzhack, I. Y., & Oshman, Y. (1985). Attitude determination from vector observations: Quaternion estimation. *IEEE Transactions on Aerospace and Electronic Systems, 21*, 128–136.

Shuster, M. D. (1993). A survey of attitude presentation. *Journal of the Astronautical Sciences, 27*, 439–517.

Reynolds, R. G. (1998). Quaternion parameterization and a simple algorithm for global attitude estimation. *Journal of Guidance, Control and Dynamics, 21*, 669–671.

Markley, F. L. (2002). Fast quaternion attitude estimation from two vector measurements. *Journal of Guidance, Control and Dynamics, 25*, 411–414.

Mortari, D. (1997). ESOQ: A closed-form solution to the Wahba problem. *Journal of Astronautical Sciences, 45*(2), 195–205.

Markley, F. L. (1993). Attitude determination using vector observations: A fast optimal matrix algorithm. *Journal of Astronautical Sciences, 41*(2), 261–280.

Davenport (1965). *A vector approach to the algebra of rotations with applications*. Technical Report, NASA, X-546-65-437.

Keat, J. (1977). *Analysis of least-squares attitude determination*. Routine DOAOP, Technical Report, Comp. Sc. Corp., CSC/TM-77/6034.

Rugh, W. J. (1993). *Linear system theory*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., pp. 4–5.

Markley, F. L., & Mortari, D. (2000). Quaternion attitude estimation using vector observations. *Journal of Astronautical Sciences, 48*(2), 359–380.

Nocedal, J., & Wright, S. J. (1999). *Numerical optimization*. New York: Springer-Verlag.

Herbison-Evans, D. (1994). *Solving quartics and cubics for graphics*. TR94-487, University of Sydney, Sydney, Australia.

Shmakov, S. L. (2011). A universal method of solving quartic equations. *International Journal of Pure and Applied Mathematics, 71*(2), 251–259.

Polyanin, Andrei D., & Manzhirov, Alexander V. (2007). *Handbook of mathematics for engineers and scientists*. Noca Raton, FL: Chapman & Hall/CRC.

Carl Christian Liebe, & Sohrab Mobasser. (2001). MEMS based sun sensor. In *IEEE proceedings of aerospace conference* (pp. 1565–1572).

Wikipedia (2011). http://en.wikipedia.org/wiki/Magnetometer.

Jer-Nan Juang, Hye Young Kim, & Junkins, John L. (2003). *An efficient and robust singular value method for star pattern recognition and attitude determination*. Technical Report, NASA/TM-2003-212142, Langley Research Center, NASA, Hampton, Virginia.

Fortescue, P., Stark, J., & Swinerd, G. (2003). *Spacecraft systems engineering*. West Sussex, England: John Wiley & Sons.

Singla, P., L Crassida, J., & Junkins, J. L. (2003). Spacecraft angular rate estimation algorithms for star tracker-based attitude determination. *Advances in the Astronautical Sciences, 114*, 1303–1316.

Oppenheim, A. V., Schafer, R. W., & Buck, J. R. (1999). *Discrete-time signal processing*. New York: Prentice Hall.

Wie, B., Weiss, H., & Arapostathis, A. (1989). Quaternion feedback regulator for spacecraft eigenaxis rotations. *Journal of Guidance, Control and Dynamics, 12*, 375–380.

Stoltz, P. M., Sivapiragasam, S., & Anthony, T. (1998). Satellite orbit-raising using LQR control with fixed thrusters. *Advances in the Astronautical Sciences, 98*, 109–120.

Won, Chang-Hee (1999). Comparative study of various control methods for attitude control of a LEO satellite. *Aerospace Science and Technology, 3*, 323–333.

Athans, M., & Falb, P. L. (1966). *Optimal control: An introduction to the theory and its applications*. New York: McGraw-Hill, Inc.

Kautsky, J., Nichols, N. K., & Van Dooren, P. (1985). Robust pole assignment in linear state feedback. *International Journal of Control, 41*, 1129–1155.
Tits, André L., & Yang, Y. (1996). Globally convergent algorithms for robust pole assignment by state feedback. *IEEE Transactions on Automatic Control, 41*, 1432–1452.
Yang, Y. (1996). *Robust system design: Pole assignment approach*. College Park, MD: University of Maryland at College Park.
Doyle, J., Glover, K., Khargonekar, P., & Francis, B. (1989). State-space solutions to standard H_2 and H_∞ control problems. *IEEE Transactions on Automatic Control, 34*, 831–847.
Yang, Y. (2012). Quaternion based LQR spacecraft control design is a robust pole assignment design. *Journal of Aerospace Engineering*. http://dx.doi.org/10.1061/(ASCE)AS.1943-5525.0000232.
HALL, A. C. (1943). *The analysis and synthesis of linear servomechanisms*. Cambridge, MA: The Technology Press.
WIENER, N. (1949). *Extrapolation, interpolation, and smoothing of stationary time series*. Cambridge, MA: MIT Press.
Kalman, R. E. (1960). Contributions to the theory of optimal control. *Boletin de la Sociedad Matematica Mexicana, 5*, 102–119.
Kalman, R. E. (1960). A new approach to linear filtering and prediction problem. *Transactions of the ASME – Journal of Basic Engineering, 82D*, 35–45.
Wonham, W. M. (1968). On the separation theorem of stochastic control. *SIAM Journal on Control, 6*, 312–326.
Khalil, Hassan K. (1992). *Nonlinear system*. New York: Macmillan Publishing Company.
Wilkinson, J. H. (1965). *The algebraic eigenvalue problem*. Oxford: Charendon Press.
Yang, Y. (1989). A new condition number of eigenvalue and its application in control theory. *Journal of Computational Mathematics, 7*, 15–21.
Sima, V., Tits, A. L., & Yang, Y. (2006). *Computational experience with robust pole assignment algorithms*. In *2006 IEEE conference on computer aided control systems design*, Munich, Germany.
Zhang, F., Duan, G. R. (2012). Integrated translational and rotational control for the terminal landing phase of a lunar modele. *Aerospace Sciences and Technologies*. <http://dx.doi.org/10.1016/j.ast.2012.07.003>.
Fejzic, A. (2008). *Development of control and autonomy algorithm for docking to complex tumbling satellites*. Master Thesis, MIT.
Song, C., Kim, S., & Nam, H. (2006). Robust control of the missile attitude based on quaternion feedback. *Control Engineering Practice, 14*, 811–818.
Stott, J. E., & Shtessel (2012). Launch vehicle attitude control using sliding mode control and observation techniques. *Journal of the Franklin Institute, 349*, 397–412.
Cheng, J., & Shuster, M. D. (2007). The speed of attitude estimation. *Advances in the Astronautical Sciences, 127*(Part 1), 101–116.
Crassidis, J. L., Markley, L., & Cheng, J. (2007). Survey of nonlinear attitude estimation methods. *Journal of Guidance, Control and Dynamics, 30*, 12–28.
Wright, S. J. (1993). Interior point methods for optimal control of discrete time systems. *Journal of Optimization Theory and Applications, 77*, 161–187.
Yang, Y. (2011). A polynomial arc-search interior-point algorithm for convex quadratic programming. *European Journal of Operational Research, 215*, 25–38.

Yaguang Yang received the B.S. (1982) and M.S. (1985) degrees from the Department of Automatic Control Engineering at Huazhong University of Science and Technology, China. From 1985 to 1990, he was a lecturer at Zhejiang University in China. In 1996, he received the Ph.D degree from the Department of Electrical and Computer Engineering at the University of Maryland, College Park. Since then, he has worked in industry and the US government. He designed and implemented control systems for astronomical telescope system at UKIRT, for fiber optical equipment manufacture system at CIENA, for free space optical communication system at ITT, for spacecraft and rocket systems at Orbital Sciences Corporation. Currently he is a research project manager at the US Nuclear Regulatory Commission. He managed Advanced Reactor Controls and Advanced Reactor Instrumentation projects for the Next Generation Nuclear Power Plant (NGNP) with hydrogen production. He has also conducted research and published journal papers in linear programming, nonlinear programming, and optimization on Riemannian Manifolds.