JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 20, No. 3, May–June 1997

New Penalty Functions and Optimal Control Formulation for Spacecraft Attitude Control Problems

Hanspeter Schaub* and John L. Junkins†
Texas A&M University, College Station, Texas 77843
and
Rush D. Robinett‡
Sandia National Laboratories, Albuquerque, New Mexico 87185

A universal attitude penalty function $g()$ is presented that renders spacecraft optimal control problem solutions independent of attitude coordinate choices. This function returns the same scalar penalty for a given physical attitude regardless of the choice of attitude coordinates used to describe this attitude. The only singularities the $g()$ function might encounter are solely due to the choice of attitude coordinates. A second attitude penalty function $G()$ is considered, which depends specifically on the modified Rodrigues parameter (MRP) vector σ. The MRPs allow for a nonsingular attitude description through a noncontinuous switch to their associated shadow set. A corresponding switching condition is presented for the MRP costates to allow the nonsingular MRP attitude description to be used in optimal control problems. Computational experiments are used to illustrate and validate the analytical results.

I. Introduction

SOLUTIONS of spacecraft optimal control problems, whose cost function rely on an attitude description, usually depend on the choice of attitude coordinates used. Coordinate choices are often considered a matter of taste, but the question of coordinate optimality arises. For example, a problem could be solved using 3–2–1 Euler angles or using classical Rodrigues parameters and yield two different optimal solutions, unless the performance index is invariant with respect to the attitude coordinate choice. Another problem arising with many attitude coordinates is that the resulting control formulation has no intrinsic sense of when a body has tumbled beyond ± 180 deg from the reference attitude. In many such cases it would be simpler and cheaper to let the body complete the revolution rather than force it to reverse the rotation and return to the desired attitude.

This paper develops a universal attitude penalty function $g()$ whose value is independent of the attitude coordinates chosen to represent it. Furthermore, this function achieves its maximum value for any principal rotation of ± 180 deg from the target state. This implicitly permits the $g()$ function to sense the shortest rotational distance back to the reference state.

An attitude penalty function $G()$, which depends specifically on the modified Rodrigues parameters (MRP) will also be presented. This MRP penalty function is simpler than the attitude coordinate independent $g()$ function, but retains the useful property of avoiding lengthy principal rotations of more than ± 180 deg and being nonsingular. These recently discovered MRPs[1–6] allow for a nonsingular, three-parameter attitude description. This is achieved by switching the MRPs to their associated shadow set, which abide by exactly the same differential equation.[1,4] Because the MRPs are discontinuous during this switching, a corresponding MRP costate switching condition is introduced that allows the nonsingular MRP attitude description to be used in optimal control problems.

Received Aug. 26, 1996; revision received Feb. 4, 1997; accepted for publication Feb. 5, 1997.

*Graduate Research Assistant, Aerospace Engineering Department. Student Member AIAA.

†George Eppright Chair Professor, Aerospace Engineering Department. Fellow AIAA.

‡Research Engineer. Senior Member AIAA.

II. Problem Statement

A. Optimal Control Problem

Most spacecraft optimal control problems have a cost function J, which depends on the control effort, the body angular velocity, and the attitude. Let $\boldsymbol{u}$ be the control torque vector, $\boldsymbol{\omega}$ the body angular velocity vector, and $\boldsymbol{\eta}$ a generic attitude coordinate vector in the following general optimal control formulation with fixed maneuver time t_f:

$$\min\ J = h(t_f) + \int_0^{t_f} p(\boldsymbol{\eta}, \boldsymbol{\omega}, \boldsymbol{u}, t)\,\mathrm{d}t \tag{1}$$

subject to

$$(\dot{\boldsymbol{\eta}}, \dot{\boldsymbol{\omega}})^T = F(\boldsymbol{\eta}, \boldsymbol{\omega}, \boldsymbol{u})$$

where typical penalty functions are

$$h(t_f) = \tfrac{1}{2}K_1 g\left(\boldsymbol{\eta}_{t_f}\right) + \tfrac{1}{2}\boldsymbol{\omega}_{t_f}^T K_2 \boldsymbol{\omega}_{t_f} \tag{2}$$

and

$$p(\boldsymbol{\eta}, \boldsymbol{\omega}, \boldsymbol{u}, t) = \tfrac{1}{2}\left[K_3 g(\boldsymbol{\eta}) + \boldsymbol{\omega}^T K_4 \boldsymbol{\omega} + \boldsymbol{u}^T R \boldsymbol{u}\right] \tag{3}$$

The weights K_1 and K_3 are scalars; the weights K_2, K_4, and R are 3×3 matrices. The function $g(\boldsymbol{\eta})$ is a general, nonnegative attitude penalty function. For spacecraft optimal control problems, the equations of motion are usually imposed as an equality constraint. They are given in Eqs. (4) and (5), where the function $f(\boldsymbol{\eta})$, obtained from kinematic analysis, returns a matrix dependent on the choice of attitude coordinates. The equations of motion are

$$\dot{\boldsymbol{\eta}} = f(\boldsymbol{\eta})\boldsymbol{\omega} \tag{4}$$

$$\Im\dot{\boldsymbol{\omega}} = -[\tilde{\omega}]\Im\boldsymbol{\omega} + \boldsymbol{u} \tag{5}$$

The matrix $\Im$ is the spacecraft inertia matrix. The tilde matrix is the cross-product operator

$$[\tilde{\omega}] = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} \tag{6}$$

The Hamiltonian H for this system is

$$H = \tfrac{1}{2}K_3 g(\boldsymbol{\eta}) + \tfrac{1}{2}\boldsymbol{\omega}^T K_4 \boldsymbol{\omega} + \tfrac{1}{2}\boldsymbol{u}^T R \boldsymbol{u} + \boldsymbol{\Lambda}_\sigma^T f(\boldsymbol{\eta})\boldsymbol{\omega} + \boldsymbol{\Lambda}_\omega^T \Im^{-1}(-[\tilde{\omega}]\Im\boldsymbol{\omega} + \boldsymbol{u}) \tag{7}$$

The costate equations are given by[7–9]

$$\dot{\Lambda}_\sigma = -\frac{\partial H}{\partial \boldsymbol{\eta}} = -\frac{1}{2}K_3\frac{\partial g}{\partial \boldsymbol{\eta}} - \frac{\partial}{\partial \boldsymbol{\eta}}[f(\boldsymbol{\eta})\boldsymbol{\omega}]^T\Lambda_\sigma \tag{8}$$

$$\dot{\Lambda}_\omega = -\frac{\partial H}{\partial \boldsymbol{\omega}} = -K_4\boldsymbol{\omega} - f(\boldsymbol{\eta})^T\Lambda_\sigma - (\Im[\tilde{\omega}] - [\widetilde{\Im\omega}])\Im^{-1}\Lambda_\omega \tag{9}$$

For unbounded control torque $\boldsymbol{u}$, the optimality condition $\partial H/\partial \boldsymbol{u} = 0$ leads to the following optimal control torque[7–9]:

$$\boldsymbol{u} = -R^{-1}\Im^{-1}\Lambda_\omega \tag{10}$$

The transversality conditions for a free final state are[7,8]

$$\Lambda_\sigma(t_f) = \frac{\partial h}{\partial \boldsymbol{\eta}}(t_f) = \frac{1}{2}K_1\frac{\partial g}{\partial \boldsymbol{\eta}}(t_f) \tag{11}$$

$$\Lambda_\omega(t_f) = \frac{\partial h}{\partial \boldsymbol{\omega}}(t_f) = K_2\boldsymbol{\omega}(t_f) \tag{12}$$

Given good estimates of initial conditions, this nonlinear optimal control problem can be solved using various standard techniques.[9]

B. Attitude Coordinates

Attitude coordinates define the rotational orientation of a rigid body relative to some reference frame. Just as there are a number of different coordinates that describe translation (Cartesian, cylindrical, spherical), there are an infinite number of different ways to describe an attitude. Common examples are the Euler angles, the classical Rodrigues parameters or the Euler parameters (quaternions). Orientations differ from spatial positions in a fundamental way. The largest difference between two physical orientations corresponds to a principal rotation of ±180 deg, a finite value, whereas the difference in two spatial positions can grow to infinity.

Minimal three-coordinate attitude representations generally contain singularities. These occur at specific attitudes at which the coordinates are not uniquely defined. The Euler parameters avoid any singularities at the cost of adding another coordinate. This redundant set has an equality constraint, which restrains the attitude vector to be of unit magnitude. Therefore, if Euler parameters are used in a simple optimal control problem otherwise having no constraints, an equality constraint is automatically added.[10,11]

Among other coordinate sets, this paper will use the very elegant set of recently developed MRPs with their shadow counterpart.[1,3–5] They allow for a nonsingular rigid-body attitude description with several other useful attributes. The MRPs can be defined through a transformation from the Euler parameters as[1,2,6]

$$\sigma_i = \beta_i/(1+\beta_0) \qquad i = 1,2,3 \tag{13}$$

or in terms of the principal rotation axis $\hat{\boldsymbol{e}}$ and the principal rotating angle χ, the MRP vector is

$$\boldsymbol{\sigma} = \hat{\boldsymbol{e}}\cdot\tan(\chi/4) \tag{14}$$

Like the Euler parameters, the MRPs are not unique. A second set of MRPs, called the shadow set, can be used to avoid the singularity of the original MRP at $\chi = \pm 360$ deg at the cost of a discontinuity at a switching point. The shadow set is found by reversing the sign of the β_i in Eq. (13). The transformation between the original and shadow sets of MRPs for any arbitrary switching surface $\boldsymbol{\sigma}^T\boldsymbol{\sigma}$ is[1,3,4]

$$\sigma_i^S = -\sigma_i/\boldsymbol{\sigma}^T\boldsymbol{\sigma} \qquad i = 1,2,3 \tag{15}$$

Keep in mind that distinguishing between original and shadow set is purely arbitrary. Both sets describe the same physical orientation. If the switching condition is set to $\boldsymbol{\sigma}^T\boldsymbol{\sigma} = 1$, the magnitude of the MRP orientation vector is bounded between $0 \le |\boldsymbol{\sigma}| \le 1$ and the principal rotation angle is restricted between $-180 \le \chi \le +180$ deg. Note that this combined set of original and shadow parameters implicitly knows the shortest rotation back to the origin.[1] Principal rotations of more than 180 deg are typically avoided.

The differential kinematic equation for the MRPs is given next.[1,2,6] This exact kinematic equation only contains second-order polynomial nonlinearities in $\boldsymbol{\sigma}$:

$$\frac{\mathrm{d}\boldsymbol{\sigma}}{\mathrm{d}t} = \frac{1}{2}\left[I\left(\frac{1-\boldsymbol{\sigma}^T\boldsymbol{\sigma}}{2}\right) + [\tilde{\sigma}] + \boldsymbol{\sigma}\boldsymbol{\sigma}^T\right]\boldsymbol{\omega} \tag{16}$$

Equation (16) holds for both the original and the shadow set. This means that the derivative is well defined even at the switching point. Let us introduce the notation $\sigma^{2n} = (\boldsymbol{\sigma}^T\boldsymbol{\sigma})^n$ and note that

$$\frac{\mathrm{d}\sigma^2}{\mathrm{d}t} = \frac{1+\sigma^2}{2}\boldsymbol{\sigma}^T\boldsymbol{\omega} \tag{17}$$

Using Eqs. (16) and (17), the general relationship between $\mathrm{d}\boldsymbol{\sigma}/\mathrm{d}t$ and $\mathrm{d}\boldsymbol{\sigma}^S/\mathrm{d}t$ for an arbitrary switching condition is

$$\frac{\mathrm{d}\boldsymbol{\sigma}^S}{\mathrm{d}t} = -\frac{1}{\sigma^2}\frac{\mathrm{d}\boldsymbol{\sigma}}{\mathrm{d}t} + \frac{1+\sigma^2}{2\sigma^4}\boldsymbol{\sigma}\boldsymbol{\sigma}^T\boldsymbol{\omega} \tag{18}$$

The partial derivative of Eq. (16) with respect to $\boldsymbol{\sigma}$ is

$$\frac{\partial}{\partial \boldsymbol{\sigma}}[f(\boldsymbol{\sigma})\boldsymbol{\omega}] = \frac{1}{2}(\boldsymbol{\sigma}\boldsymbol{\omega}^T - [\tilde{\omega}] - \boldsymbol{\omega}\boldsymbol{\sigma}^T + \boldsymbol{\sigma}^T\boldsymbol{\omega}I) \tag{19}$$

The direction cosine matrix in terms of the MRPs is[1,2,6]

$$C(\boldsymbol{\sigma}) = \frac{1}{(1+\sigma^2)^2} \times \begin{bmatrix} 4\left(2\sigma_1^2 - \sigma^2\right) + \Sigma^2 & 8\sigma_1\sigma_2 + 4\sigma_3\Sigma & 8\sigma_1\sigma_3 - 4\sigma_2\Sigma \\ 8\sigma_1\sigma_2 - 4\sigma_3\Sigma & 4\left(2\sigma_2^2 - \sigma^2\right) + \Sigma^2 & 8\sigma_2\sigma_3 + 4\sigma_1\Sigma \\ 8\sigma_1\sigma_3 + 4\sigma_2\Sigma & 8\sigma_2\sigma_3 - 4\sigma_1\Sigma & 4\left(2\sigma_3^2 - \sigma_2\right) + \Sigma^2 \end{bmatrix}$$

$$\Sigma = 1 - \sigma^2 \tag{20}$$

III. Universal Attitude Penalty Function

A scalar attitude penalty function is sought that will return the same value for a given physical orientation, regardless of the choice of attitude coordinates used to describe this function. This allows for a universal solution to many spacecraft optimal control problems and removes the dependency on the attitude coordinate choice. We introduce the following nonnegative measure of attitude displacement from a reference orientation:

$$g([C]) = \tfrac{1}{4}\{3 - \mathrm{tr}([C])\} \qquad \in \Re^+ \tag{21}$$

This penalty function is given in terms of a proper orthogonal direction cosine matrix $[C]$. This rotation matrix is the most fundamental way to describe a rotation; unfortunately, it is also the most redundant. If there is no rotational displacement, the $[C]$ matrix is the identity matrix and $g([C]) = 0$.

The largest difference between two attitudes is a principal rotation of ±180 deg. Here the $[C]$ matrix is a diagonal matrix with two entries being −1 and one being +1. In this case $g([C]) = 1$. Therefore, the $g(\,)$ function is bounded for all possible motion between

$$0 \le g(\,) \le 1 \tag{22}$$

This penalty function can be written explicitly in terms of the principal rotation angle χ as

$$g\{[C(\hat{\boldsymbol{e}},\chi)]\} = \sin^2(\chi/2) \tag{23}$$

This description makes the bound in Eq. (22) very easy to understand. The attitude cost is the highest only if the body is turned ±180 deg from the reference state. The $g(\,)$ function is plotted relative to the principal rotation angle χ in Fig. 1. Using this type of attitude penalty function will typically avoid lengthy rotations. It intrinsically lowers the cost once the attitude has moved beyond ±180 deg. The advantage of defining the $g(\,)$ function initially in terms of the $[C]$ matrix is that this rotation matrix can be parameterized by any attitude coordinates and, thus, making it universally valid for any choice of attitude coordinates. The following are a

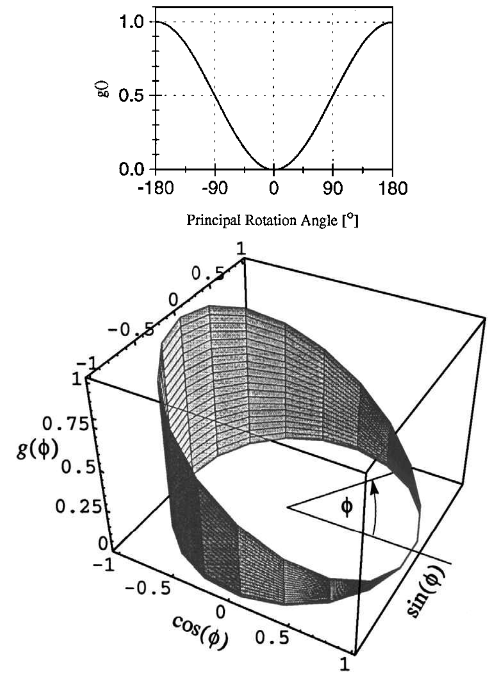

Fig. 1 Universal attitude penalty function $g()$.

few sample parameterizations. Let $\boldsymbol{\beta}$ be an Euler parameter vector; then

$$g\{[C(\boldsymbol{\beta})]\} = \beta_1^2 + \beta_2^2 + \beta_3^2 \tag{24}$$

and

$$\frac{\partial g}{\partial \boldsymbol{\beta}} = [0 \quad 2\beta_1 \quad 2\beta_2 \quad 2\beta_3]^T \tag{25}$$

The Euler parameters are a nonsingular once-redundant set of attitude coordinates. Their drawback for optimization problems is that the redundancy introduces an additional equality constraint on the attitude vector. This redundancy also requires care to avoid other numerical problems when inverting some sets of equations.[10]

Another popular attitude coordinate set is the classical Rodrigues parameter vector $\boldsymbol{q}$. It parameterizes the $g()$ function as

$$g\{[C(\boldsymbol{q})]\} = \frac{\boldsymbol{q}^T\boldsymbol{q}}{1 + \boldsymbol{q}^T\boldsymbol{q}} \tag{26}$$

and

$$\frac{\partial g}{\partial \boldsymbol{q}} = 2\boldsymbol{q}\frac{1}{(1 + \boldsymbol{q}^T\boldsymbol{q})^2} \tag{27}$$

These coordinates are a minimal three-coordinate set and do not have any problems with redundancies. However, like most three-parameters sets, they contain a singular orientation. The classical Rodrigues parameters go singular for any principal rotation of ± 180 deg. If it is a priori known that no such rotations will be encountered, this choice for attitude coordinates does allow for a large range of possible rotations.

Probably the most popular choice of attitude coordinates are any one of the 12 sets of Euler angles. Let $(\theta_1, \theta_2, \theta_3)$ be the set of 3–1–3 Euler angles. They parameterize the $g()$ function as

$$g\{[C(\boldsymbol{\theta})]\} = \tfrac{1}{4}[3 - (1 + \cos\theta_2)\cos(\theta_1 + \theta_3) - \cos\theta_2] \tag{28}$$

and

$$\frac{\partial g}{\partial \boldsymbol{\theta}} = \frac{1}{4}\begin{bmatrix} (1 + \cos\theta_2)\sin(\theta_1 + \theta_3) \\ \sin\theta_2\cos(\theta_1 + \theta_3) + \sin\theta_2 \\ (1 + \cos\theta_2)\sin(\theta_1 + \theta_3) \end{bmatrix} \tag{29}$$

The advantage of the Euler angles is that they are easy to visualize, especially for small angles. However, any attitude description with Euler angles is never more than 90 deg away from a singularity. This makes these coordinates difficult to use for large arbitrary rotations. Further, the kinematic equations for the Euler angles are in terms of trigonometric functions, making them more computationally intensive than having only polynomial equations.

A very attractive attitude description are the MRPs. They parameterize the $g()$ function as

$$g\{[C(\boldsymbol{\sigma})]\} = 4\frac{\boldsymbol{\sigma}^T\boldsymbol{\sigma}}{(1 + \boldsymbol{\sigma}^T\boldsymbol{\sigma})^2} \tag{30}$$

and

$$\frac{\partial g}{\partial \boldsymbol{\sigma}} = 8\boldsymbol{\sigma}\left[\frac{1 - \boldsymbol{\sigma}^T\boldsymbol{\sigma}}{(1 + \boldsymbol{\sigma}^T\boldsymbol{\sigma})^3}\right] \tag{31}$$

The MRPs are a minimal attitude description, which are also nonsingular when combined with their corresponding shadow set. They are well suited to describe any large arbitrary rotation while their equations retain a simple polynomial form.

The $g([C])$ attitude penalty function can be parameterized by any other attitude coordinate representation of $[C]$. Note that all $g()$ equations shown return the same penalty for a given physical orientation. They only differ in their $[C()]$ representation. This attitude coordinate independent attitude penalty function effectively removes the dependency of the optimal control solution on the choice of attitude coordinates. However, the optimal attitude costate vector $\boldsymbol{\Lambda}_\sigma$ depends on the attitude coordinates used inasmuch as Eqs. (8) and (11) depend on the partial derivative of $g()$ with respect to the particular attitude coordinates.

IV. MRP Attitude Penalty Function

Although the universal attitude penalty function $g()$ has some very appealing properties, it is usually more complicated than just using the standard sum squared of the attitude coordinates typically seen as an optimal control performance measure. For example, using the simpler attitude penalty function

$$G(\boldsymbol{\sigma}) = \boldsymbol{\sigma}^T\boldsymbol{\sigma} \tag{32}$$

where $\boldsymbol{\sigma}$ is a MRP vector, retains all properties of $g()$ defined in Eq. (21), except being universal with respect to attitude coordinate choice. By switching between original and shadow MRP trajectories on the $\boldsymbol{\sigma}^T\boldsymbol{\sigma} = 1$ surface, the attitude penalties in Eq. (32) are bounded within [0, 1]. Using Eq. (14), the penalty function can be written in terms of the principal rotation angle χ as

$$G[\boldsymbol{\sigma}(\hat{\boldsymbol{e}}, \chi)] = \tan^2(\chi/4) \tag{33}$$

The $G()$ function is plotted relative to the principal rotation angle χ in Fig. 2. Note that like the universal attitude penalty function $g()$, the maximum attitude penalty is also attained at a principal rotation of ± 180 deg.

By using the MRPs, this penalty function is globally nonsingular using a minimal attitude coordinate description. The nonsingularity comes at the price of having the switching condition defined in Eq. (15). However, both $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}^S$ are well defined for both the original and the shadow parameters. By choosing the switching surface $\boldsymbol{\sigma}^T\boldsymbol{\sigma} = 1$, Eqs. (15) and (18) are simplified to

$$\boldsymbol{\sigma}^S = -\boldsymbol{\sigma} \tag{34}$$

$$\dot{\boldsymbol{\sigma}}^S = \dot{\boldsymbol{\sigma}} - [\tilde{\sigma}]\boldsymbol{\omega} \tag{35}$$

Observe from Eq. (35) that for pure single-axis rotations $\dot{\boldsymbol{\sigma}}$ simply equals $\dot{\boldsymbol{\sigma}}^S$ on the $\boldsymbol{\sigma}^T\boldsymbol{\sigma} = 1$ switching surface. Because the derivatives of the costates depend on the attitude coordinates, they will also have a discontinuity as the attitude vector is switched.

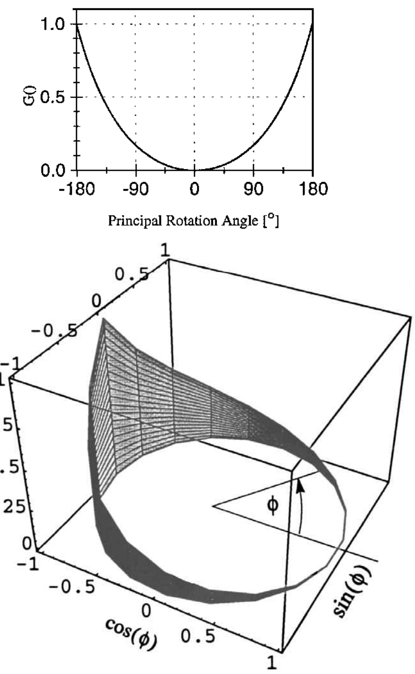

Fig. 2 MRP attitude penalty function G( ).

V. MRP Costate Switching Condition

The MRPs have many useful attributes. However, to avoid a singularity, this minimal attitude coordinate description needs to switch between the original and the shadow MRP set in a discontinuous fashion.[1,4] This switching can occur on any surface $\boldsymbol{\sigma}^T\boldsymbol{\sigma} = c^2$, where $c \geq 0$. The most attractive switching surface is $\boldsymbol{\sigma}^T\boldsymbol{\sigma} = 1$ since it bounds $|\boldsymbol{\sigma}|$ to be within unit magnitude. The optimality conditions for the optimal control problem used were derived assuming that all states were piecewise smooth and continuous. The continuity is no longer guaranteed when using both sets of MRPs.

Because the original and shadow MRP satisfy exactly the same differential equation, the MRP costate differential equation in Eq. (8) is also the same for either set of MRPs. To be able to switch the MRPs along the optimal trajectory, an analogous mapping of the MRP costates at the time of the switch is required. Because the attitude penalty function $g(\,)$ renders the optimal trajectory independent of attitude coordinate choice, the MRP switching could be triggered by any $\boldsymbol{\sigma}^T\boldsymbol{\sigma} = c^2$ surface. If the $G(\,)$ function is used, then one must switch the MRPs on the $\boldsymbol{\sigma}^T\boldsymbol{\sigma} = 1$ surface; otherwise the cost function J will be discontinuous.

Deriving the necessary corner conditions for discontinuous states has been covered in the literature, such as in Ref. 9. However, the state discontinuities, and, therefore, the variations of the state before and after a jump, are assumed to be arbitrary in this reference. With the MRPs we are in the unique situation where the discontinuity is well defined through Eq. (15). This results in the state variations before and after a jump being related and leads to a different costate switching condition than what is found in the literature.

The Weierstrass–Erdmann corner conditions were developed for the case where the state derivative is discontinuous.[7,12] The same initial assumptions used in deriving the Weierstrass–Erdmann corner conditions also hold if the state, not the derivative of the state, is discontinuous. Without loss of generality, let us assume that $\boldsymbol{\sigma}$ is only discontinuous at t_1, where $0 < t_1 < t_f$. The cost function J, in terms of the system Hamiltonian H and the costates $\boldsymbol{\Lambda}$, can now be written as

$$J = h(t_f) + \int_0^{t_1^-} (H - \boldsymbol{\Lambda}^T\dot{\boldsymbol{x}})\,\mathrm{d}t + \int_{t_1^+}^{t_f} (H - \boldsymbol{\Lambda}^T\dot{\boldsymbol{x}})\,\mathrm{d}t$$

$$= J_1 + J_2 + J_3 \tag{36}$$

where $\boldsymbol{x} = \mathrm{col}(\boldsymbol{\sigma}, \boldsymbol{\omega})$ and $\boldsymbol{\Lambda} = \mathrm{col}(\boldsymbol{\Lambda}_\sigma, \boldsymbol{\Lambda}_\omega)$. For notational compactness, let us define $\boldsymbol{\sigma}_- = \boldsymbol{\sigma}(t_1^-)$, $\boldsymbol{\sigma}_+ = \boldsymbol{\sigma}(t_1^+)$, $\boldsymbol{\Lambda}_{i^-} = \boldsymbol{\Lambda}_i(t_1^-)$, and $\boldsymbol{\Lambda}_{i^+} = \boldsymbol{\Lambda}_i(t_1^+)$. Each integral can now be evaluated without state discontinuity problems. The first variation of J must satisfy[7]

$$\delta J = 0 = \delta J_1 + \delta J_2 + \delta J_3 \tag{37}$$

The first variation of J_1 is

$$\delta J_1 = \frac{\partial h}{\partial \boldsymbol{x}}[\boldsymbol{x}(t_f)]^T\,\delta\boldsymbol{x}(t_f) \tag{38}$$

Taking the first variation of J_2, we take into account that t_1 is a free final time:

$$\delta J_2 = \left[H\left(t_1^-\right) - \boldsymbol{\Lambda}\left(t_1^-\right)^T\dot{\boldsymbol{x}}\left(t_1^-\right)\right]\delta t_1 + \int_0^{t_1^-}\left(\frac{\partial H}{\partial \boldsymbol{x}}^T\delta\boldsymbol{x} - \boldsymbol{\Lambda}^T\delta\dot{\boldsymbol{x}} + \frac{\partial H}{\partial \boldsymbol{u}}^T\delta\boldsymbol{u}\right)\mathrm{d}t \tag{39}$$

Because the states are smooth and continuous within the integral, $\delta\dot{\boldsymbol{x}}$ can be written as $\mathrm{d}/\mathrm{d}t(\delta\boldsymbol{x})$. This permits δJ_2 to be integrated by parts:

$$\delta J_2 = \left[H\left(t_1^-\right) - \boldsymbol{\Lambda}\left(t_1^-\right)^T\dot{\boldsymbol{x}}\left(t_1^-\right)\right]\delta t_1 - \boldsymbol{\Lambda}\left(t_1^-\right)^T\delta\boldsymbol{x}\left(t_1^-\right) + \int_0^{t_1^-}\left[\left(\frac{\partial H}{\partial \boldsymbol{x}} + \dot{\boldsymbol{\Lambda}}\right)^T\delta\boldsymbol{x} + \frac{\partial H}{\partial \boldsymbol{u}}^T\delta\boldsymbol{u}\right]\mathrm{d}t \tag{40}$$

Let $\delta\boldsymbol{x}(t_1^-) = [\delta\boldsymbol{\sigma}(t_1^-), \delta\boldsymbol{\omega}(t_1^-)]^T$. Then the state variations at $t_1^- + \delta t_1$ are defined as[7,8]

$$\delta\boldsymbol{\sigma}_- = \delta\boldsymbol{\sigma}\left(t_1^- + \delta t_1\right) = \delta\boldsymbol{\sigma}\left(t_1^-\right) + \dot{\boldsymbol{\sigma}}\left(t_1^-\right)\delta t_1 \tag{41}$$

$$\delta\boldsymbol{\omega}_- = \delta\boldsymbol{\omega}\left(t_1^- + \delta t_1\right) = \delta\boldsymbol{\omega}\left(t_1^-\right) + \dot{\boldsymbol{\omega}}\left(t_1^-\right)\delta t_1 \tag{42}$$

which reduces δJ_2 to the simple form

$$\delta J_2 = H\left(t_1^-\right)\delta t_1 - \boldsymbol{\Lambda}_{\sigma^-}^T\delta\boldsymbol{\sigma}_- - \boldsymbol{\Lambda}_{\omega^-}^T\delta\boldsymbol{\omega}_- + \int_0^{t_1^-}\left[\left(\frac{\partial H}{\partial \boldsymbol{x}} + \dot{\boldsymbol{\Lambda}}\right)^T\delta\boldsymbol{x} + \frac{\partial H}{\partial \boldsymbol{u}}^T\delta\boldsymbol{u}\right]\mathrm{d}t \tag{43}$$

Similarly, δJ_3 can be found assuming that the initial and final states and the time t_1 are free:

$$\delta J_3 = \boldsymbol{\Lambda}_{\sigma^+}^T\delta\boldsymbol{\sigma}_+ + \boldsymbol{\Lambda}_{\omega^+}^T\delta\boldsymbol{\omega}_+ - H\left(t_1^+\right)\delta t_1 - \boldsymbol{\Lambda}(t_f)^T\delta\boldsymbol{x}_f + \int_{t_1^+}^{t_f}\left[\left(\frac{\partial H}{\partial \boldsymbol{x}} + \dot{\boldsymbol{\Lambda}}\right)^T\delta\boldsymbol{x} + \frac{\partial H}{\partial \boldsymbol{u}}^T\delta\boldsymbol{u}\right]\mathrm{d}t \tag{44}$$

Because the body angular velocity is continuous, $\delta\boldsymbol{\omega}_- = \delta\boldsymbol{\omega}_+ = \delta\boldsymbol{\omega}$. After enforcing the optimality and transversality conditions, the total variation δJ becomes[9]

$$\delta J = \boldsymbol{\Lambda}_{\sigma^+}^T\delta\boldsymbol{\sigma}_+ - \boldsymbol{\Lambda}_{\sigma^-}^T\delta\boldsymbol{\sigma}_- + (\boldsymbol{\Lambda}_{\omega^+} - \boldsymbol{\Lambda}_{\omega^-})^T\delta\boldsymbol{\omega} - \left[H\left(t_1^+\right) - H\left(t_1^-\right)\right]\delta t_1 = 0 \tag{45}$$

Because the variations $\delta\boldsymbol{\omega}$ and δt_1 in Eq. (45) are independent from other variations, the following conclusions can be made:

$$\boldsymbol{\Lambda}_{\omega^+} = \boldsymbol{\Lambda}_{\omega^-} \tag{46}$$

$$H\left(t_1^+\right) = H\left(t_1^-\right) \tag{47}$$

Before any conclusions can be made about $\boldsymbol{\Lambda}_{\sigma^-}$ and $\boldsymbol{\Lambda}_{\sigma^+}$, further development is needed to establish what constitutes an admissible variation $\delta\boldsymbol{\sigma}_-$ and $\delta\boldsymbol{\sigma}_+$, what is their relationship, and whether they are nonzero.

It is assumed that at time t_1^- the optimal attitude $\boldsymbol{\sigma}_-^*$ is on the constraint surface $\boldsymbol{\sigma}^T\boldsymbol{\sigma} = c^2$. Let $\boldsymbol{\sigma}(t_1^- + \delta t_1)$ be a variation of $\boldsymbol{\sigma}_-^*$. Because the variation of $\boldsymbol{\sigma}_-^*$ must also be on the sphere surface, the following condition must hold:

$$\left|\boldsymbol{\sigma}_-^*\right|^2 = \left|\boldsymbol{\sigma}\left(t_1^- + \delta t_1\right)\right|^2 = c^2 \tag{48}$$

Let $\boldsymbol{\sigma}(t_1 + \delta t_1)$ be related to the optimal $\boldsymbol{\sigma}_-^*$ through

$$\boldsymbol{\sigma}(t_1 + \delta t_1) = \boldsymbol{\sigma}^*(t_1) + \delta\boldsymbol{\sigma}_1 \tag{49}$$

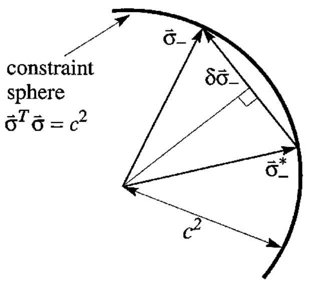

Fig. 3 Constraint illustration of $\delta\boldsymbol{\sigma}_-$.

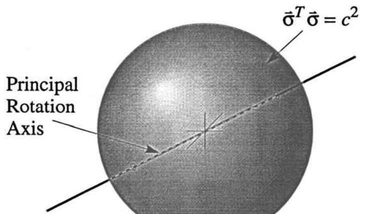

Fig. 4 Lemma illustration.

as shown in Fig. 3. The variation $\delta\boldsymbol{\sigma}_-$ must be such that $\boldsymbol{\sigma}_-$ still lies on the unit sphere. Using Eqs. (48) and (49), it can be shown that

$$\left(\boldsymbol{\sigma}_-^* + \tfrac{1}{2}\delta\boldsymbol{\sigma}_-\right)^T \delta\boldsymbol{\sigma}_- = 0 \tag{50}$$

The condition in Eq. (50) holds if the orthogonality is satisfied or if $\delta\boldsymbol{\sigma}_-$ is zero. The variation $\delta\boldsymbol{\sigma}_+$ must satisfy the same type of condition. For general rotations, Eq. (50) is satisfied through the orthogonality condition. However, for a single-axis rotation, Eq. (50) is satisfied by forcing $\delta\boldsymbol{\sigma}_-$ and $\delta\boldsymbol{\sigma}_+$ to be zero.

Lemma: The variations $\delta\boldsymbol{\sigma}_-$ and $\delta\boldsymbol{\sigma}_+$ must be zero if a single-axis rotation is being performed.

Proof: From the definition of the MRPs in Eq. (14) it is clear that for a single-axis rotation all MRP vectors will lie along the constant principal rotation vector. This straight line will touch the unit sphere constraint surface only at two points, as shown in Fig. 4. It is impossible to be at such a surface point and have a small variation while remaining on the constraint surface. □

Because a switch occurs at t_1, $\boldsymbol{\sigma}_+$ would be the shadow set of $\boldsymbol{\sigma}_-$. This mapping from $\boldsymbol{\sigma}_-$ to $\boldsymbol{\sigma}_+$ is well defined in Eq. (15); therefore, the variations $\delta\boldsymbol{\sigma}_-$ and $\delta\boldsymbol{\sigma}_+$ must be related. Their relative mapping is found by taking the first variation of Eq. (15),

$$\delta\boldsymbol{\sigma}_- = \boldsymbol{\sigma}_+^{-4}\left[2\boldsymbol{\sigma}_+\boldsymbol{\sigma}_+^T - \boldsymbol{\sigma}_+^T\boldsymbol{\sigma}_+ I\right]\delta\boldsymbol{\sigma}_+ \tag{51}$$

which can be further reduced to the following useful form using Eq. (15):

$$\delta\boldsymbol{\sigma}_- = \left[2\boldsymbol{\sigma}_-\boldsymbol{\sigma}_-^T - \boldsymbol{\sigma}_-^T\boldsymbol{\sigma}_- I\right]\delta\boldsymbol{\sigma}_+ \tag{52}$$

Let us first examine the case where $\delta\boldsymbol{\sigma}_-$ and $\delta\boldsymbol{\sigma}_+$ are not zero. In this case Eq. (45) shows that

$$\boldsymbol{\Lambda}_{\sigma^+}^T \delta\boldsymbol{\sigma}_+ = \boldsymbol{\Lambda}_{\sigma^-}^T \delta\boldsymbol{\sigma}_- \tag{53}$$

which can be expanded using Eq. (52) to

$$\left(\boldsymbol{\Lambda}_{\sigma^+} - \left[2\boldsymbol{\sigma}_-\boldsymbol{\sigma}_-^T - \boldsymbol{\sigma}_-^T\boldsymbol{\sigma}_- I\right]\boldsymbol{\Lambda}_{\sigma^-}\right)^T \delta\boldsymbol{\sigma}_+ = 0 \tag{54}$$

Because Eq. (54) must hold for any admissible variations $\delta\boldsymbol{\sigma}_+$, the following costate switching condition is found:

$$\boldsymbol{\Lambda}_{\sigma^+} = \left[2\boldsymbol{\sigma}_-\boldsymbol{\sigma}_-^T - \left(\boldsymbol{\sigma}_-^T\boldsymbol{\sigma}_-\right) I\right]\boldsymbol{\Lambda}_{\sigma^-} \tag{55}$$

Note that Eq. (55) yields a general mapping for the switching of a costate $\boldsymbol{\Lambda}_{\sigma^-}$ to its shadow costate $\boldsymbol{\Lambda}_{\sigma^+}$. This mapping is valid for any switching condition $\boldsymbol{\sigma}^T\boldsymbol{\sigma} = c^2$, but has its simplest form if the switching surface $\boldsymbol{\sigma}^T\boldsymbol{\sigma} = 1$ is chosen.

Equation (47) provides another condition that must be satisfied for optimality. Because the attitude penalty functions $g\{[C(\boldsymbol{\sigma})]\}$ and $G(\boldsymbol{\sigma})$ are such that $p(t_1^-) = p(t_1^+)$, Eq. (47) can be further reduced using Eq. (46) to

$$\boldsymbol{\Lambda}_{\sigma^-}^T \dot{\boldsymbol{\sigma}}\left(t_1^-\right) = \boldsymbol{\Lambda}_{\sigma^+}^T \dot{\boldsymbol{\sigma}}\left(t_1^+\right) \tag{56}$$

By making use of the costate switching condition in Eq. (55) and of Eq. (16), this condition can be shown to be always true. Therefore, Eq. (47) provides no further information for the case where $\delta\boldsymbol{\sigma}_-$ and $\delta\boldsymbol{\sigma}_+$ are nonzero.

By our Lemma, if the optimal rotation is a single-axis rotation, then $\delta\boldsymbol{\sigma}_- \equiv 0$ and $\delta\boldsymbol{\sigma}_+ \equiv 0$. Because of this, Eq. (45) does not reveal any information about the costate $\boldsymbol{\Lambda}_\sigma$ at t_1. To derive the necessary costate switching condition for $\boldsymbol{\Lambda}_\sigma$ Eq. (47), which was reduced to Eq. (56), will be used.

Let $\hat{e}$ be the constant axis of rotation. Using Eq. (14) the attitude vector can be written as $\boldsymbol{\sigma} = \hat{\boldsymbol{e}}|\boldsymbol{\sigma}| = \hat{\boldsymbol{e}}\sigma$. The body angular velocity vector is given by $\boldsymbol{\omega} = \hat{\boldsymbol{e}}|\boldsymbol{\omega}| = \hat{\boldsymbol{e}}\omega$. From Eqs. (8), (11), and (19) it is clear that for a single-axis rotation the attitude costate $\boldsymbol{\Lambda}_\sigma$ can be written as $\boldsymbol{\Lambda}_\sigma = \hat{\boldsymbol{e}}|\boldsymbol{\Lambda}_\sigma| = \hat{\boldsymbol{e}}\Lambda_\sigma$. Using Eqs. (16) and (18) in Eq. (56), the following costate switching condition can be found for the single-axis rotation case:

$$\Lambda_{\sigma^+} = \left(\boldsymbol{\sigma}_-^T\boldsymbol{\sigma}_-\right)\Lambda_{\sigma^-} \tag{57}$$

This condition shows that the only instance for which $\boldsymbol{\Lambda}_\sigma$ does not have a discontinuity during the switching is the case of a single-axis rotation with the switching surface $\boldsymbol{\sigma}^T\boldsymbol{\sigma} = 1$. Note that even though the costate switching condition in Eq. (55) was not derived for the case of a pure single-axis rotation, it does simplify to Eq. (57) when a single-axis rotation is imposed. This allows the costate switching condition for both cases to be unified into one costate switching condition, Eq. (55). These developments prove the following theorem.

Theorem (MRP costate switching condition): Let the MRP switching surface be $\boldsymbol{\sigma}^T\boldsymbol{\sigma} = c^2$ and let the attitude penalty function be continuous with respect to $\boldsymbol{\sigma}$, then the costate $\boldsymbol{\Lambda}_\omega$ will remain continuous during the switching of the MRPs to their shadow set. The costate $\boldsymbol{\Lambda}_\sigma$, however, will have a discontinuity defined by Eq. (55).

Note that this theorem allows for the MRPs to switch on any surface $\boldsymbol{\sigma}^T\boldsymbol{\sigma} = c^2$. A numerical method would not have to pinpoint the time where this surface is penetrated. It is sufficient to monitor at each time step whether a surface penetration has occurred, meaning that $\boldsymbol{\sigma}^T\boldsymbol{\sigma} > c^2$. If yes, then the attitude state and costate would be switched as shown. This theorem leads directly to the following corollary regarding the costate magnitude $|\boldsymbol{\Lambda}_\sigma|$ during the MRP switching.

Corollary: The costate magnitude $|\boldsymbol{\Lambda}_\sigma|$ will remain continuous during the MRP switching if the $\boldsymbol{\sigma}^T\boldsymbol{\sigma} = 1$ switching surface is used.

Proof: This corollary is verified by using the theorem to find $\boldsymbol{\Lambda}_\sigma^T\boldsymbol{\Lambda}_\sigma$ before and after the MRP switching and using the fact that during the switching $\boldsymbol{\sigma}^T\boldsymbol{\sigma}$ is equal to 1. □

This corollary shows that the MRP costates $\boldsymbol{\Lambda}_\sigma$ behave very similarly to the MRP vector $\boldsymbol{\sigma}$ during the switching. Both switch on the surface of a sphere. The difference is that $\boldsymbol{\sigma}$ switches on a unit sphere, where $\boldsymbol{\Lambda}_\sigma$ switches on a sphere of arbitrary radius.

Instead of using the calculus of variations development, the MRP costate switching conditions can also be derived by transforming the given optimal control problem with internal discontinuous states into a standard optimal control problem with a terminal constraint surface. This is accomplished by doubling the number of system states by assigning the states before and after the switch at t_1 to be separate, independent states. The states before the switch are integrated forward in time on the interval $[0, t_1]$, where the states after the switch are integrated backward in time on $[t_f, t_1]$. By folding the time around the state discontinuity, the previously internal MRP switching at t_1 now effectively becomes a terminal state constraint.[13] The MRP costate switching conditions are then found by using the standard necessary conditions for a continuous optimal control problem with the terminal state constraint surfaces[9] $\boldsymbol{\sigma}^S(t_1^+) + \boldsymbol{\sigma}(t_1^-)/\sigma^2(t_1^-) = 0$ and $\boldsymbol{\omega}(t_1^+) - \boldsymbol{\omega}(t_1^-) = 0$.

VI. Single-Axis Analytical Result

To verify the MRP costate switching conditions, a simple single-axis optimal control problem is solved analytically using the MRPs as attitude parameters. For generality, the switching surface is set to $\sigma^2 = c^2$. Let us minimize the cost function J, which depends solely on the control u,

$$J = \frac{1}{2}\int_0^1 u^2\,\mathrm{d}t \tag{58}$$

subject to the simple one-dimensional equations of motion for a body with unit inertia

$$\dot{\sigma} = \tfrac{1}{4}(1+\sigma^2)\omega \tag{59}$$

$$\dot{\omega} = u \tag{60}$$

and subject to the state constraints

$$\sigma(t_0 = 0) = \sigma_0 \qquad \sigma(t_f = 1) = \sigma_f \qquad \omega(t_0) = \omega(t_f) = 0 \tag{61}$$

The optimal control torque u^* for this cost function J is known to be of the form[10]

$$u^*(t) = k(1-2t) \tag{62}$$

where k is simply a scaling factor that guarantees that the body is at σ_f at t_f. Note that the optimal trajectory is independent of the choice of attitude coordinates. This allows the optimal control problem to be solved using either the original MRP set or the shadow set. By comparing the resulting costate Λ_σ history for the original and shadow costates, we verify next the costate switching condition for single-axis rotations in Eq. (57). The optimality condition in Eq. (10) states that

$$u^*(t) = -\Lambda_\omega(t) \tag{63}$$

Because u^* is continuous, so is Λ_ω, as predicted in Eq. (46). If at some point in time the MRPs are switched to their shadow set it obviously has no effect on the continuity of Λ_ω.

To find a time history of the costate Λ_σ, Eq. (9) is used:

$$\dot{\Lambda}_\omega = -\tfrac{1}{4}(1+\sigma^2)\Lambda_\sigma \tag{64}$$

Because $\dot{\Lambda}_\omega = 2k$ this can be solved for Λ_σ:

$$\Lambda_\sigma = -[8k/(1+\sigma^2)] \tag{65}$$

Let σ^S and Λ_σ^S be the shadow attitude and costate vector. Analogous to the preceding, the solution for Λ_σ^S is

$$\Lambda_\sigma^S = -\frac{8k}{1+(\sigma^S)^2} \tag{66}$$

Equation (66) can be written in terms of σ by using Eq. (15):

$$\Lambda_\sigma^S = -\frac{8k}{1+(1/\sigma^2)} = -\frac{8k}{1+\sigma^2}\sigma^2 \tag{67}$$

Substituting Eq. (65) into Eq. (67) a direct relationship between Λ_σ^S and Λ_σ is obtained:

$$\Lambda_\sigma^S = \sigma^2\Lambda_\sigma \tag{68}$$

By switching between the two possible MRP attitude descriptions the costate Λ_σ must be switched as well according to Eq. (68). This result verifies the single-axis rotation MRP costate switching condition found in Eq. (57).

VII. Three-Dimensional Numerical Result

To verify the general transformation given in the MRP costate switching condition theorem, a three-dimensional optimal control problem was solved as outlined in the problem statement. The attitude penalty function was chosen to be the $g()$ given in Eq. (21). With this penalty function the answer did not depend on the attitude coordinate choice. Therefore, the optimal solution using the combined set of $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}^S$ should be the same as the optimal solution obtained by using only $\boldsymbol{\sigma}$ or $\boldsymbol{\sigma}^S$.

The optimization problem was solved numerically by a steepest descent method.[9] The only modification needed to use the combined set of original and shadow MRP vectors was to check whether $\boldsymbol{\sigma}^T\boldsymbol{\sigma}$ had grown larger than one. If yes, then the attitude vector was switched to its shadow counterpart. At the same time the corresponding attitude costate vector was also switched to its shadow counterpart using the MRP costate switching condition.

The three-dimensional optimal control problem had a fixed maneuver time of $t_f = 10$ s. The body inertia matrix was $\Im = \mathrm{diag}(0.5, 1.0, 0.7)$ kgm^2. The cost function weights were $K_1 = 2$, $K_2 = 10$, $K_3 = 1$, $K_4 = 5$, and $R = 20$. The initial states were $\boldsymbol{\sigma}(0) = (0.87, 0, 0)$ and $\boldsymbol{\omega}(0) = (80.21, 51.57, 45.84)$ deg/s. Note that the initial orientation has the body almost turned upside down with a large initial angular velocity driving it to the upside-down orientation. This optimal control problem penalizes any nonzero state and torque during the maneuver and any nonzero final state. Note that the final state is left free though. Trying to minimize torque for this maneuver, it is intuitively reasonable to let the body rotate through the upside-down orientation and then reduce the states instead of forcefully reversing the existing motion. We show key results in Figs. 5–8.

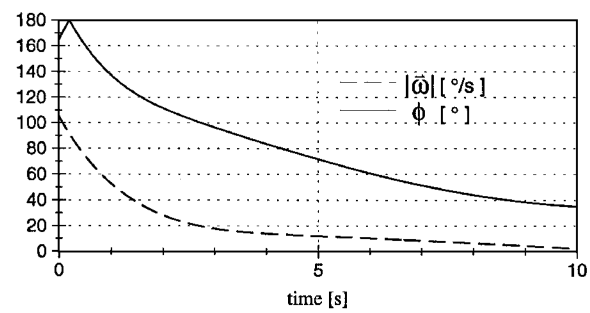

Fig. 5 Optimal states for all three cases.

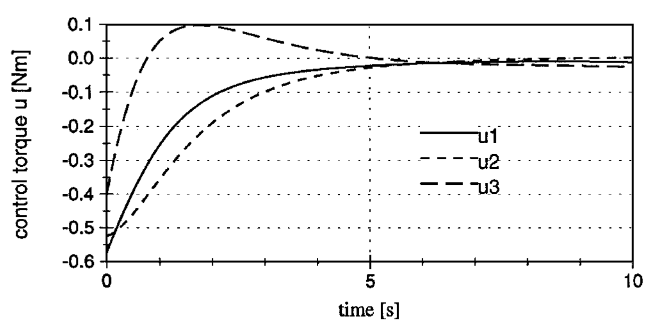

Fig. 6 Optimal control torque *u*.

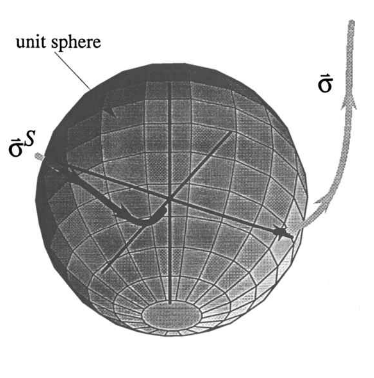

Fig. 7 Three-dimensional illustration of attitude vectors.

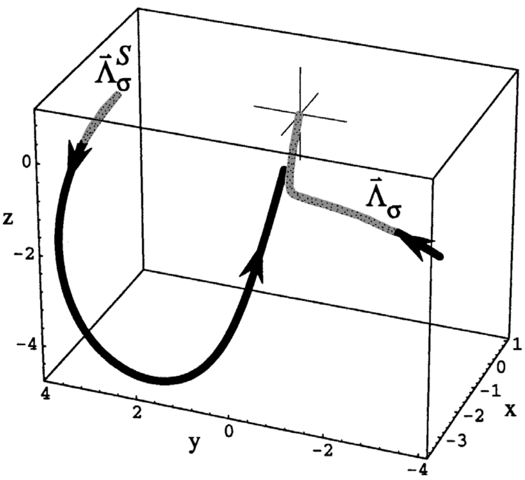

Fig. 8 Three-dimensional illustration of costate Λ_σ.

Three separate optimal control problems were solved using σ/σ^S, σ, or σ^S as the attitude coordinates. As expected, all three optimizations converged to the same solution. The principal rotation angle χ and the magnitude of the angular velocity are shown in Fig. 5. The optimal solution indeed lets the body rotate through the $\chi = 180$ deg point and diminishes simultaneously the angular velocity and attitude errors as the final maneuver time is approached.

The optimal control torque for the maneuver is shown in Fig. 6. The attitude coordinate vector time histories were different for each problem because different attitude coordinates were used. The combined set σ/σ^S started out identical to σ, because the initial attitude vector had less than unit magnitude. As $|\sigma|$ grew larger than one, the combined set σ/σ^S trajectory is switched to the shadow set σ^S trajectory. This is shown in Fig. 7. The black line denotes the trajectory of the combined σ and σ^S set, which remains within the unit sphere. Note that this trajectory converged exactly with the σ and σ^S trajectories whenever they, too, were within the unit sphere.

The ultimate test of the MRP costate switching condition theorem is to see if the costate Λ_σ exhibits the same behavior. The costate trajectories are shown in Fig. 8. Again the black line is the solution obtained using the combined σ/σ^S set and using the MRP costate switching condition. Indeed, the costate Λ_σ switches exactly from the costate trajectory of the pure σ solution to the costate trajectory of the pure σ^S solution, thus verifying the theorem presented.

VIII. Conclusion

A universal attitude penalty function $g(\,)$ for optimal control problems is presented that makes the optimization independent of the choice on attitude coordinates. This function also has other beneficial properties such as being bounded between 0 and 1 and being nonsingular. Another attitude penalty function $G(\,)$ was presented that made use of many good properties of the MRP. The MRP costate switching condition introduced makes the use of MRPs combined with their shadow set possible for general motion optimal control problems.

Acknowledgments

This work was sponsored by the U.S. Department of Energy under Contract DE-AC04-94AL85000. The authors would also like to thank the reviewers for their very insightful comments.

References

[1] Schaub, H., and Junkins, J. L., "Stereographic Orientation Parameters for Attitude Dynamics: A Generalization of the Rodrigues Parameters," *Journal of the Astronautical Sciences*, Vol. 44, No. 1, 1996, pp. 1–19.

[2] Tsiotras, P., "Stabilization and Optimality Results for the Attitude Control Problem," *Journal of Guidance, Control, and Dynamics,* Vol. 19, No. 4, 1996, pp. 772–779.

[3] Schaub, H., Tsiotras, P., and Junkins, J. L., "Principal Rotation Representations of Proper N × N Orthogonal Matrices," *International Journal of Engineering Science*, Vol. 33, No. 15, 1995, pp. 2277–2295.

[4] Marandi, S. R., and Modi, V. J., "A Preferred Coordinate System and the Associated Orientation Representation in Attitude Dynamics," *Acta Astronautica*, Vol. 15, No. 15, 1987, pp. 833–843.

[5] Wiener, T. F., *Theoretical Analysis of Gimballess Inertial Reference Equipment Using Delta-Modulated Instruments*, Ph.D. Dissertation, Dept. of Aeronautics and Astronautics, Massachusetts Inst. of Technology, Cambridge, MA, March 1962.

[6] Shuster, M. D., "A Survey of Attitude Representations," *Journal of the Astronautical Sciences*, Vol. 41, No. 4, 1993, pp. 439–517.

[7] Kirk, D. E., *Optimal Control Theory, An Introduction*, Networks Series, Prentice–Hall, Englewood Cliffs, NJ, 1970, pp. 184–236.

[8] Luenberger, D. G., *Optimization by Vector Space Methods*, Wiley, New York, 1969, pp. 254–270.

[9] Bryson, A. E., and Ho, Y., *Applied Optimal Control*, Hemisphere, Washington, DC, 1975, pp. 106–108.

[10] Junkins, J. L., and Turner, J. D., *Optimal Spacecraft Rotational Maneuvers*, Elsevier Science, Amsterdam, 1986, pp. 293–301.

[11] Vadali, S. R., "On the Euler Parameter Constraint," *Journal of Astronautical Sciences*, Vol. 36, No. 3, 1988, pp. 259–265.

[12] Leitmann, G., *The Calculus of Variations and Optimal Control*, Plenum, New York, 1981.

[13] Denbow, C. H., "A Generalized Form of the Problem of Bolza," *Contributions to the Calculus of Variations (1933–1937)*, Univ. of Chicago Press, Chicago, IL, 1937.