Analytical Mechanics of Aerospace Systems

Hanspeter Schaub
and
John L. Junkins

January 1, 2002

Contents

Preface

Part I

BASIC MECHANICS

Chapter One

Particle Kinematics

KINEMATICS is a branch of dynamics that studies aspects of motion apart from considerations of masses and forces. Essentially, Kinematics is a collection of vector/matrix methods to *describe* positions, velocities and accelerations of particles and rigid bodies, as viewed from various *reference frames*. The sub-field of Particle Kinematics considers only the motion of particles. This in itself can be quite challenging at times. As an example, consider a person driving a car on the highway. The road itself is fixed to a constantly rotating Earth which in turn is orbiting the sun. What is your velocity and acceleration relative to a Sun-fixed coordinate system? This chapter will help answer these and many related questions.

1.1 Particle Position Description

1.1.1 Basic Geometry

When studying the kinematics of particle motion, one is not concerned about the physical dimensions or mass of a particle. Let P be a point in a three-dimensional space as illustrated in Figure 1.1. To define the position of the point P, a coordinate system along with its origin must be chosen. Without this coordinate system, it is difficult to describe the position of point P. To visualize this problem, imagine one person A telling another person B that their location is "10 miles." Without knowing from what reference point person A measured 10 miles and in what direction it was measured, it is impossible for person B to know the meaning of "10 miles."

A coordinate system is defined by two things. First, a coordinate system origin O must be established to specify its position in space. Second, the orientation of the coordinate system must be chosen. By choosing the orientation of the coordinate system a person will know what is considered "up" or "east" as measured within this coordinate system. Three perpendicular (or orthogonal) right-handed unit vectors are traditionally used to denote unit displacement

directions along the orthogonal axes. In Figure 1.1 a standard cartesian coordinate system labeled as $\mathcal{E}$ is shown. The three unit vectors $\hat{\boldsymbol{e}}_1$, $\hat{\boldsymbol{e}}_2$ and $\hat{\boldsymbol{e}}_3$ are used to define the orientation of $\mathcal{E}$ and the coordinate system origin is denoted by $O_{\mathcal{E}}$. We will label all unit vectors with a ($\hat{}$) symbol. When assigning the unit vectors to the coordinate system, the first two unit vectors typically span the local "horizontal plane," while the third unit vector points in the "upwards" direction normal to the plane of the first two unit vectors. However, this sequence and interpretation is not required.

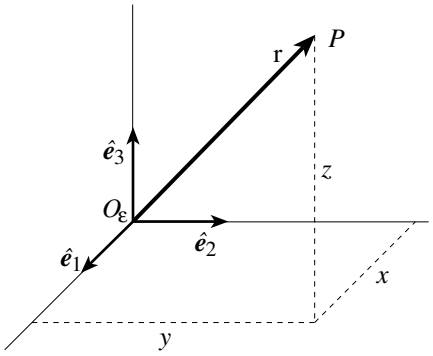


Figure 1.1: The Cartesian Coordinate System

A coordinate system, defined through the origin and the three unit direction vectors, is often referred to as a *reference frame*. Vectors with components taken in different coordinate systems are said to be written in different reference frames. More generally, think of a reference frame as a rigid body. While the Earth is a rigid body, there is an infinite set of coordinate systems that could be embedded in the Earth-fixed reference frame. For the present, we will usually associate only one coordinate system with a reference frame (rigid body).

Let $\boldsymbol{r} = \overline{O_{\mathcal{E}}P}$ be the vector pointing from the coordinate origin $O_{\mathcal{E}}$ to the point P. Note that there are an infinite number of ways to parameterize that vector in terms of orthogonal coordinate axis components. To write the position vector $\boldsymbol{r}$ in the cartesian coordinate system $\mathcal{E}$ shown in Figure 1.1, it is broken down (i.e. projected orthogonally) into the three components along the coordinate system unit axes. Let the $\hat{\boldsymbol{e}}_1$ component of $\boldsymbol{r}$ be called x, the $\hat{\boldsymbol{e}}_2$ component be called y and the $\hat{\boldsymbol{e}}_3$ component be called z. Then the vector $\boldsymbol{r}$ is written in the $\mathcal{E}$ cartesian coordinate system components as

$$\boldsymbol{r} = {}^{\mathcal{E}}\boldsymbol{r} = x\hat{\boldsymbol{e}}_1 + y\hat{\boldsymbol{e}}_2 + z\hat{\boldsymbol{e}}_3 \tag{1.1}$$

The short hand notation ${}^{\mathcal{E}}\boldsymbol{r}$ is used when we wish to specify that the vector components of $\boldsymbol{r}$ are taken along the unit directions vectors of the $\mathcal{E}$ coordinate system. The superscript coordinate system label is often omitted when it is clear in which system the components are taken or, more likely, one wishes to preserve the freedom to choose a particular coordinate system at a later point.

When it is clear in context, we can also use ${}^{\mathcal{E}}\boldsymbol{r}$ to denote the $\mathcal{E}$ frame base vector components of $\boldsymbol{r}$ as the 3x1 column vector (matrix)

$$ {}^{\mathcal{E}}\boldsymbol{r} = {}^{\mathcal{E}}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{1.2} $$

For cartesian coordinate systems, the i-th entry of the column vector is the component of the $\boldsymbol{r}$ vector along the i-th unit vector $\hat{\boldsymbol{e}}_i$.

Care must be taken when performing vector operations if multiple coordinate systems are used. Writing a vector addition as

$$ \boldsymbol{q} = \boldsymbol{r} + \boldsymbol{p} $$

is correct since no coordinate systems have been assigned yet; this equation has an infinity of possible component descriptions. We mention that one of the subtle and powerful facts of vector algebra is the ability to derive vector equations that hold for all possible component parameterizations of the vectors. However, if the vectors have specific coordinate systems components as shown in Eq. (1.2), then the following matrix vector addition *would not be correct*.

$$ {}^{\mathcal{E}}\begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} = {}^{\mathcal{E}}\begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} + {}^{\mathcal{B}}\begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} $$

The vector $\boldsymbol{p}$ is here written in $\mathcal{B}$ frame components while all other vectors are expressed in the $\mathcal{E}$ frame. To add the $\mathcal{B}$ frame components of the $\boldsymbol{p}$ vector to $\mathcal{E}$ frame vectors, these components would first have to be transformed (projected) from the $\mathcal{B}$ frame to the $\mathcal{E}$ frame. Later on in Chapter 3 it will be shown how the direction cosine matrix can be used to perform this transformation.

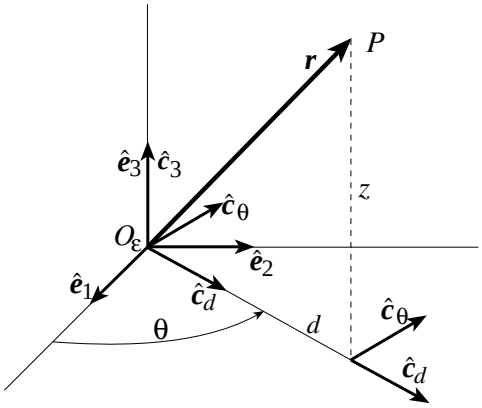


Figure 1.2: The Cylindrical Coordinate System

1.1.2 Cylindrical and Spherical Coordinate Systems

While the cartesian coordinate system is the most common and the easiest one to visualize, many times it is not the easiest to use. This is particularly true if the motion of Point P is of a rotational type or if the dominant forces are radial. In these cases it is usually easier to use either a cylindrical or spherical coordinate system. When we address dynamics in Chapter 2, we will provide some insight on coordinate system selection in the context of solving example problems.

A cylindrical coordinate system $\mathcal{C}$ is illustrated in Figure 1.2. Its orientation is defined through the triad of unit vectors $\{\hat{\boldsymbol{c}}_d, \hat{\boldsymbol{c}}_\theta, \hat{\boldsymbol{c}}_3\}$. This system is particularly useful in describing particles rotating about an axis $\hat{\boldsymbol{e}}_3$ which are free to move parallel to the axis $\hat{\boldsymbol{e}}_3$. For a large number of problems having rotational symmetry of force fields or constraint surfaces, cylindrical coordinates would be an attractive choice. For example, consider a particle constrained to move on the surface of a cylinder. Contrary to the inertially fixed cartesian coordinate system $\mathcal{N}$, two unit orientation vectors of the cylindrical coordinate system are varying with θ as seen from $\mathcal{N}$. These are the unit vector $\hat{\boldsymbol{c}}_d$ and $\hat{\boldsymbol{c}}_\theta$. They rotate in the horizontal plane perpendicular to the $\hat{\boldsymbol{c}}_3$ unit vector. The vector $\hat{\boldsymbol{c}}_d$ tracks the heading of the projection of the $\boldsymbol{r}$ position vector in this horizontal plane. The position vector $\boldsymbol{r}$ of point P is expressed in cylindrical coordinates as

$$\boldsymbol{r} = {}^{\mathcal{C}}\boldsymbol{r} = d\hat{\boldsymbol{c}}_d + z\hat{\boldsymbol{c}}_3 = \begin{matrix}\mathcal{C}\\ \end{matrix}\begin{pmatrix} d \\ 0 \\ z \end{pmatrix} \tag{1.3}$$

where the scalar d is the radial distance of point P from the $\hat{\boldsymbol{c}}_3$ axis. The second entry of the cylindrical coordinate system column vector in Eq. (1.3) will always be zero. Any particle position vector expressed in a cylindrical coordinate system will never have a component along the $\hat{\boldsymbol{c}}_\theta$ direction. Note that in Eq. (1.3) the unit vector $\hat{\boldsymbol{c}}_d$ has a variable direction as observed from $\mathcal{N}$. The angle θ describes how far $\hat{\boldsymbol{c}}_d$ has rotated from the $\hat{\boldsymbol{e}}_1$ axis. Therefore, instead of using (x, y, z) cartesian coordinates to describe a position, cylindrical coordinates use d and z, and the angle θ provides the azimuth angle of the unit vector $\hat{\boldsymbol{c}}_d$ relative to $\hat{\boldsymbol{e}}_1$. Assuming $\hat{\boldsymbol{c}}_3$ is aligned with $\hat{\boldsymbol{e}}_3$, the unit vectors $\hat{\boldsymbol{c}}_d$ and $\hat{\boldsymbol{c}}_\theta$ can be related to $\hat{\boldsymbol{e}}_1$ and $\hat{\boldsymbol{e}}_2$ through

$$\hat{\boldsymbol{c}}_d = \cos\theta\hat{\boldsymbol{e}}_1 + \sin\theta\hat{\boldsymbol{e}}_2 \tag{1.4a}$$

$$\hat{\boldsymbol{c}}_\theta = -\sin\theta\hat{\boldsymbol{e}}_1 + \cos\theta\hat{\boldsymbol{e}}_2 \tag{1.4b}$$

A spherical coordinate system $\mathcal{S}$ is illustrated in Figure 1.3 with its orientation defined through the triad of unit vectors $\{\hat{\boldsymbol{s}}_r, \hat{\boldsymbol{s}}_\theta, \hat{\boldsymbol{s}}_\phi\}$. Note that all three unit orientation vectors are time varying for the spherical coordinate system as seen from $\mathcal{N}$. The unit vector $\hat{\boldsymbol{s}}_r$ now points from $\mathcal{O}_\epsilon$ towards point P. Let the scalar r be the radial distance from the coordinate system center $O_\mathcal{E}$ to the point P. Then the position vector $\boldsymbol{r}$ is expressed as components along the

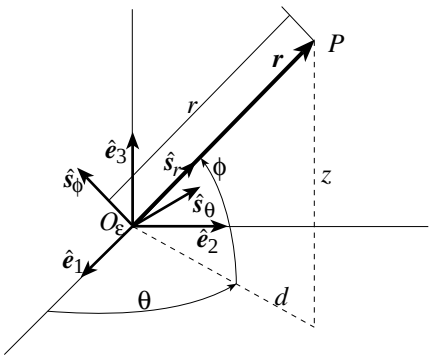


Figure 1.3: The Spherical Coordinate System

spherical coordinate triad $\{\hat{\boldsymbol{s}}_r, \hat{\boldsymbol{s}}_\theta, \hat{\boldsymbol{s}}_\phi\}$ as

$$\boldsymbol{r} = {}^{\mathcal{S}}\boldsymbol{r} = r\hat{\boldsymbol{s}}_r = \begin{pmatrix} r \\ 0 \\ 0 \end{pmatrix}^{\mathcal{S}} \tag{1.5}$$

A particle position vector written as a column vector with components taken in the $\mathcal{S}$ frame will have a non-zero entry only in the first position. As shown in Figure 1.3, the two angles θ and ϕ completely describe the orientation of the unit vectors $\hat{\boldsymbol{s}}_r$, $\hat{\boldsymbol{s}}_\theta$ and $\hat{\boldsymbol{s}}_\phi$ relative to the three $\hat{\boldsymbol{e}}_i$ (i = 1,2,3). Therefore, the $\{\hat{\boldsymbol{s}}_r, \hat{\boldsymbol{s}}_\theta, \hat{\boldsymbol{s}}_\phi\}$ projection onto $\{\hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_2, \hat{\boldsymbol{e}}_3\}$ with components a function of (r, θ, ϕ) are

$$\hat{\boldsymbol{s}}_r = \cos\phi\cos\theta\hat{\boldsymbol{e}}_1 + \cos\phi\sin\theta\hat{\boldsymbol{e}}_2 + \sin\phi\hat{\boldsymbol{e}}_3 \tag{1.6a}$$

$$\hat{\boldsymbol{s}}_\theta = -\sin\theta\hat{\boldsymbol{e}}_1 + \cos\theta\hat{\boldsymbol{e}}_2 \tag{1.6b}$$

$$\hat{\boldsymbol{s}}_\phi = -\sin\phi\cos\theta\hat{\boldsymbol{e}}_1 - \sin\phi\sin\theta\hat{\boldsymbol{e}}_2 + \cos\phi\hat{\boldsymbol{e}}_3 \tag{1.6c}$$

Spherical coordinates and the associated triad of unit vectors $\{\hat{\boldsymbol{s}}_r, \hat{\boldsymbol{s}}_\theta, \hat{\boldsymbol{s}}_\phi\}$ are very useful when describing a particle motion on the surface of a sphere or a particle orbiting a body.

Example 1.1: Given a vector $\boldsymbol{r}$ written in the cartesian coordinate system $\mathcal{E}$ as

$$\boldsymbol{r} = {}^{\mathcal{E}}\boldsymbol{r} = 2\hat{\boldsymbol{e}}_1 - 3\hat{\boldsymbol{e}}_2 + 5\hat{\boldsymbol{e}}_3$$

Express $\boldsymbol{r}$ in terms of the cylindrical coordinate system $\mathcal{C}$ where $\hat{\boldsymbol{c}}_3 = \hat{\boldsymbol{e}}_3$. From Eqs. (1.4), we can express $\hat{\boldsymbol{e}}_1$ and $\hat{\boldsymbol{e}}_2$ in terms of $\hat{\boldsymbol{c}}_d$ and $\hat{\boldsymbol{c}}_\theta$ as

$$\hat{\boldsymbol{e}}_1 = \cos\theta\hat{\boldsymbol{c}}_d - \sin\theta\hat{\boldsymbol{c}}_\theta$$
$$\hat{\boldsymbol{e}}_2 = \sin\theta\hat{\boldsymbol{c}}_d + \cos\theta\hat{\boldsymbol{c}}_\theta$$

Using this relationship the vector $\boldsymbol{r}$ is expressed in the $\mathcal{C}$ frame as

$$\boldsymbol{r} = {}^{\mathcal{C}}\boldsymbol{r} = (2\cos\theta - 3\sin\theta)\,\hat{\boldsymbol{c}}_d - (2\sin\theta + 3\cos\theta)\,\hat{\boldsymbol{c}}_\theta + 5\hat{\boldsymbol{c}}_3$$

The angle θ is resolved noting that in the $\mathcal{C}$ frame the $\hat{\boldsymbol{c}}_\theta$ component must be zero. Therefore θ must be

$$\theta = -\tan^{-1}\left(\frac{3}{2}\right) = -56.31^\circ$$

which brings ${}^{\mathcal{C}}\boldsymbol{r}$ to the desired result

$$\boldsymbol{r} = {}^{\mathcal{C}}\boldsymbol{r} = 3.61\hat{\boldsymbol{c}}_d + 5\hat{\boldsymbol{c}}_3$$

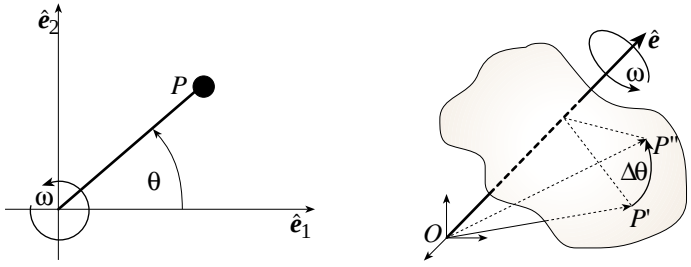
1.2 Vector Differentiation

1.2.1 Angular Velocity Vector

In planar motion it is easy to define and visualize the concept of angular velocity as is shown in Figure 1.4(i). For this single axis $\hat{\boldsymbol{e}}_3$ rotation case, the rotation angles and rotation rates (angular velocities) are only scalar quantities. The instantaneous angular rate ω of a particle is given by

$$\omega = \dot{\theta} \tag{1.7}$$

where a positive rotation or rotation rate is defined to be in the increasing θ (counterclockwise) direction shown. Angular velocity of a particle in a plane simply describes at what rate the radius vector locating the particle is orbiting the origin.



(i) Planar Case **(ii)** Three-Dimensional Case

Figure 1.4: The Angular Velocity Vector

For the general three dimensional case, we will prove in Chapter 3 that a general large angular displacement is not a vector quantity; however, paradoxically, angular velocity is a vector quantity. For the present, we limit ourselves

to an argument based upon small angular displacements to introduce the angular velocity vector. As the rigid body shown in Figure 1.4(ii) rotates about the body- and space-fixed $\hat{\boldsymbol{e}}$ axis by the small angle $\Delta\theta$, the body-fixed point at position P' rotates to position P''. This rotation is described through the rotation vector $\Delta\boldsymbol{\theta}$ defined as

$$\Delta\boldsymbol{\theta} = \Delta\theta\hat{\boldsymbol{e}} \tag{1.8}$$

The angular velocity vector is the instantaneous angular rate at which this rotation occurs. Let the angular velocity vector magnitude be ω, then the vector $\boldsymbol{\omega}$ can be written as

$$\boldsymbol{\omega} = \omega\hat{\boldsymbol{e}} \tag{1.9}$$

The unit direction vector $\hat{\boldsymbol{e}}$ defines an axis about which the rigid body or coordinate system is instantaneously rotating. For the case of planar rotations in Figure 1.4(i) the rotation axis is simply $\hat{\boldsymbol{e}}_3$. Note that any orientation of a rigid body can be defined by the orientation of any body-fixed coordinate system. Therefore position descriptions for rotating rigid bodies and rotating coordinate systems are actually the same problem geometrically and there is no need to formally distinguish between the two. For the case of constant $\hat{\boldsymbol{e}}$ it is natural to define

$$\boldsymbol{\omega} = \lim_{\Delta t \to 0} \frac{\Delta\boldsymbol{\theta}}{\Delta t} \tag{1.10}$$

The angular velocity vector $\boldsymbol{\omega}$ of a rigid body or coordinate system $\mathcal{B}$ relative to another coordinate system $\mathcal{N}$ is typically expressed in $\mathcal{B}$ frame components.

$$^{\mathcal{B}}\boldsymbol{\omega} = \omega_1\hat{\boldsymbol{b}}_1 + \omega_2\hat{\boldsymbol{b}}_2 + \omega_3\hat{\boldsymbol{b}}_3 \tag{1.11}$$

Each component ω_i expresses the instantaneous angular rate of the body $\mathcal{B}$ about the i-th coordinate axis $\hat{\boldsymbol{b}}_i$ as shown in Figure 1.5. The ω_i components are obviously the orthogonal components of $\boldsymbol{\omega}$. As will be evident in Chapter 3, it is often convenient to describe $\boldsymbol{\omega}$ with non-orthogonal components as well.

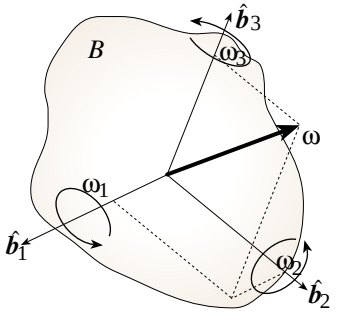


Figure 1.5: Illustration of Angular Velocity Body Frame Components

1.2.2 Rotation about a Fixed Axis

It is instructive to study in detail the rotation of a rigid body about a fixed axis. In particular, the velocity vector $\dot{\boldsymbol{r}}$ of a body-fixed point P is examined. Let a body $\mathcal{B}$ have a rod attached to it which is fixed in space at points A and B as shown in Figure 1.6 so the rod is the axis of rotation. The rigid body $\mathcal{B}$ is rotating about this rod with an angular velocity $\boldsymbol{\omega}$. The origin O of the coordinate system for $\mathcal{B}$ is located on the axis of rotation. Let P be a body-fixed point located relative to O by the vector $\boldsymbol{r}$. The angle between the angular velocity vector $\boldsymbol{\omega}$ and the position vector $\boldsymbol{r}$ is θ.

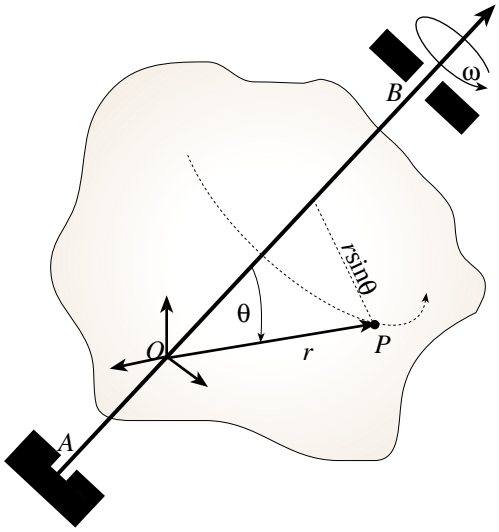


Figure 1.6: Rigid Body Rotation about a Fixed Axis

Studying Figure 1.6 it is quite clear that the body-fixed point P will have no velocity component parallel to the angular velocity vector $\boldsymbol{\omega}$; i.e., P moves in a plane perpendicular to the $\boldsymbol{\omega}$ axis. If one would look down the angular velocity vector one would see P moving on a circle with radius $r\sin\theta$ while being "transported" with the rotating rigid body. Thus the speed of P is given by

$$|\dot{\boldsymbol{r}}| = (r\sin\theta)\,\omega \tag{1.12}$$

Studying Figure 1.6 further it is apparent that the inertial velocity vector of P will always be normal to the plane of $\boldsymbol{r}$ and $\boldsymbol{\omega}$. This provides the direction of $\dot{\boldsymbol{r}}$ which can then be written as

$$\dot{\boldsymbol{r}} = (r\sin\theta)\,\omega\left(\frac{\boldsymbol{\omega}\times\boldsymbol{r}}{|\boldsymbol{\omega}\times\boldsymbol{r}|}\right) \tag{1.13}$$

However, note that $|\boldsymbol{\omega}\times\boldsymbol{r}| = \omega r\sin\theta$, so the transport velocity is

$$\dot{\boldsymbol{r}} = \boldsymbol{\omega}\times\boldsymbol{r} \tag{1.14}$$

The only restriction for Eq. (1.14) is that $\boldsymbol{r}$ must be a *body-fixed* vector within $\mathcal{B}$. As was mentioned earlier, the concepts of rigid bodies and reference frames can be used interchangeably. The above result would also hold if we are finding the velocity vector fixed to any reference frame which is rotating relative to another; as is evident below, this easily generalizes for three-dimensional motion.

1.2.3 Transport Theorem

As was mentioned earlier, it is simpler to define a particle position in terms of cylindrical or spherical coordinate systems. However, when computing the velocity of the particle and taking the time derivative of the position vector, one must take into account that the base vector directions of the chosen coordinate system may be time varying also. The following *transport theorem* allows one to take the derivative of a vector in one coordinate system, even though the vector itself has its components taken in another, possibly rotating, coordinate system.

Let $\mathcal{N}$ be an inertially fixed reference frame with a corresponding triad of $\mathcal{N}$-fixed orthogonal base vectors $\{\hat{\boldsymbol{n}}_1, \hat{\boldsymbol{n}}_2, \hat{\boldsymbol{n}}_3\}$. Let $\mathcal{B}$ be another reference frame with the $\mathcal{B}$-fixed base vectors $\{\hat{\boldsymbol{b}}_1, \hat{\boldsymbol{b}}_2, \hat{\boldsymbol{b}}_3\}$. For simplicity, let the origin of the two associated reference frames be coincident. Let ${}^{\mathcal{B}}\boldsymbol{r}$ be a vector written in the $\mathcal{B}$ coordinate system.

$$\boldsymbol{r} = {}^{\mathcal{B}}\boldsymbol{r} = r_1\hat{\boldsymbol{b}}_1 + r_2\hat{\boldsymbol{b}}_2 + r_3\hat{\boldsymbol{b}}_3 \tag{1.15}$$

We introduce the following notation: the angular velocity vector $\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}$ defines the angular velocity of the $\mathcal{B}$ frame relative to the $\mathcal{N}$ frame. An angular velocity vector is typically written in the $\mathcal{B}$ frame. Therefore we write $\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}$ as

$${}^{\mathcal{B}}\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}} = \omega_1\hat{\boldsymbol{b}}_1 + \omega_2\hat{\boldsymbol{b}}_2 + \omega_3\hat{\boldsymbol{b}}_3 \tag{1.16}$$

At this point we introduce the notion of taking the vector time derivative while accounting for the reference frame from which the vector's time variations are being observed. Imagine this: you are standing still on Earth's surface. Let $\mathcal{B}$ be an Earth fixed coordinate system with the origin in the center of the Earth. Your position vector would point from the Earth's center to your feet on the surface. By calculating the derivative of your position vector within $\mathcal{B}$, you are determining how quickly this vector changes direction and/or magnitude as seen from the $\mathcal{B}$ system. You would find the time variation of your position to be zero when viewed from the Earth-fixed frame. This should be no big surprise; after all, you are standing still and not walking around on Earth. Now, let's introduce another coordinate system $\mathcal{N}$ with the same origin, but this one is non-rotating and therefore fixed in space. Calculating the derivative of your position vector in the $\mathcal{N}$ frame, you wish to know how fast this vector is changing with respect to the *fixed coordinate system* $\mathcal{N}$. Since Earth itself is rotating, in this case your position derivative would be non-zero. This is because relative to $\mathcal{N}$, you are moving at constant speed along a circle about the Earth's spin axis.

To indicate that a derivative is taken of a generic vector $\boldsymbol{x}$ as seen in the $\mathcal{B}$ frame, we write

$$\frac{{}^{\mathcal{B}}d}{dt}(\boldsymbol{x})$$

The derivative of ${}^{\mathcal{B}}\boldsymbol{r}$ given in Eq. (1.15) with components taken in the $\mathcal{B}$ coordinate system is denoted by

$$\frac{{}^{\mathcal{B}}d}{dt}(\boldsymbol{r}) = \frac{{}^{\mathcal{B}}d}{dt}\left({}^{\mathcal{B}}\boldsymbol{r}\right) = \dot{r}_1\hat{\boldsymbol{b}}_1 + \dot{r}_2\hat{\boldsymbol{b}}_2 + \dot{r}_3\hat{\boldsymbol{b}}_3 \tag{1.17}$$

since the unit vectors $\hat{\boldsymbol{b}}_i$ are fixed (i.e. time invariant) within the $\mathcal{B}$ frame and therefore the terms ${}^{\mathcal{B}}d/dt\left(\hat{\boldsymbol{b}}_i\right)$ are zero. When taking the inertial derivative of ${}^{\mathcal{B}}\boldsymbol{r}$ however, these unit vectors must now be considered time varying as seen in $\mathcal{N}$. Therefore, using the chain rule of differentiation, we get[1]

$$\frac{{}^{\mathcal{N}}d}{dt}\left({}^{\mathcal{B}}\boldsymbol{r}\right) = \dot{r}_1\hat{\boldsymbol{b}}_1 + \dot{r}_2\hat{\boldsymbol{b}}_2 + \dot{r}_3\hat{\boldsymbol{b}}_3 + r_1\frac{{}^{\mathcal{N}}d}{dt}\left(\hat{\boldsymbol{b}}_1\right) + r_2\frac{{}^{\mathcal{N}}d}{dt}\left(\hat{\boldsymbol{b}}_2\right) + r_3\frac{{}^{\mathcal{N}}d}{dt}\left(\hat{\boldsymbol{b}}_3\right) \tag{1.18}$$

However, since $\hat{\boldsymbol{b}}_i$ are body-fixed vectors within $\mathcal{B}$, Eq. (1.14) can be used to find their derivative in $\mathcal{N}$.

$$\frac{{}^{\mathcal{N}}d}{dt}\left(\hat{\boldsymbol{b}}_i\right) = \boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}} \times \hat{\boldsymbol{b}}_i \tag{1.19}$$

Using Eqs. (1.17) and (1.19), Eq. (1.18) is rewritten as

$$\frac{{}^{\mathcal{N}}d}{dt}(\boldsymbol{r}) = \frac{{}^{\mathcal{N}}d}{dt}\left({}^{\mathcal{B}}\boldsymbol{r}\right) = \frac{{}^{\mathcal{B}}d}{dt}\left({}^{\mathcal{B}}\boldsymbol{r}\right) + \boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}} \times {}^{\mathcal{B}}\boldsymbol{r} \tag{1.20}$$

However, note that it is not necessary for the vector $\boldsymbol{r}$ to be written in the $\mathcal{B}$ coordinate frame for Eq. (1.20) to hold, because ${}^{\mathcal{B}}\boldsymbol{r}$ is simply one of the infinity of possible components of the unique vector $\boldsymbol{r}$. Rather, components can be written in any arbitrary coordinate frame. This result leads to the general form of the *transport theorem*.

Theorem 1.1 (Transport Theorem) *Let $\mathcal{N}$ and $\mathcal{B}$ be two frames with a relative angular velocity vector of $\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}$, and let $\boldsymbol{r}$ be a generic vector, then the derivative of $\boldsymbol{r}$ in the $\mathcal{N}$ frame can be related to the derivative of $\boldsymbol{r}$ in the $\mathcal{B}$ frame as*

$$\frac{{}^{\mathcal{N}}d}{dt}(\boldsymbol{r}) = \frac{{}^{\mathcal{B}}d}{dt}(\boldsymbol{r}) + \boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}} \times \boldsymbol{r} \tag{1.21}$$

This formula allows one to relate a vector derivative taken relative to frame $\mathcal{B}$ to the corresponding vector derivative taken in frame $\mathcal{N}$, where $\mathcal{B}$ and $\mathcal{N}$ are arbitrarily moving reference frames. This permits one to relate the derivative of $\boldsymbol{r}$ as it would be seen from the $\mathcal{N}$ frame to the analogous rate of change of $\boldsymbol{r}$ as seen in the $\mathcal{B}$ frame. It is a very fundamental and important result that is used

almost every time kinematic equations are derived. In particular, we will find that vectors are typically differentiated with respect to an inertial frame called $\mathcal{N}$. However, the notation ${}^{\mathcal{N}}d/dt\,(\boldsymbol{x})$ becomes cumbersome at times. When we want to compact the equation, we will use the following shorthand notation:

$$\frac{{}^{\mathcal{N}}d}{dt}(\boldsymbol{x}) \equiv \dot{\boldsymbol{x}} \tag{1.22}$$

Example 1.2: The inertial velocity and acceleration vectors are sought for a general planar motion described in terms of polar coordinates with components taken along $\{\hat{\boldsymbol{e}}_r, \hat{\boldsymbol{e}}_\theta, \hat{\boldsymbol{e}}_3\}$. The origin and base vectors of the polar coordinate system $\mathcal{E}$ are denoted

$$\mathcal{E} = \{O, \hat{\boldsymbol{e}}_r, \hat{\boldsymbol{e}}_\theta, \hat{\boldsymbol{e}}_3\}$$

as shown in Figure 1.7. The inertial coordinate system having the same origin O is denoted

$$\mathcal{N} = \{O, \hat{\boldsymbol{n}}_1, \hat{\boldsymbol{n}}_2, \hat{\boldsymbol{n}}_3\}$$

where $\hat{\boldsymbol{n}}_3 = \hat{\boldsymbol{e}}_3$. The position vector ${}^{\mathcal{E}}\boldsymbol{r}$ written in the $\mathcal{E}$ coordinate system is

$$\boldsymbol{r} = {}^{\mathcal{E}}\boldsymbol{r} = r\hat{\boldsymbol{e}}_r$$

Let $\boldsymbol{\omega}_{\mathcal{E}/\mathcal{N}}$ be the angular velocity vector of $\mathcal{E}$ with respect to $\mathcal{N}$. As is evident in Figure 1.7 this is simply

$$\boldsymbol{\omega}_{\mathcal{E}/\mathcal{N}} = \dot{\theta}\hat{\boldsymbol{e}}_3 = \dot{\theta}\hat{\boldsymbol{n}}_3$$

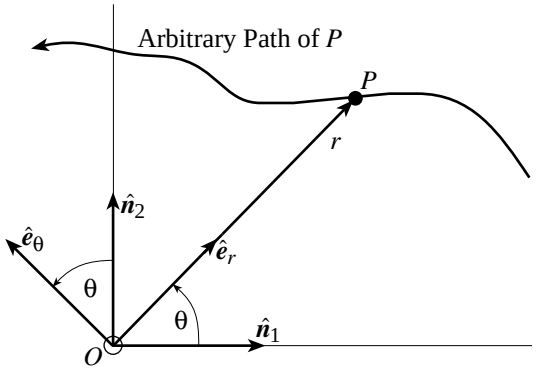


Figure 1.7: Polar Coordinates Illustration

Using the transport theorem in Eq. (1.21), the inertial velocity vector of $\boldsymbol{r}$ is found to be

$$\dot{\boldsymbol{r}} = \frac{{}^{\mathcal{E}}d}{dt}\left({}^{\mathcal{E}}\boldsymbol{r}\right) + \boldsymbol{\omega}_{\mathcal{E}/\mathcal{N}} \times {}^{\mathcal{E}}\boldsymbol{r}$$

Using the definition of ${}^{\mathcal{E}}\boldsymbol{r} = r\hat{\boldsymbol{e}}_r$ it is clear that

$$\frac{{}^{\mathcal{E}}d}{dt}(\boldsymbol{r}) = \frac{{}^{\mathcal{E}}d}{dt}(r\hat{\boldsymbol{e}}_r) = \dot{r}\hat{\boldsymbol{e}}_r$$

After carrying out the cross product term, the inertial velocity vector $\dot{\boldsymbol{r}}$ is reduced to

$$\dot{\boldsymbol{r}} = {}^{\mathcal{E}}\dot{\boldsymbol{r}} = \dot{r}\hat{\boldsymbol{e}}_r + r\dot{\theta}\hat{\boldsymbol{e}}_\theta \tag{1.23}$$

where $\dot{r}$ and $r\dot{\theta}$ are is the radial and the transverse velocity components, respectively.

The inertial acceleration $\ddot{\boldsymbol{r}}$ is found by taking the inertial derivative of $\dot{\boldsymbol{r}}$ using the transport theorem.

$$\ddot{\boldsymbol{r}} = \frac{{}^{\mathcal{E}}d}{dt}(\dot{\boldsymbol{r}}) + \boldsymbol{\omega}_{\mathcal{E}/\mathcal{N}} \times \dot{\boldsymbol{r}}$$

Using the result for $\dot{\boldsymbol{r}}$ that was just found, we obtain

$$\frac{{}^{\mathcal{E}}d}{dt}(\dot{\boldsymbol{r}}) = \ddot{r}\hat{\boldsymbol{e}}_r + \left(\dot{r}\dot{\theta} + r\ddot{\theta}\right)\hat{\boldsymbol{e}}_\theta$$

Again after carrying out the cross product and collecting terms, the inertial acceleration vector $\ddot{\boldsymbol{r}}$ is found to be

$$\ddot{\boldsymbol{r}} = {}^{\mathcal{E}}\ddot{\boldsymbol{r}} = \left(\ddot{r} - r\dot{\theta}^2\right)\hat{\boldsymbol{e}}_r + \left(r\ddot{\theta} + 2\dot{r}\dot{\theta}\right)\hat{\boldsymbol{e}}_\theta \tag{1.24}$$

where $\ddot{r}$ is the $radial$ component, $r\dot{\theta}^2$ is the $centrifugal$ component, $r\ddot{\theta}$ is the $tangential$ component and $2\dot{r}\dot{\theta}$ is the $coriolis$ acceleration component.

It is instructive to obtain Eq. (1.24) by "brute force." Notice we can write the $\mathcal{N}$ frame rectangular components of position, velocity and acceleration as

$$\begin{aligned}\boldsymbol{r} &= x\hat{\boldsymbol{n}}_1 + y\hat{\boldsymbol{n}}_2\\ \dot{\boldsymbol{r}} &= \dot{x}\hat{\boldsymbol{n}}_1 + \dot{y}\hat{\boldsymbol{n}}_2\\ \ddot{\boldsymbol{r}} &= \ddot{x}\hat{\boldsymbol{n}}_1 + \ddot{y}\hat{\boldsymbol{n}}_2\end{aligned}$$

Since $\hat{\boldsymbol{n}}_i$ are fixed in $\mathcal{N}$, the transport theorem is not required. Upon substituting the polar coordinate transformations

$$\begin{aligned}x &= r\cos\theta\\ y &= r\sin\theta\end{aligned}$$

and taking two time derivatives, you can obtain the lengthy trigonometric functions $a_x(r,\theta)$ and $a_y(r,\theta)$ in

$$\ddot{\boldsymbol{r}} = a_x(r,\theta)\hat{\boldsymbol{n}}_1 + a_y(r,\theta)\hat{\boldsymbol{n}}_2$$

Finally, substituting

$$\begin{aligned}\hat{\boldsymbol{n}}_1 &= \cos\theta\hat{\boldsymbol{e}}_r - \sin\theta\hat{\boldsymbol{e}}_\theta\\ \hat{\boldsymbol{n}}_2 &= \sin\theta\hat{\boldsymbol{e}}_r + \cos\theta\hat{\boldsymbol{e}}_\theta\end{aligned}$$

and performing considerable algebra, you will find all trigonometric functions of θ cancel, leaving you with the same result as in Eq. (1.24).

1.2.4 Particle Kinematics with Moving Frames

So far all coordinate systems or reference frames discussed were considered non-translating. Their origins were fixed inertially in space. Now a more general problem will be discussed where the coordinate frame origins are free to translate, while the frame orientations (defined through the three respective unit direction vectors) might be rotating.

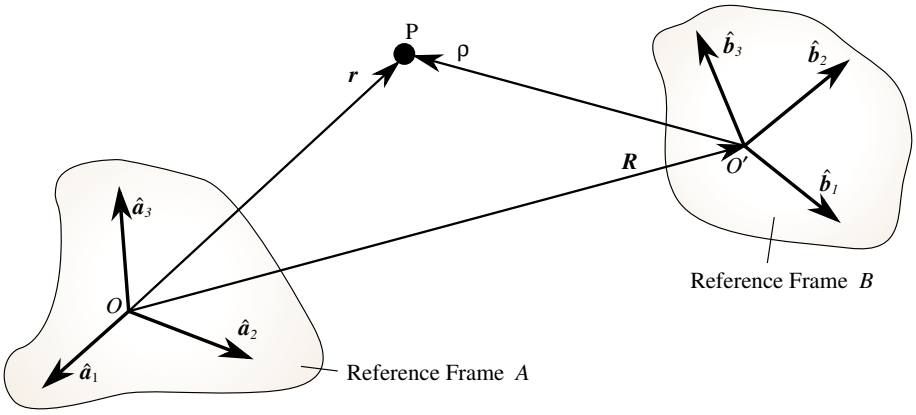


Figure 1.8: Two Coordinate Frames with Moving Origins

Let P be a generic particle in a three-dimensional space. Assume two different frames $\mathcal{A} = \{O, \hat{\boldsymbol{a}}_1, \hat{\boldsymbol{a}}_2, \hat{\boldsymbol{a}}_3\}$ and $\mathcal{B} = \{O', \hat{\boldsymbol{b}}_1, \hat{\boldsymbol{b}}_2, \hat{\boldsymbol{b}}_3\}$ exist as shown in Figure 1.8. The position of O' relative to O is given by the vector R. Note that these two coordinate frames could be actually attached to some rigid bodies and define their position and orientation in space, or they could simply be some artificial coordinate sets placed there without any other physical significance. In this discussion, however, it is frequently useful to think of reference frames $\mathcal{A}$ and $\mathcal{B}$ as rigid bodies.

Let the vectors $\boldsymbol{r}$ and $\boldsymbol{\rho}$ be the position vectors of particle P in the $\mathcal{A}$ and $\mathcal{B}$ frames respectively. The angular velocity vector of $\mathcal{B}$ relative to $\mathcal{A}$ is given by $\boldsymbol{\omega}_{\mathcal{B}/\mathcal{A}}$. Observe that the position vector $\boldsymbol{r}$ of P in the $\mathcal{A}$ frame can be related to $\boldsymbol{R}$ and $\boldsymbol{\rho}$ through the vector addition

$$\boldsymbol{r} = \boldsymbol{R} + \boldsymbol{\rho} \tag{1.25}$$

The shorthand notation $\left(\boldsymbol{v}^P\right)_{\mathcal{B}}$ is used in this section to express the velocity vector of particle P with the derivative taken relative to the $\mathcal{B}$ frame.

$$\left(\boldsymbol{v}^P\right)_{\mathcal{B}} \equiv \frac{{}^{\mathcal{B}}d}{dt}\left(\boldsymbol{\rho}\right) \tag{1.26}$$

The velocity vector $\left(\boldsymbol{v}^P\right)_{\mathcal{A}}$ of P relative to the $\mathcal{A}$ frame is given by

$$\left(\boldsymbol{v}^P\right)_{\mathcal{A}} \equiv \frac{{}^{\mathcal{A}}d}{dt}\left(\boldsymbol{r}\right) = \frac{{}^{\mathcal{A}}d}{dt}\left(\boldsymbol{R} + \boldsymbol{\rho}\right) = \frac{{}^{\mathcal{A}}d}{dt}\left(\boldsymbol{R}\right) + \frac{{}^{\mathcal{A}}d}{dt}\left(\boldsymbol{\rho}\right) \tag{1.27}$$

The velocity vector of the origin O' in the $\mathcal{A}$ frame is defined to be

$$\left(\boldsymbol{v}^{O'}\right)_{\mathcal{A}} = \frac{{}^{\mathcal{A}}d}{dt}\left(\boldsymbol{R}\right) \tag{1.28}$$

Using the transport theorem and the definition in Eq. (1.28), the velocity vector $\left(\boldsymbol{v}^{P}\right)_{\mathcal{A}}$ of Eq. (1.27) can be written as

$$\left(\boldsymbol{v}^{P}\right)_{\mathcal{A}} = \left(\boldsymbol{v}^{O'}\right)_{\mathcal{A}} + \left(\boldsymbol{v}^{P}\right)_{\mathcal{B}} + \boldsymbol{\omega}_{\mathcal{B}/\mathcal{A}} \times \boldsymbol{\rho} \tag{1.29}$$

To find the acceleration $\left(\boldsymbol{a}^{P}\right)_{\mathcal{A}}$ of particle P in the $\mathcal{A}$ frame, the derivative of Eq. (1.29) is taking in the $\mathcal{A}$ frame.

$$\left(\boldsymbol{a}^{P}\right)_{\mathcal{A}} = \frac{{}^{\mathcal{A}}d}{dt}\left(\left(\boldsymbol{v}^{P}\right)_{\mathcal{A}}\right) = \frac{{}^{\mathcal{A}}d}{dt}\left(\left(\boldsymbol{v}^{O'}\right)_{\mathcal{A}} + \left(\boldsymbol{v}^{P}\right)_{\mathcal{B}} + \boldsymbol{\omega}_{\mathcal{B}/\mathcal{A}} \times \boldsymbol{\rho}\right) \tag{1.30}$$

Allowing the differentiation operator to apply term-by-term in the last term, and using the transport theorem, $\left(\boldsymbol{a}^{P}\right)_{\mathcal{A}}$ becomes

$$\begin{aligned}\left(\boldsymbol{a}^{P}\right)_{\mathcal{A}} = \frac{{}^{\mathcal{A}}d}{dt}\left(\left(\boldsymbol{v}^{O'}\right)_{\mathcal{A}}\right) + \frac{{}^{\mathcal{B}}d}{dt}\left(\left(\boldsymbol{v}^{P}\right)_{\mathcal{B}}\right) + \boldsymbol{\omega}_{\mathcal{B}/\mathcal{A}} \times \left(\boldsymbol{v}^{P}\right)_{\mathcal{B}} + \\ \frac{{}^{\mathcal{A}}d}{dt}\left(\boldsymbol{\omega}_{\mathcal{B}/\mathcal{A}}\right) \times \boldsymbol{\rho} + \boldsymbol{\omega}_{\mathcal{B}/\mathcal{A}} \times \left(\frac{{}^{\mathcal{B}}d}{dt}\left(\boldsymbol{\rho}\right) + \boldsymbol{\omega}_{\mathcal{B}/\mathcal{A}} \times \boldsymbol{\rho}\right)\end{aligned} \tag{1.31}$$

Looking at the first term, the acceleration of the origin O' in the $\mathcal{A}$ frame is defined to be

$$\left(\boldsymbol{a}^{O'}\right)_{\mathcal{A}} = \frac{{}^{\mathcal{A}}d}{dt}\left(\left(\boldsymbol{v}^{O'}\right)_{\mathcal{A}}\right) \tag{1.32}$$

While looking at the second term, the acceleration of particle P in the $\mathcal{B}$ frame is

$$\left(\boldsymbol{a}^{P}\right)_{\mathcal{B}} = \frac{{}^{\mathcal{B}}d}{dt}\left(\left(\boldsymbol{v}^{P}\right)_{\mathcal{B}}\right) \tag{1.33}$$

The angular acceleration vector of the $\mathcal{B}$ frame relative to the $\mathcal{A}$ frame is defined to be

$$\boldsymbol{\alpha}_{\mathcal{B}/\mathcal{A}} = \frac{{}^{\mathcal{A}}d}{dt}\left(\boldsymbol{\omega}_{\mathcal{B}/\mathcal{A}}\right) \tag{1.34}$$

Using the definitions in Eqs. (1.26) and (1.32) – (1.34), the particle P acceleration vector $\left(\boldsymbol{a}^{P}\right)_{\mathcal{A}}$ can be written as the useful result[2]

$$\begin{aligned}\left(\boldsymbol{a}^{P}\right)_{\mathcal{A}} = \left(\boldsymbol{a}^{O'}\right)_{\mathcal{A}} + \left(\boldsymbol{a}^{P}\right)_{\mathcal{B}} + \boldsymbol{\alpha}_{\mathcal{B}/\mathcal{A}} \times \boldsymbol{\rho} + 2\boldsymbol{\omega}_{\mathcal{B}/\mathcal{A}} \times \left(\boldsymbol{v}^{P}\right)_{\mathcal{B}} + \\ \omega_{\mathcal{B}/\mathcal{A}} \times \left(\omega_{\mathcal{B}/\mathcal{A}} \times \boldsymbol{\rho}\right)\end{aligned} \tag{1.35}$$

The term $2\boldsymbol{\omega}_{\mathcal{B}/\mathcal{A}} \times \left(\boldsymbol{v}^P\right)_{\mathcal{B}}$ defines the coriolis acceleration and the term $\omega_{\mathcal{B}/\mathcal{A}} \times \left(\omega_{\mathcal{B}/\mathcal{A}} \times \boldsymbol{\rho}\right)$ is the centrifugal acceleration. The latter term can also be expressed as

$$\omega_{\mathcal{B}/\mathcal{A}} \times \left(\omega_{\mathcal{B}/\mathcal{A}} \times \boldsymbol{\rho}\right) = (\omega_{\mathcal{B}/\mathcal{A}} \cdot \boldsymbol{\rho})\omega_{\mathcal{B}/\mathcal{A}} - |\omega_{\mathcal{B}/\mathcal{A}}|^2 \boldsymbol{\rho} \tag{1.36}$$

which immediately reveals the centripical acceleration vector components along $\omega_{\mathcal{B}/\mathcal{A}}$ and $\boldsymbol{\rho}$. Note that Eq. (1.35) holds between *any two reference frames*. It is not necessary that $\mathcal{A}$ or $\mathcal{B}$ be inertially fixed. The vector components used in the various terms on the right hand side of Eq. (1.35) can be taken along *any* choice of unit vectors. It is important that we recognize the complete freedom we have to use any basis vectors we wish to express components of any vector in Eq. (1.35).

Example 1.3: A disk of radius ρ, attached to a rod of length L, is rolling on the inside of a circular tube of radius R as shown in Figure 1.9. The rod is rotating at constant rate $\omega = \dot{\theta}$. Three different reference frames are defined. The inertially fixed frame is $\mathcal{N} = \{O, \hat{\boldsymbol{n}}_1, \hat{\boldsymbol{n}}_2, \hat{\boldsymbol{n}}_3\}$ with the origin at the center of the tube. The second coordinate frame $\mathcal{E} = \{O, \hat{\boldsymbol{e}}_L, \hat{\boldsymbol{e}}_\theta, \hat{\boldsymbol{e}}_3\}$ has the same origin, but the direction axes track the center of disk O'. The third frame $\mathcal{B} = \{O', \hat{\boldsymbol{b}}_r, \hat{\boldsymbol{b}}_\phi, \hat{\boldsymbol{b}}_3\}$ has the origin in the center of the disk and the direction unit vectors track a point P on the disk edge. Note that $\hat{\boldsymbol{n}}_3$ and $\hat{\boldsymbol{e}}_3$ point out of the paper and $\hat{\boldsymbol{b}}_3 = -\hat{\boldsymbol{n}}_3$ points into the paper. What is the inertial acceleration $\ddot{\boldsymbol{r}}$ of point P expressed in $\mathcal{E}$ frame components? Note that since three frames are present, we cannot directly use Eq. (1.35). Instead the result will be derived by differentiation of the position vector by applying the transport theorem.

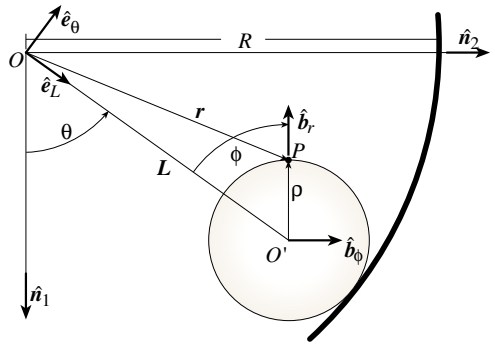


Figure 1.9: Disk Rolling inside Circular Tube

First, let's determine an expression for relating the angular rates $\dot{\phi}$ and $\dot{\theta} = \omega$. Since there is no slippage between the disk and the tube, then notice that the "contact arcs" must be equal on the tube and the cylinder, giving the constraint

$$\theta R = \phi \rho$$

Taking the derivative of the above expression and using $\dot{\theta} = \omega$, the term $\dot{\phi}$ is given as

$$\dot{\phi} = \frac{R}{\rho}\omega$$

The angular velocity vectors of frame $\mathcal{E}$ relative to $\mathcal{N}$ and frame $\mathcal{B}$ relative to $\mathcal{E}$ are

$$\boldsymbol{\omega}_{\mathcal{E}/\mathcal{N}} = \omega \hat{\boldsymbol{n}}_3$$
$$\boldsymbol{\omega}_{\mathcal{B}/\mathcal{E}} = \dot{\phi}\hat{\boldsymbol{b}}_3 = -\frac{R}{\rho}\omega\hat{\boldsymbol{n}}_3$$

The angular velocity vector of frame $\mathcal{B}$ relative to frame $\mathcal{N}$ is

$$\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}} = \boldsymbol{\omega}_{\mathcal{B}/\mathcal{E}} + \boldsymbol{\omega}_{\mathcal{E}/\mathcal{N}} = -\frac{R-\rho}{\rho}\omega\hat{\boldsymbol{n}}_3$$

The position vector $\boldsymbol{r}$ of point P relative to the origin O is

$$\boldsymbol{r} = L\hat{\boldsymbol{e}}_L + \rho\hat{\boldsymbol{b}}_r$$

Using the transport theorem in Eq. (1.21), the inertial velocity vector $\dot{\boldsymbol{r}}$ of P is

$$\dot{\boldsymbol{r}} = \frac{{}^{\mathcal{E}}d}{dt}\left(L\hat{\boldsymbol{e}}_L\right) + \boldsymbol{\omega}_{\mathcal{E}/\mathcal{N}} \times L\hat{\boldsymbol{e}}_L + \frac{{}^{\mathcal{B}}d}{dt}\left(\rho\hat{\boldsymbol{b}}_r\right) + \boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}} \times \rho\hat{\boldsymbol{b}}_r$$

Note that since L and ρ are constants for this system, the derivatives within the $\mathcal{E}$ and $\mathcal{B}$ frames are zero since $\hat{\boldsymbol{e}}_L$ is fixed in $\mathcal{E}$ and $\hat{\boldsymbol{b}}_r$ is fixed in $\mathcal{B}$, so

$$\dot{\boldsymbol{r}} = \omega L\hat{\boldsymbol{e}}_\theta + (R-\rho)\,\omega\hat{\boldsymbol{b}}_\phi$$

The inertial acceleration vector $\ddot{\boldsymbol{r}}$ of P is found by taking the derivative of $\dot{\boldsymbol{r}}$ in the $\mathcal{N}$ frame.

$$\ddot{\boldsymbol{r}} = \frac{{}^{\mathcal{E}}d}{dt}\left(\omega L\hat{\boldsymbol{e}}_\theta\right) + \boldsymbol{\omega}_{\mathcal{E}/\mathcal{N}} \times \left(\omega L\hat{\boldsymbol{e}}_\theta\right) + \frac{{}^{\mathcal{B}}d}{dt}\left((R-\rho)\,\omega\hat{\boldsymbol{b}}_\phi\right) + \boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}} \times \left((R-\rho)\,\omega\hat{\boldsymbol{b}}_\phi\right)$$

Since ω is constant, the inertial acceleration is then written as the simple expression

$$\ddot{\boldsymbol{r}} = -\omega^2 L\hat{\boldsymbol{e}}_L - \frac{(R-\rho)^2}{\rho}\omega^2\hat{\boldsymbol{b}}_r$$

To express the inertial acceleration only in unit direction vectors of, for example, the $\mathcal{E}$ frame, we eliminate $\hat{\boldsymbol{b}}_r$ by making use of the identity

$$\hat{\boldsymbol{b}}_r = -\cos\phi\hat{\boldsymbol{e}}_L + \sin\phi\hat{\boldsymbol{e}}_\theta$$

to obtain the final result

$${}^{\mathcal{E}}\ddot{\boldsymbol{r}} = -\left(\omega^2 L - \frac{(R-\rho)^2}{\rho}\omega^2\cos\phi\right)\hat{\boldsymbol{e}}_L - \frac{(R-\rho)^2}{\rho}\omega^2\sin\phi\hat{\boldsymbol{e}}_\theta$$

Although the result in Eq. (1.35) can be quite useful at times, when more than two frames are present it is typically easier to derive the acceleration terms by differentiating the position vector twice as in this example.

Problems

1.1 The particle P moves along a space curve described by the cartesian coordinates

$$x(t) = cos(t)$$
$$y(t) = sin(t)$$
$$z(t) = sin(t)$$

Describe the given motion in terms of cylindrical and spherical coordinates by finding explicit equations for the coordinates.

1.2 The planar point acceleration vector is given in the cartesian coordinates as

$$\ddot{\boldsymbol{r}} = \ddot{x}\hat{\boldsymbol{e}}_1 + \ddot{y}\hat{\boldsymbol{e}}_2$$

Directly transform this vector into polar coordinates r, θ, $\hat{\boldsymbol{e}}_r$ and $\hat{\boldsymbol{e}}_\theta$ by substituting $x = r\cos\theta$, $y = r\sin\theta$. Verify the result in Eq. (1.24) obtained through the transport theorem.

1.3 Let a particle P be free to slide radially in a rotating tube as shown in Figure 1.10. Assume the tube is rotating at a constant angular velocity ω. What is the inertial velocity and acceleration of the particle P? Express your answer as functions of r, θ, $\hat{\boldsymbol{e}}_r$ and $\hat{\boldsymbol{e}}_\theta$.

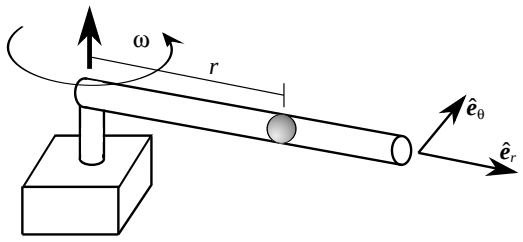


Figure 1.10: Particle in Rotating Tube

1.4 ♣ Let $\mathcal{N} = \{O, \hat{\boldsymbol{n}}_1, \hat{\boldsymbol{n}}_2, \hat{\boldsymbol{n}}_3\}$ be an inertial, non-rotating reference frame with its center in the center of Earth. The Earth-fixed, equatorial coordinate frame $\mathcal{E} = \{O, \hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_2, \hat{\boldsymbol{n}}_3\}$ has the same origin, but the unit direction vectors are fixed in the Earth. The Earth-fixed, topocentric coordinate frame $\mathcal{T} = \{O', \hat{\boldsymbol{u}}, \hat{\boldsymbol{e}}, \hat{\boldsymbol{n}}\}$ tracks a point on Earth as shown in Figure 1.11. Notice the local "geometric" interpretation: $\hat{\boldsymbol{u}}$ = "up", $\hat{\boldsymbol{e}}$ = "east" and $\hat{\boldsymbol{n}}$ = "north". Assuming that a stationary person is at a latitude of $\phi = 40^{\text{ffi}}$ and a longitude of $\lambda = 35^{\text{ffi}}$, what is the inertial velocity and acceleration of the point O'? Express your answer in both $\{\hat{\boldsymbol{n}}\}$ and $\{\hat{\boldsymbol{e}}\}$ components as functions of r, θ, λ, ϕ and derivatives thereof.

1.5 When launching a vehicle into orbit, one typically tries to make use of Earth's rotation when choosing a launch site. From what place on Earth would it be the simplest (i.e. require least additional energy to be added) to launch vehicles into space and how much initial eastward velocity (as seen in an Earth-fixed frame) would a vehicle have there thanks to Earth's rotation?

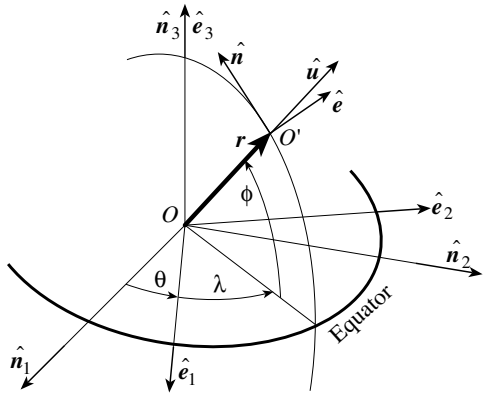


Figure 1.11: Coordinate Frames of a Person on Earth

1.6 ♣ The person in Problem 1.4 has boarded a high-speed train and is traveling due south at a constant 450 km/h as seen in an Earth-fixed reference frame. What is the inertial velocity and acceleration now?

1.7 A constantly rotating disk is mounted on a moving train as shown in Figure 1.12. The train itself is moving with a time varying linear velocity of $\boldsymbol{v}(t)$. Assume the particle P is fixed on the disk, what are its inertial velocity and acceleration? Express your answer with $\{\hat{\boldsymbol{d}}\}$ components as functions of r, ω and $v(t)$.

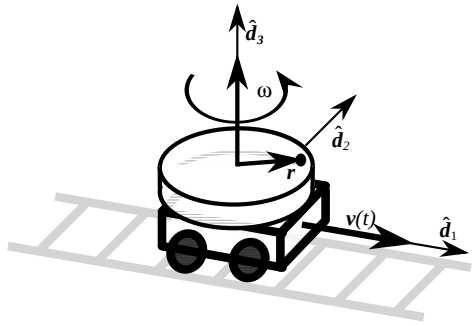


Figure 1.12: Rotating Disk on Train

1.8 Repeat Problem 1.7, but this time assume that the particle P is free to move radially on the disk. Again find the corresponding inertial velocity and acceleration.

1.9 Two rotating disks are are arranged as shown in Figure 1.13. Relative to an inertial reference frame $\mathcal{N}$, Disk $\mathcal{A}$ has a relative angular velocity $\boldsymbol{\omega}_{\mathcal{A}/\mathcal{N}}$ and disk $\mathcal{B}$ has a relative angular velocity $\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}$. Each disk has a particle A or B respectively fixed to its rim. The orientation of the $\mathcal{A}$ frame is given by $\{\hat{\boldsymbol{a}}_r, \hat{\boldsymbol{a}}_t, \hat{\boldsymbol{a}}_3\}$ and the orientation of the $\mathcal{B}$ frame is given by $\{\hat{\boldsymbol{b}}_r, \hat{\boldsymbol{b}}_t, \hat{\boldsymbol{b}}_3\}$.

a) What is the relative inertial velocity $\dot{\boldsymbol{\rho}}$ and acceleration $\ddot{\boldsymbol{\rho}}$ of particle B versus A?

b) As seen from particle A, what is the relative velocity and acceleration of particle B?

It is recommended that this problem be solved in two ways: (1) By using Eq. (1.35) and (2) by differentiation of the position and velocity vector using the transport theorem.

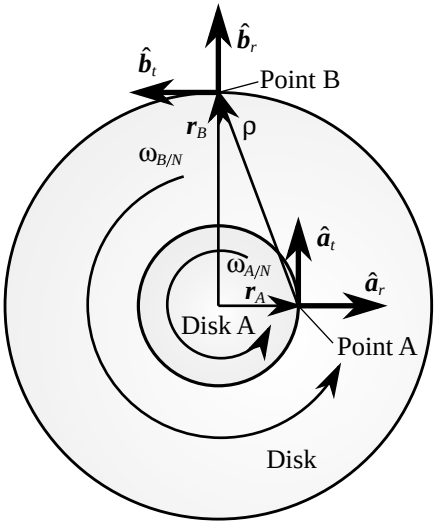


Figure 1.13: Two Rotating Disks

1.10 Consider the overly simplified planetary system shown in Figure 1.14. The Earth is assumed to have a circular orbit of radius R about the sun and is orbiting at a constant rate $\dot{\phi}$. The moon is orbiting Earth also in a circular orbit at a constant radius r at a constant rate $\dot{\theta}$. Assume the sun is inertially fixed in space by the frame $\{\hat{\boldsymbol{n}}_1, \hat{\boldsymbol{n}}_2, \hat{\boldsymbol{n}}_3\}$. Further, a UFO is orbiting the sun at a radius R_2 at fixed rate $\dot{\gamma}$. Let the Earth frame $\mathcal{E}$ be given by the direction vectors $\{\hat{\boldsymbol{e}}_r, \hat{\boldsymbol{e}}_\phi, \hat{\boldsymbol{e}}_3\}$, the moon frame $\mathcal{M}$ by $\{\hat{\boldsymbol{m}}_r, \hat{\boldsymbol{m}}_\theta, \hat{\boldsymbol{m}}_3\}$ and the UFO frame $\mathcal{U}$ by $\{\hat{\boldsymbol{u}}_r, \hat{\boldsymbol{u}}_\gamma, \hat{\boldsymbol{u}}_3\}$.

a) Find the inertial velocity and acceleration of the moon relative to the sun.

b) Find the position vector of the moon relative to the UFO.

c) Find the angular velocity vectors $\boldsymbol{\omega}_{\mathcal{E}/\mathcal{U}}$ and $\boldsymbol{\omega}_{\mathcal{M}/\mathcal{U}}$.

d) What are the velocity and acceleration vectors of the moon *as seen by the UFO frame*?

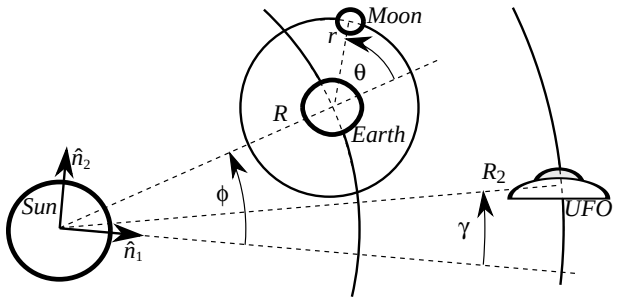


Figure 1.14: Planar Planetary System

1.11 A disk of constant radius r is attached to a telescoping rod which is extending at a constant rate as shown in Figure 1.15. Both the disk and the rod are rotating at a constant rate. Find the inertial velocity and acceleration of point P at the rim of the disk.

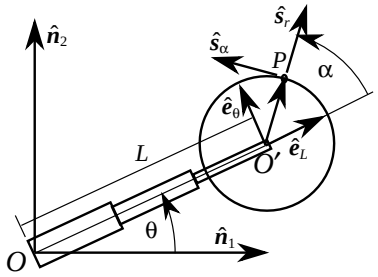


Figure 1.15: Rotating Disk Attached to Telescoping Rod

1.12 A disk is rolling at a constant rate $\dot{\theta}$ on a moving conveyor belt as shown in Figure 1.16. The conveyor belt speed v is constant. Find the inertial velocity and acceleration of Point P.

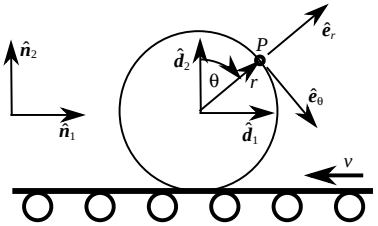


Figure 1.16: Disk Rolling on a Conveyor Belt

1.13 A vertical disk of radius r is attached to a horizontal shaft of length R as shown in Figure 1.17. The shaft is rotating at a time varying rate $\dot{\phi}$. A fixed point P is on the rim of the disk, while a missile is flying overhead at a fixed height h with the trajectory $\boldsymbol{r}_m = h\hat{\boldsymbol{n}}_3 - t\hat{\boldsymbol{n}}_2$.

a) Find the inertial velocity and acceleration of point P.

b) What is the velocity and acceleration of point P *as seen* by the missile.

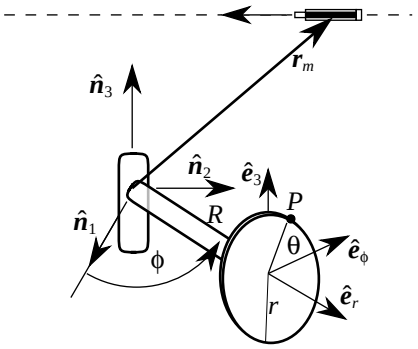


Figure 1.17: Grinding Disk

1.14 Two disks are rotating at constant rates $\dot{\theta}$ and $\dot{\phi}$ a fixed distance L apart as shown in Figure 1.18. The radius of the left disk is r and the radius of the right disk is R.

a) What is the inertial velocity and acceleration of point B on the right disk?

b) As seen from Point A on the left disk, what is the relative velocity and acceleration of Point B?

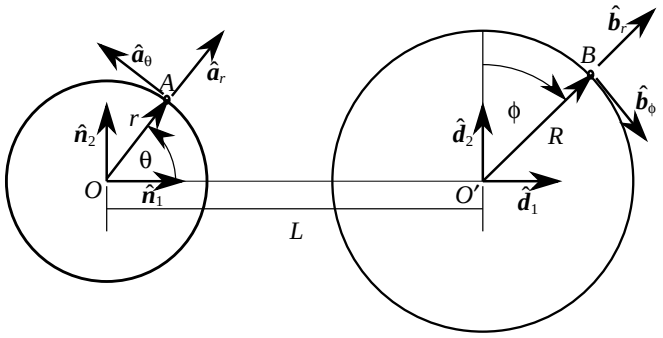


Figure 1.18: Two Rotating Disks

1.15 A person A is descending in an elevator at a constant velocity v. A second person B is riding a a big wheel of radius R whose center is a distance L away from the elevator as shown in Figure 1.19. What is the relative velocity and acceleration of person B *as seen* from person A?

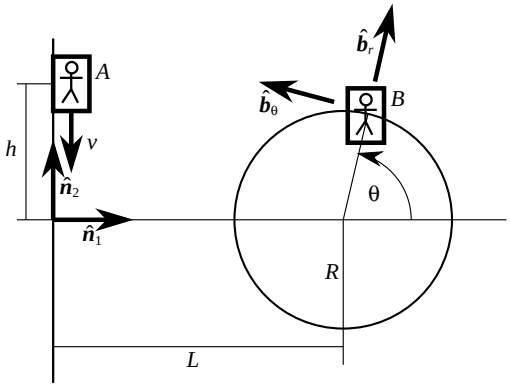


Figure 1.19: Person Riding Large Wheel

Bibliography

[1] Likins, P. W., *Elements of Engineering Mechanics*, McGraw-Hill, New York, 1973.

[2] Greenwood, D. T., *Principles of Dynamics*, Prentice-Hall, Inc, Englewood Cliffs, New Jersey, 2nd ed., 1988.

Chapter Two

Newtonian Mechanics

The previous chapter on Particle Kinematics dealt with vector methods for *describing* a motion. Now we would like to be able to establish complete motion models which permit us to *solve* for the motion once the system forces and torques are given. Mass distribution and point of application of forces of a dynamical system clearly affect the resulting motion and must be taken into account. The motions are found by solving the system equations of motion which form the cause/effect model between the forces acting on the system and the resulting translational, rotational and deformational accelerations.

In this chapter, we will first consider the dynamics of a single particle and then that of a system of particles. An example of a system of particles would be the solar system with the various planets within it idealized as particles. The particle mechanics results will then be generalized to derive formulations for the dynamics of continuous systems such as vibrating beams or some generally deformable collection of matter (such as a bowl of Jello) where the system shape may be time varying.

2.1 Newton's Laws

The following laws of nature were discovered by Sir Isaac Newton over 200 years ago in England. Later in the early 20th century Albert Einstein theorized that these basic laws are only a low-speed approximation in his papers about special relativity. However, *relativistic effects* only become significant when the velocity of a particle or body approaches that of the speed of light. In this discussion we will assume that all systems studied are moving much slower than the speed of light and we will therefore neglect relativistic effects. The following three laws are commonly known as *Newton' s laws of motion.*[1–3]

Newton's First Law: *Unless acted upon by a force, a particle will maintain a straight line motion with constant inertial velocity.*

Newton's First Law is the most easily overlooked Law because it is a special

case of the second law. It simply states that unless something pushes against the particle, it will keep on moving in the same direction with constant velocity.

Newton's Second Law: *Let the vector $\boldsymbol{F}$ be the sum of all forces acting on a particle having a mass m with the inertial position vector $\boldsymbol{r}$. Assume that $\mathcal{N}$ is an inertial reference frame, then*

$$\boldsymbol{F} = \frac{{}^{\mathcal{N}}d}{dt}\left(m\dot{\boldsymbol{r}}\right) \tag{2.1}$$

Or in words, the force acting on m is equal to the inertial time rate of change of the particle linear momentum $\boldsymbol{p} = m\dot{\boldsymbol{r}}$. If the mass m is constant then this results simplifies to the well known result

$$\boldsymbol{F} = m\ddot{\boldsymbol{r}} \tag{2.2}$$

We observe that if units are not chosen consistent with Eqs. (2.1) and (2.2), Newton's second law requires an additional proportionality factor. Note that all derivatives taken in Newton's Second Law must be inertial derivatives. Since it is typically necessary to also describe a position vector in a non-inertial coordinate frame, the importance of proper kinematics skills becomes apparent. Without correctly formulated kinematics, the dynamical system description will be incorrect from the start. We mention that a large fraction of errors made in practice have their origin in kinematics errors formulating $\ddot{\boldsymbol{r}}$ and similar vector derivatives.

Newton's Third Law: *If mass m_1 is exerting a force $\boldsymbol{F}_{21}$ on mass m_2, then the force $\boldsymbol{F}_{12}$ experienced by m_1 due to interaction with m_2 will be*

$$\boldsymbol{F}_{12} = -\boldsymbol{F}_{21} \tag{2.3}$$

This conforms to our intuitive experience. Anytime one pushes against an object, the reaction force from the object to our hand is an equal force. Be sure to keep that in mind when contemplating punching a solid wall, or jumping from a canoe.

In order to write down Newton's laws, it is important to make use of force and moment sketches known as Free Body Diagrams (FBDs). In essence, FBDs are used to specify and determine the force vector $\boldsymbol{F}$ in Eq. (2.2). Figure 2.1 is an example of a FBD. There are several conventions for free body diagrams, we adopt the following rule. The FBD should show all forces and moments acting on the system. We exclude from our FBDs acceleration vectors and so-called "inertia forces" which are subsets of the $m\ddot{\boldsymbol{r}}$ terms in Eq. (2.2) that may arise in rotating coordinate systems.

Sir Isaac Newton is probably best known for the development of calculus and the laws of gravity which by popular account were initiated when an apple fell on his head while he was sitting under a tree. However, his laws of motion form the foundation of all modern sciences and engineering.

Newton's Law of Universal Gravitation: *Let the vector* $\boldsymbol{r}_{12} = \boldsymbol{r}_2 - \boldsymbol{r}_1$ *describe the position of mass* m_2 *relative to mass* m_1 *as shown in Figure 2.1. Then the mutually attractive gravitational force between the objects will be*

$$\boldsymbol{F}_{12} = -\boldsymbol{F}_{21} = \frac{Gm_1m_2}{|\boldsymbol{r}_{12}|^2}\frac{\boldsymbol{r}_{12}}{|\boldsymbol{r}_{12}|} \tag{2.4}$$

where $G \cong 6.6732 \cdot 10^{-11}\frac{m^3}{s^2 kg}$ *is the universal gravity constant.*

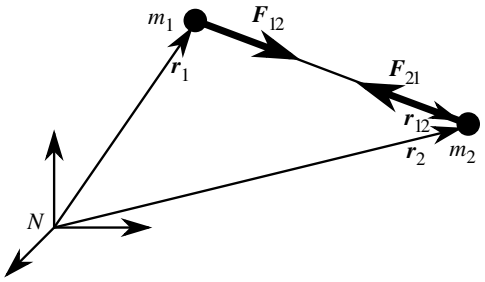


Figure 2.1: Newton's Law of Universal Gravitation

For example, this law of universal gravitation allows one to model accurately the attractive forces between spacecraft and planets. Note however, that since the universal gravity constant G is relatively small, the gravitational attraction between two everyday objects such as a house and a car is very small and typically ignored. Even Mount Everest makes a barely measurable perturbation in the Earth's total gravitational attraction on objects in the immediate vicinity of Mount Everest.

One important aspect of the law of universal gravitation is that the gravity force is conservative and can be calculated from a gravity field potential energy function. A general potential energy function $V(\boldsymbol{r})$ is a scalar function which depends on the system position vector $\boldsymbol{r}$. The potential function measures how much work has to be done to the system to move an object from rest a reference position $\boldsymbol{r}_0$ to rest at position $\boldsymbol{r}$. A conservative force is defined as a force derivable by taking the gradient of a corresponding potential energy function $V(\boldsymbol{r})$ as

$$\boldsymbol{F}(\boldsymbol{r}) = -\nabla V(\boldsymbol{r}) \tag{2.5}$$

Given V, we can derive $\boldsymbol{F}$ from the gradient operator as in Eq. (2.5). Given $\boldsymbol{F}$, we can derive V by integration. Note that conservative forces only depend on the position vector $\boldsymbol{r}$ and not the velocity vector $\dot{\boldsymbol{r}}$ or time t. For example, the classical viscous drag force $\boldsymbol{F} = -c\dot{\boldsymbol{r}}$ would not be a conservative force.

The gravity potential energy function V_G experienced by the masses m_1 and m_2 is[1, 3]

$$V_G(\boldsymbol{r}_{12}) = -\frac{Gm_1m_2}{|\boldsymbol{r}_{12}|} = -\frac{Gm_1m_2}{r_{12}} \tag{2.6}$$

$V_G(\boldsymbol{r}_{12})$ is energy required to separate the two masses from the current distance of $|\boldsymbol{r}_{12}|$ to an infinite separation. We will subsequentially consider (in section 2.2.3) the relationship of potential energy and work in more detail. Let's describe the $\boldsymbol{r}_{12}$ vector through cartesian coordinates as

$$\boldsymbol{r}_{12} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \tag{2.7}$$

The magnitude of $\boldsymbol{r}_{12}$ is defined as

$$|\boldsymbol{r}_{12}| = \sqrt{x_1^2 + x_2^2 + x_3^2} \tag{2.8}$$

and the partial derivatives of $|\boldsymbol{r}_{12}|$ with respect to the cartesian coordinates x_i are given by

$$\frac{\partial |\boldsymbol{r}_{12}|}{\partial x_i} = \frac{x_i}{|\boldsymbol{r}_{12}|} \tag{2.9}$$

The gradient of the potential field V_G is given by

$$\frac{\partial V_G}{\partial x_i} = \frac{Gm_1m_2}{|\boldsymbol{r}_{12}|^2} \frac{\partial |\boldsymbol{r}_{12}|}{\partial x_i} = \frac{Gm_1m_2}{|\boldsymbol{r}_{12}|^2} \frac{x_i}{|\boldsymbol{r}_{12}|} \tag{2.10}$$

The gravitational force $\boldsymbol{F}_{21}$ the mass m_2 experiences due to the mass m_1 at the relative position $\boldsymbol{r}_{12}$ is given by

$$\boldsymbol{F}_{21} = -\nabla V_G = -\frac{Gm_1m_2}{|\boldsymbol{r}_{12}|^2} \frac{1}{|\boldsymbol{r}_{12}|} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = -\frac{Gm_1m_2}{|\boldsymbol{r}_{12}|^3} \boldsymbol{r}_{12} \tag{2.11}$$

Another example of a conservative force is the force exerted by a spring. Let the spring have a spring constant k and a linear deflection x. Then its potential function V_S is given by

$$V_S(x) = \frac{1}{2} kx^2 \tag{2.12}$$

The current potential energy indicates how much work was performed to stretch the spring from a zero reference deflection state to the deflection x. The force exerted by the spring on a mass m is given by the famous Hook's Law.

$$F = -\nabla V_G = -kx \tag{2.13}$$

Example 2.1: Let us find a first order approximation of the gravity potential function in Eq. (2.6) that a body with m would experience near the Earth's surface. Assume a spherical Earth with radius R_e and mass m_e. The radial distance r of the body to the center of Earth is written as

$$r = R_e + h$$

where h is the height above the Earth's surface. The gravity potential experienced by the body m due to Earth is

$$V(r) = -\frac{Gm_e m}{r}$$

The function $V(r)$ can be approximated about the distance R_e through the Taylor series expansion

$$V(r) = V(R_e) + \frac{1}{1!}\frac{\partial V}{\partial r}\bigg|_{R_e} h + \frac{1}{2!}\frac{\partial^2 V}{\partial r^2}\bigg|_{R_e} h^2 + \cdots$$

The local gravity potential V_{local} uses R_e as its reference potential and can approximated by

$$V_{local}(h) = V(r) - V(R_e) \simeq \frac{\partial V}{\partial r}\bigg|_{R_e} h + \mathcal{O}(h^2)$$

After carrying out the partial derivative the local gravity potential function for the special case of a *constant* gravity field is found to be

$$V_{local}(h) = \frac{Gm_e}{R_e^2} mh = mgh$$

where $g = Gm_e/R_e^2$ is the local gravitational acceleration.

2.2 Single Particle Dynamics

The equation of motion for a single particle is given by Newton's second law in Eq. (2.2) where it is assumed that the particle mass m is constant and $\ddot{\boldsymbol{r}}$ is the second inertial derivative of the position vector $\boldsymbol{r}$. The following two sections treat two cases of this simple dynamical system. In the first case the force being applied to the mass is assumed to be constant and in the second case it is assumed to be time varying.

2.2.1 Constant Force

If the force $\boldsymbol{F}$ being applied to the mass m is a constant vector, then the equations of motion

$$m\ddot{\boldsymbol{r}} = \boldsymbol{F} = \text{constant} \tag{2.14}$$

can be solved for the time varying position vector $\boldsymbol{r}(t)$. Eq. (2.14) can be solved for the inertial acceleration vector $\ddot{\boldsymbol{r}}$ as

$$\ddot{\boldsymbol{r}}(t) = \frac{\boldsymbol{F}}{m} \tag{2.15}$$

After integrating this equation once from an initial time t_0 to an arbitrary time t we obtain the following velocity formulation for mass m.

$$\dot{\boldsymbol{r}}(t) = \dot{\boldsymbol{r}}(t_0) + \frac{\boldsymbol{F}}{m}(t - t_0) \tag{2.16}$$

After integrating the velocity formulation an expression for the time varying position vector $\boldsymbol{r}(t)$ of mass m is found.

$$\boldsymbol{r}(t) = \boldsymbol{r}(t_0) + \dot{\boldsymbol{r}}(t_0)(t - t_0) + \frac{\boldsymbol{F}}{2m}(t - t_0)^2 \tag{2.17}$$

Note that Eqs. (2.15) through (2.17) are actually each three sets of equations since $\boldsymbol{r} = (x_1, x_2, x_3)^T$ and $\boldsymbol{F} = (F_1, F_2, F_3)^T$ are each three-dimensional vectors. Given an initial velocity vector $\dot{\boldsymbol{r}}(t_0)$, the time required to reach a final velocity under constant driving force $\boldsymbol{F}$ can be solved from Eq. (2.16).

$$(t - t_0) = (\dot{x}_i(t) - \dot{x}_i(t_0))\frac{m}{F_i} \tag{2.18}$$

Given an initial position vector $\boldsymbol{r}(t_0)$, the time required to reach a final position vector under constant driving force is found by solving the quadratic equation in Eq. (2.17) for the time t.

$$t - t_0 = \frac{m}{F_i}\left(\pm\sqrt{\dot{x}_i^2(t_0) + \frac{2F_i}{m}(x_i(t) - x_i(t_0))} - \dot{x}_i(t_0)\right) \tag{2.19}$$

Given an initial position and velocity vector and a final position vector, the corresponding final velocity vector is found by substituting Eq. (2.18) into Eq. (2.17) and solving for $\dot{\boldsymbol{r}}(t)$.

$$\dot{x}_i^2(t) = \dot{x}_i^2(t_0) + 2\frac{F_i}{m}(x_i(t) - x_i(t_0)) \tag{2.20}$$

Example 2.2: The trajectory of a mass m is studied as it travels in a vertical plane under the influence of a constant gravitational force F. Determine an equation that relates an arbitrary target location (x_1, x_2) to the corresponding launch velocity v_0 and flight path angle γ_0. As shown in Figure 2.2, the mass is at the coordinate center at time zero with a speed of v_0 and a elevation angle of γ_0. The cartesian components of the initial position and velocity vectors are therefore given by

$$\boldsymbol{r}(t_0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \dot{\boldsymbol{r}}(t_0) = v_0 \begin{pmatrix} \cos\gamma_0 \\ \sin\gamma_0 \end{pmatrix}$$

Since the gravitational force F only acts along the vertical direction, the equations of motion are given as

$$\ddot{\boldsymbol{r}}(t) = \frac{1}{m}\begin{pmatrix} 0 \\ -F \end{pmatrix} = \begin{pmatrix} 0 \\ -g \end{pmatrix}$$

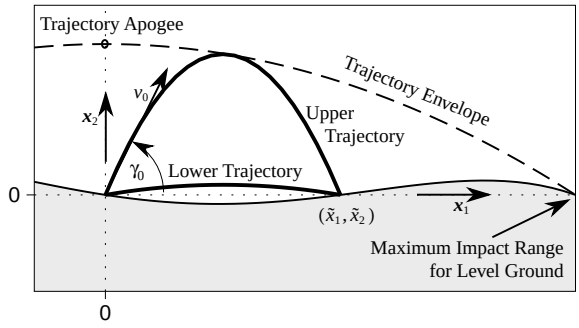


Figure 2.2: Ballistic Trajectories under Constant Gravity Force

where $g = F/m$ is the local constant gravitational acceleration. Using Eq. (2.16) the velocity vector $\dot{\boldsymbol{r}}(t)$ is

$$\dot{\boldsymbol{r}}(t) = v_0 \begin{pmatrix} \cos\gamma_0 \\ \sin\gamma_0 \end{pmatrix} - \begin{pmatrix} 0 \\ gt \end{pmatrix}$$

The position vector $\boldsymbol{r}(t)$ is found through Eq. (2.17).

$$\boldsymbol{r}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = v_0 t \begin{pmatrix} \cos\gamma_0 \\ \sin\gamma_0 \end{pmatrix} - \begin{pmatrix} 0 \\ gt^2 2 \end{pmatrix}$$

By solving the $x_1(t)$ equation for the time t and substituting it into the $x_2(t)$ equation, one obtains the parabola expression relating x_2 to x_1 (the equation of the path or trajectory):

$$x_2 = x_1 \tan\gamma_0 - \frac{g \sec^2\gamma_0}{2v_0^2} x_1^2$$

An interesting question now arises. Given an initial speed v_0, what would the initial elevation angle γ_0 have to be to make the mass m hit a target at coordinates $(\tilde{x}_1, \tilde{x}_2)$? To answer this we rewrite the above expression relating x_1 and x_2 making use of the trig identity $\sec^2\gamma_0 = 1 + \tan^2\gamma_0$.

$$\tan^2\gamma_0 - \frac{2v_0^2}{g\tilde{x}_1}\tan\gamma_0 + \frac{2v_0^2\tilde{x}_2}{g\tilde{x}_1^2} + 1 = 0$$

This quadratic equation can be solved explicitly for $\tan\gamma_0$.

$$(\tan\gamma_0)_{1/2} = \frac{v_0^2}{g\tilde{x}_1} \pm \frac{v_0}{g\tilde{x}_1}\sqrt{v_0^2 - 2g\tilde{x}_2 - \frac{g^2\tilde{x}_1^2}{v_0^2}}$$

If the point $(\tilde{x}_1, \tilde{x}_2)$ is within the range limit, then this formula will return two real answers. One corresponds to a lower trajectory and the other to a higher trajectory as illustrated in Figure 2.2. If the point $(\tilde{x}_1, \tilde{x}_2)$ is on the range limit, then the formula will return a double root. If the real point $(\tilde{x}_1, \tilde{x}_2)$ is outside the range limit, then two complex variables will be returned, indicating the reasonable truth that no real solutions exist.

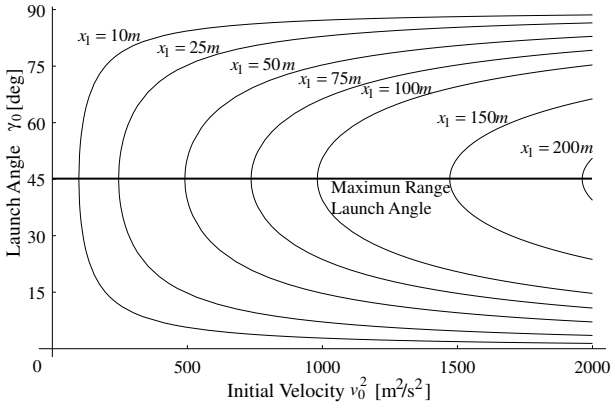


Figure 2.3: Ballistic Trajectories under Constant Gravity Force

To find the envelope of all possible trajectories the case where only double roots exist is examined. Setting the square root term to zero, the following parabola is found.

$$x_2 = \frac{v_0^2}{2g} - \frac{g}{2v_0^2}x_1^2$$

Any targets that are accessible with the given v_0 must lie within this parabola. The trajectory envelope parabola is shown as a dashed line in Figure 2.2. As can be verified, the special case where $\tilde{x}_1 = 0$ gives $\tilde{x}_2 = v_0^2/2g$. You can readily show that this is the apogee of a vertically launched projectile with launch velocity v_0. Another special case where is where $\tilde{x}_2 = 0$, which provides the maximum impact range $x_1 = v_0^2/g$ if the surface is flat. Figure 2.3 compares the various launch angles required to hit a target a distance x_1 away with a given initial velocity v_0^2. For this constant gravity field case, the maximum range launch angle is always 45 degrees. Later on this problem is revisited in celestial mechanics where the inverse square gravity field case is considered.

2.2.2 Time-Varying Force

When the force $\boldsymbol{F}$ acting on the mass m is time varying, then there are typically no closed form solutions for the velocity and position vectors. The equations of motion are given as

$$\ddot{\boldsymbol{r}} = \frac{1}{m}\boldsymbol{F}(t) \tag{2.21}$$

Upon integrating Eq. (2.21) from t_0 to t the velocity vector $\dot{\boldsymbol{r}}(t)$ at time t is given as

$$\dot{\boldsymbol{r}}(t) = \dot{\boldsymbol{r}}(t_0) + \frac{1}{m}\int_{t_0}^{t} \boldsymbol{F}(\tau)d\tau \tag{2.22}$$

The position vector $\boldsymbol{r}(t)$ is obtain by integrating the velocity vector.

$$\boldsymbol{r}(t) = \boldsymbol{r}(t_0) + \dot{\boldsymbol{r}}(t_0)\,(t - t_0) + \frac{1}{m}\int_{t_0}^{t}\int_{t_0}^{\tau_2} \boldsymbol{F}(\tau_1)d\tau_1 d\tau_2 \tag{2.23}$$

Finding the time required to accelerate from one velocity to another or to travel from one position to another under the influence of $\boldsymbol{F}(t)$ cannot be found generically as for the case of constant $\boldsymbol{F}$. These results would have to be found explicitly for a given problem statement or through a numerical method if no closed form solution exists.

Example 2.3: Let the mass m be restricted to travel only in one dimension. It is attached to the coordinate frame origin through a linear spring with spring constant k. The force acting on mass m is then given through Hook's Law as

$$F = -kx$$

and the equations of motion are then given through Newton's second law in Eq. (2.21) as

$$\ddot{x} = \frac{1}{m}\,(-kx)$$

This can be rewritten in the form of the standard unforced oscillator differential equation.

$$m\ddot{x} + kx = 0$$

The oscillator problem is known to have a solution of the type

$$x(t) = A\cos\omega t + B\sin\omega t$$

Where the constants A, B and ω are yet to be determined. The velocity and acceleration expressions are then given as

$$\dot{x}(t) = -A\omega\sin\omega t + B\omega\cos\omega t$$
$$\ddot{x}(t) = -A\omega^2\cos\omega t - B\omega^2\sin\omega t = -\omega^2 x(t)$$

Substituting the expression for $\ddot{x}(t)$ into the equation of motion the following expression is obtained

$$\left(-m\omega^2 + k\right)x = 0$$

which must hold for any position x. Therefore the *natural frequency* ω is given by[4]

$$\omega = \sqrt{\frac{k}{m}}$$

The constants A and B would be found through enforcing the solution to satisfy the initial conditions $x(t_0) = A$ and $\dot{x}(t_0) = \omega B$.

2.2.3 Kinetic Energy

The kinetic energy T of a particle of mass m is given by

$$T = \frac{1}{2} m \dot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}} \tag{2.24}$$

To find the work done on the particle we investigate the time derivative of the kinetic energy T.

$$\frac{dT}{dt} = m \ddot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}} \tag{2.25}$$

After using Eq. (2.14) the kinetic energy rate or *power* is given as

$$\frac{dT}{dt} = \boldsymbol{F} \cdot \dot{\boldsymbol{r}} \tag{2.26}$$

If the force $\boldsymbol{F}$ is conservative it can be expressed as the negative gradient of a potential function V.

$$\frac{dT}{dt} = -\frac{\partial V}{\partial \boldsymbol{r}} \cdot \dot{\boldsymbol{r}} \tag{2.27}$$

Noting that $\frac{\partial V}{\partial \boldsymbol{r}} \dot{\boldsymbol{r}} = \frac{dV}{dt}$ Eq. (2.27) can be written as

$$\frac{dT}{dt} + \frac{dV}{dt} = 0 \tag{2.28}$$

Therefore the total system energy $E = T + V$ is conserved. For conservative systems it is often convenient to obtain an expression relating coordinates and their time derivatives using the system energy. This avoids having to perform difficult integrations of the acceleration expressions to obtain the same relationship.

Let W be the work performed between times t_1 and t_2. Upon integrating Eq. (2.26) from time t_1 to t_2 the following *work/energy equation* is obtained.

$$T(t_2) - T(t_1) = \int_{t_1}^{t_2} \boldsymbol{F} \cdot \dot{\boldsymbol{r}} dt = \int_{r(t_1)}^{r(t_2)} \boldsymbol{F} \cdot d\boldsymbol{r} \equiv W \tag{2.29}$$

Example 2.4: A mass m of 10 kg has an initial kinetic energy of 40 Joules (1 Joule = 1 J = 1 kg m^2/s^2 = 1 Nm). A constant force $F = 4$ N is acting on this mass from the initial position $r(t_0) = 0$ m to the final position at $r(t_f = 10$ m. What is the work done on the mass and what is the final velocity at t_f?

Using Eq. (2.29), the work W done to the mass m is

$$W = \int_{r(t_1)}^{r(t_f)} F \cdot dr = \int_{0m}^{10m} 4N \cdot dr = 40Nm = 40J$$

The energy at t_f is given by

$$T(t_f) = T(t_0) + W = 40J + 40J = 80J$$

Using Eq. (2.24) the final velocity $\dot{r}(t_f)$ is found to be

$$\dot{r}(t_f) = \sqrt{\frac{2T(t_f)}{m}} = 4m/s$$

2.2.4 Linear Momentum

The linear momentum vector $\boldsymbol{p}$ of a particle is defined as

$$\boldsymbol{p} = m\dot{\boldsymbol{r}} \tag{2.30}$$

The momentum measure provides a sense of how difficult it will be to change a motion of a particle. Assume a locomotive has a large mass m and a very small inertial velocity $\dot{\boldsymbol{r}}$. Despite the slow motion, it makes intuitive sense that it would be very difficult to stop the motion of this large object. The linear momentum $\boldsymbol{p}$ of the locomotive is large due to the large mass. Similarly, consider a bullet with a small mass and a very high inertial velocity. Again, it makes intuitive sense that it would be difficult to deflect the motion of the bullet once it has been fired. In this case the linear momentum of the bullet is large not because of its mass, but because of its very large inertial velocity.

Using the linear momentum definition, we are able to rewrite Newton's Second Law in Eq. (2.1) in terms of $\boldsymbol{p}$ as

$$\boldsymbol{F} = \frac{{}^{\mathcal{N}}d}{dt}(m\dot{\boldsymbol{r}}) = \frac{{}^{\mathcal{N}}d}{dt}(\boldsymbol{p}) \tag{2.31}$$

Thus, the force acting on a particle can be defined as the inertial time rate of change of the linear momentum of the particle. If no force is acting on the particle, then $\dot{\boldsymbol{p}}$ is zero and the linear momentum is constant. For the single particle system, this is a rather trivial result. However, using the analogous arguments on a multi-particle system will yield some very powerful conclusions.

2.2.5 Angular Momentum

Let P be an arbitrary point in space with the inertial position vector $\boldsymbol{r}_P$ and the mass m have an inertial position vector $\boldsymbol{r}$. The relative position of m to point P is given through

$$\boldsymbol{\sigma} = \boldsymbol{r} - \boldsymbol{r}_P \tag{2.32}$$

The angular momentum vector $\boldsymbol{H}_P$ of the particle m about point P is defined as

$$\boldsymbol{H}_P = \boldsymbol{\sigma} \times m\dot{\boldsymbol{\sigma}} \tag{2.33}$$

Taking the time derivative of $\boldsymbol{H}_P$ we find

$$\dot{\boldsymbol{H}}_P = \dot{\boldsymbol{\sigma}} \times m\dot{\boldsymbol{\sigma}} + \boldsymbol{\sigma} \times m\ddot{\boldsymbol{\sigma}} \tag{2.34}$$

After noting that $\ddot{\boldsymbol{\sigma}} = \ddot{\boldsymbol{r}} - \ddot{\boldsymbol{r}}_P$ and that a vector cross product with itself is zero, the vector $\dot{\boldsymbol{H}}_P$ is

$$\dot{\boldsymbol{H}}_P = \boldsymbol{\sigma} \times m\ddot{\boldsymbol{r}} - \boldsymbol{\sigma} \times m\ddot{\boldsymbol{r}}_P \tag{2.35}$$

Using Eq. (2.14) this is rewritten as

$$\dot{\boldsymbol{H}}_P = \boldsymbol{\sigma} \times \boldsymbol{F} + m\ddot{\boldsymbol{r}}_P \times \boldsymbol{\sigma} \tag{2.36}$$

Note that the term $\boldsymbol{\sigma} \times \boldsymbol{F}$ is the moment (or torque) vector $\boldsymbol{L}_P$ due to force $\boldsymbol{F}$ about point P. The angular momentum time derivative can then be written in its most general form

$$\dot{\boldsymbol{H}}_P = \boldsymbol{L}_P + m\ddot{\boldsymbol{r}}_P \times \boldsymbol{\sigma} \tag{2.37}$$

Note that if the reference point P is inertially non-accelerating or $\boldsymbol{r} = \boldsymbol{r}_P$, then Eq. (2.37) is reduced to the famous *Euler's equation.*[1, 2]

$$\dot{\boldsymbol{H}}_P = \boldsymbol{L}_P \tag{2.38}$$

Example 2.5: A weightless cylinder of radius R with a mass m embedded in it is rolling down a slope of angle α without slip under the influence of a constant gravity field as shown in Figure 2.4. The mass is offset from the cylinder center by a distance l. Let $\mathcal{N} : \{O, \hat{\boldsymbol{n}}_1, \hat{\boldsymbol{n}}_2, \hat{\boldsymbol{n}}_3\}$ be an inertial frame and $\mathcal{E} : \{O', \hat{\boldsymbol{e}}_r, \hat{\boldsymbol{e}}_\theta, \hat{\boldsymbol{e}}_3\}$ be a rotating frame tracking the point mass within the cylinder. Note that $\hat{\boldsymbol{e}}_3 = -\hat{\boldsymbol{n}}_3$.

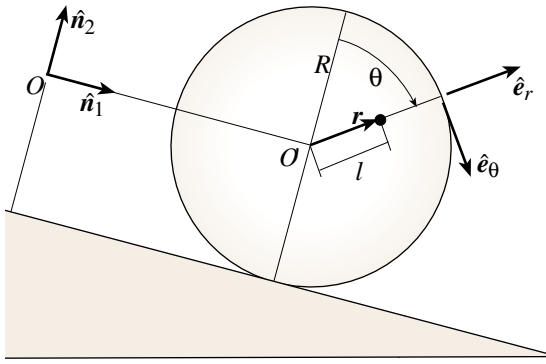


Figure 2.4: Cylinder with Offset Mass Rolling Down a Slope

The angular velocity vector between the $\mathcal{E}$ and the $\mathcal{N}$ frame is

$$\boldsymbol{\omega}_{\mathcal{E}/\mathcal{N}} = \dot{\theta}\hat{\boldsymbol{e}}_3 = -\theta\hat{\boldsymbol{n}}_3$$

Because of the no slip condition, the distance d that the center of the cylinder travels downhill is related to rotation angle θ through

$$d = R\theta$$

The position vector $\boldsymbol{r}$ of the point mass relative to O is written as

$$\boldsymbol{r} = d\hat{\boldsymbol{n}}_1 + l\hat{\boldsymbol{e}}_r = R\theta\hat{\boldsymbol{n}}_1 + l\hat{\boldsymbol{e}}_r$$

Using the transport theorem, the inertial velocity and acceleration vectors are found to be

$$\dot{\boldsymbol{r}} = R\dot{\theta}\hat{\boldsymbol{n}}_1 + l\dot{\theta}\hat{\boldsymbol{e}}_\theta$$
$$\ddot{\boldsymbol{r}} = R\ddot{\theta}\hat{\boldsymbol{n}}_1 + l\ddot{\theta}\hat{\boldsymbol{e}}_\theta - l\dot{\theta}^2\hat{\boldsymbol{e}}_r$$

The $\mathcal{E}$ frame unit vectors are expressed in terms of $\mathcal{N}$ frame components as

$$\hat{\boldsymbol{e}}_r = \sin\theta\hat{\boldsymbol{n}}_1 + \cos\theta\hat{\boldsymbol{n}}_2$$
$$\hat{\boldsymbol{e}}_\theta = \cos\theta\hat{\boldsymbol{n}}_1 - \sin\theta\hat{\boldsymbol{n}}_2$$

The acceleration vector of the point mass m is then expressed in the $\mathcal{N}$ frame as

$${}^{\mathcal{N}}\ddot{\boldsymbol{r}} = \left(R\ddot{\theta} + l\ddot{\theta}\cos\theta - l\dot{\theta}^2\sin\theta\right)\hat{\boldsymbol{n}}_1 - \left(l\ddot{\theta}\sin\theta + l\dot{\theta}^2\cos\theta\right)\hat{\boldsymbol{n}}_2$$

The forces acting on the rolling cylinder are the gravitational force $\boldsymbol{F}_g$,

$$\boldsymbol{F}_g = mg\left(\sin\alpha\hat{\boldsymbol{n}}_1 - \cos\alpha\hat{\boldsymbol{n}}_2\right)$$

the normal force $\boldsymbol{N}$ pushing perpendicular from the surface,

$$\boldsymbol{N} = N\hat{\boldsymbol{n}}_2$$

and the frictional force $\boldsymbol{F}_f$ which is keeping the cylinder from slipping.

$$\boldsymbol{F}_f = -F_f\hat{\boldsymbol{n}}_1$$

Newton's second law states that

$$m\ddot{\boldsymbol{r}} = \boldsymbol{F}_g + \boldsymbol{N} + \boldsymbol{F}_f$$

After substituting ${}^{\mathcal{N}}\ddot{\boldsymbol{r}}$ and the expressions for the forces into the above equation and equating the $\mathcal{N}$ frame components, the following two relationships are found.

$$m\left(R\ddot{\theta} + l\cos\theta\ddot{\theta} - l\dot{\theta}^2\sin\theta\right) = mg\sin\alpha - F_f$$
$$-m\left(l\sin\theta\ddot{\theta} + l\dot{\theta}^2\cos\theta\right) = -mg\cos\alpha + N$$

Once an expression for $\ddot{\theta}$ is found, the second equation could be used to solve for the time varying normal force component N. To solve the first equation for the angular acceleration, an expression for the frictional force component F_f must be found. To do so we examine the angular momentum vector of

the point mass about the $\mathcal{E}$ frame origin O'. The relative position vector $\boldsymbol{\sigma}$ of the point mass to O' and its inertial derivative are given by

$$\boldsymbol{\sigma} = l\hat{\boldsymbol{e}}_r \qquad \dot{\boldsymbol{\sigma}} = l\dot{\theta}\hat{\boldsymbol{e}}_\theta$$

The angular momentum vector $\boldsymbol{H}_{O'}$ can then be written as

$$\boldsymbol{H}_{O'} = \boldsymbol{\sigma} \times m\dot{\boldsymbol{\sigma}} = -ml^2\dot{\theta}\hat{\boldsymbol{n}}_3$$

and its inertial derivative is given by

$$\dot{\boldsymbol{H}}_{O'} = -ml^2\ddot{\theta}\hat{\boldsymbol{n}}_3$$

The torque $\boldsymbol{L}_{O'}$ about point O' is written as

$$\begin{aligned}\boldsymbol{L}_{O'} &= \boldsymbol{\sigma} \times \boldsymbol{F}_g - R\hat{\boldsymbol{n}}_2 \times (\boldsymbol{F}_f + \boldsymbol{N}) \\ &= -mgl\sin(\theta+\alpha)\,\hat{\boldsymbol{n}}_3 - RF_f\hat{\boldsymbol{n}}_3\end{aligned}$$

The inertial position vector $\boldsymbol{r}_{O'}$ of point O' and its second inertial derivative are given by

$$\boldsymbol{r}_{O'} = d\hat{\boldsymbol{n}}_1 = R\theta\hat{\boldsymbol{n}}_1 \qquad \ddot{\boldsymbol{r}}_{O'} = R\ddot{\theta}\hat{\boldsymbol{n}}_1$$

Euler's equation with moments about a general point in Eq. (2.37) is for this case

$$\dot{\boldsymbol{H}}_{O'} = \boldsymbol{L}_{O'} + m\ddot{\boldsymbol{r}}_{O'} \times \boldsymbol{\sigma}$$

which leads to the desired expression for F_f in terms of $\ddot{\theta}$.

$$RF_f = ml^2\ddot{\theta} - mgl\sin(\theta+\alpha) + mRl\ddot{\theta}\cos\theta$$

Substituting this expression back into the previous equation relating $\ddot{\theta}$ and F_f results in the equations of motion in terms of the rotation angle θ.

$$\left(R^2 + l^2 + 2Rl\cos\theta\right)\ddot{\theta} - Rl\dot{\theta}^2\sin\theta - gR\sin\alpha - gl\sin(\theta+\alpha) = 0$$

This equation could be solved for the angular acceleration $\ddot{\theta}$ which could then be used to find the normal force component N purely in terms of θ and $\dot{\theta}$.

2.3 Dynamics of a System of Particles

2.3.1 Equations of Motion

Until now we have only considered dynamical systems with a single particle. In this section we will discuss systems of N particles each with a constant mass m_i. An example to visualize such dynamical systems would be our solar system. To study the translational (orbital) motion of the planets and moons, due to the large distances involved, they can usually be considered to be point masses with each having different masses m_i.

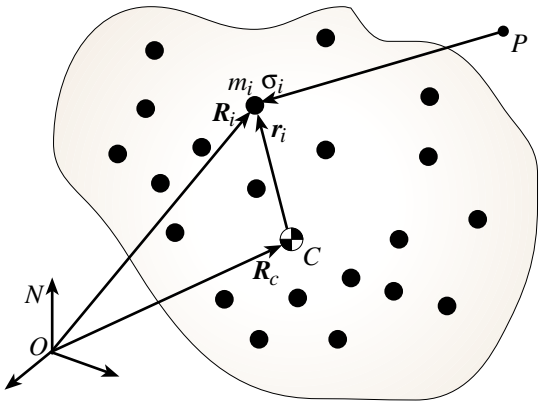


Figure 2.5: System of N Particles

Since we are now dealing with a finite number of masses, we write Newton's second law in index form as

$$\boldsymbol{F}_i = m_i \ddot{\boldsymbol{R}}_i \tag{2.39}$$

where $\ddot{\boldsymbol{R}}_i$ is the inertial acceleration vector of m_i as shown in Figure 2.5. The force acting on m_i can be broken down into two subsets of forces as

$$\boldsymbol{F}_i = \boldsymbol{F}_{iE} + \sum_{j=1}^{N} \boldsymbol{F}_{ij} \tag{2.40}$$

where $\boldsymbol{F}_{iE}$ is the vector sum of all external forces acting on mass m_i and $\boldsymbol{F}_{ij}$ is an internal force vector due to the influence of the j-th masses on the i-th mass. The total force vector $\boldsymbol{F}$ acting on the system of N particles is defined to be

$$\boldsymbol{F} = \sum_{i=1}^{N} \boldsymbol{F}_i = \sum_{i=1}^{N} \boldsymbol{F}_{iE} \tag{2.41}$$

The internal forces $\boldsymbol{F}_{ij}$ don't appear in $\boldsymbol{F}$ because of Newton's third law which states that $\boldsymbol{F}_{ij} = -\boldsymbol{F}_{ji}$, i.e., internal forces cancel in pairs. The total mass M of the N particles is defined as

$$M = \sum_{i=1}^{N} m_i \tag{2.42}$$

The system center of mass position vector $\boldsymbol{R}_c$ is defined such that

$$\sum_{i=1}^{N} m_i \boldsymbol{r}_i = 0 \tag{2.43}$$

where $\boldsymbol{r}_i = (\boldsymbol{R}_i - \boldsymbol{R}_c)$ is the position vector of m_i relative to $\boldsymbol{R}_c$. Thus Eq. (2.43) can be rewritten as

$$\sum_{i=1}^{N} m_i \boldsymbol{R}_c = \sum_{i=1}^{N} m_i \boldsymbol{R}_i \tag{2.44}$$

which is further simplified using the system mass definition in Eq. (2.42) to

$$M\boldsymbol{R}_c = \sum_{i=1}^{N} m_i \boldsymbol{R}_i \tag{2.45}$$

The center of mass position vector $\boldsymbol{R}_c$ is expressed in terms of the individual inertial mass position vectors $\boldsymbol{R}_i$ as

$$\boldsymbol{R}_c = \frac{1}{M} \sum_{i=1}^{N} m_i \boldsymbol{R}_i \tag{2.46}$$

After taking two inertial derivatives of Eq. (2.45) we obtain

$$M\ddot{\boldsymbol{R}}_c = \sum_{i=1}^{N} m_i \ddot{\boldsymbol{R}}_i = \sum_{i=1}^{N} \boldsymbol{F}_i \tag{2.47}$$

After substituting Eq. (2.41) we obtain the final result

$$M\ddot{\boldsymbol{R}}_c = \boldsymbol{F} \tag{2.48}$$

also known as the *Super Particle Theorem*. The dynamics of the mass center of the system of N particles under the influence of the total external force vector $\boldsymbol{F}$ is the same as the dynamics of the "superparticle" M. Note that the superparticle theorem only tracks the center of mass motion of the system. No information is obtained about the size, shape or orientation of the cloud of N particles.

Example 2.6: Let three masses be connected through springs with a spring stiffness constant k as shown in Figure 2.6. The second and third mass each are subjected to a constant force where $F_2 = f$ and $F_3 = 2f$.

The total system mass M is given through

$$M = 2m + m + m = 4m$$

and the total external force F being applied to the system is

$$F = f + 2f = 3f$$

The center of mass of the three-mass system is given found through Eq. (2.45) to

$$r_c = \frac{2mr_1 + mr_2 + mr_3}{M} = \frac{2r_1 + r_2 + r_3}{4}$$

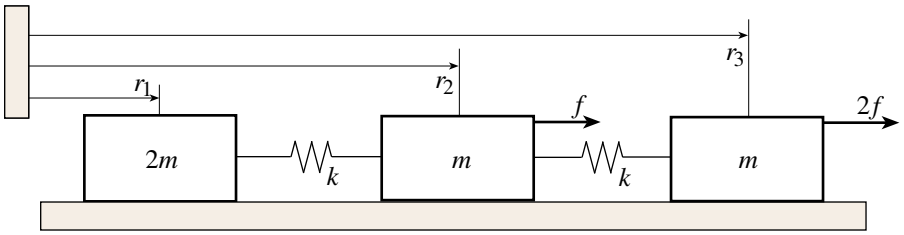


Figure 2.6: Three-Mass System

Using the super particle theorem in Eq. (2.48), the equations of motion for the center of mass of the three-mass system is

$$4m\ddot{r}_c = 3f$$

Assuming that the r_c is originally at rest at the origin, the system center of mass location is then integrated to obtain

$$r_c(t) = \frac{3f}{8m}t^2$$

To find the equations of motion of the individual masses, we need to write Eq. (2.39) for each mass.

$$\begin{aligned} 2m\ddot{r}_1 &= k(r_2 - r_1) \\ m\ddot{r}_2 &= -k(r_2 - r_1) + k(r_3 - r_2) + f \\ m\ddot{r}_3 &= -k(r_3 - r_2) + 2f \end{aligned}$$

This can be written in a standard ODE matrix form for a vibrating system

$$\begin{bmatrix} 2m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & m \end{bmatrix} \begin{pmatrix} \ddot{r}_1 \\ \ddot{r}_2 \\ \ddot{r}_3 \end{pmatrix} + \begin{bmatrix} k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & k \end{bmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} = \begin{pmatrix} 0 \\ f \\ 2f \end{pmatrix}$$

which can be solved given a set of initial conditions for $r_i(t_0)$ and $\dot{r}_i(t_0)$.

2.3.2 Kinetic Energy

The total kinetic energy T of the cloud of N particles can be written as the sum of the kinetic energies of each particle.

$$T = \frac{1}{2}\sum_{i=1}^{N} m_i \dot{\boldsymbol{R}}_i \cdot \dot{\boldsymbol{R}}_i \tag{2.49}$$

After making use of the expression $\dot{\boldsymbol{R}}_i = \dot{\boldsymbol{R}}_c + \dot{\boldsymbol{r}}_i$, the total kinetic energy is rewritten as

$$T = \frac{1}{2}\left(\sum_{i=1}^{N} m_i\right)\dot{\boldsymbol{R}}_c \cdot \dot{\boldsymbol{R}}_c + \dot{\boldsymbol{R}}_c \cdot \left(\sum_{i=1}^{N} m_i \dot{\boldsymbol{r}}_i\right) + \frac{1}{2}\sum_{i=1}^{N} m_i \dot{\boldsymbol{r}}_i \cdot \dot{\boldsymbol{r}}_i \tag{2.50}$$

where the middle term $\sum_{i=1}^{N} m_i \dot{\boldsymbol{r}}_i$ is zero due to the definition of the center of mass in Eq. (2.43). The total kinetic energy of a system of N constant mass particles m_i can therefore be written as

$$T = \frac{1}{2} M \dot{\boldsymbol{R}}_c \cdot \dot{\boldsymbol{R}}_c + \frac{1}{2} \sum_{i=1}^{N} m_i \dot{\boldsymbol{r}}_i \cdot \dot{\boldsymbol{r}}_i \tag{2.51}$$

where the first term contains the system translational kinetic energy and the second contains the system rotation and deformation kinetic energy.

To find the work done on the system we examine the energy rate dT/dt.

$$\frac{dT}{dt} = M \ddot{\boldsymbol{R}}_c \cdot \dot{\boldsymbol{R}}_c + \sum_{i=1}^{N} m_i \ddot{\boldsymbol{r}}_i \cdot \dot{\boldsymbol{r}}_i \tag{2.52}$$

After making use of the facts that $M \ddot{\boldsymbol{R}}_c = \boldsymbol{F}$ and that $\ddot{\boldsymbol{r}}_i = \ddot{\boldsymbol{R}}_i - \ddot{\boldsymbol{R}}_c$, the energy rate is written as

$$\frac{dT}{dt} = \boldsymbol{F} \cdot \dot{\boldsymbol{R}}_c + \sum_{i=1}^{N} m_i \ddot{\boldsymbol{R}}_i \cdot \dot{\boldsymbol{r}}_i - \ddot{\boldsymbol{R}}_c \cdot \left(\sum_{i=1}^{N} m_i \dot{\boldsymbol{r}}_i \right) \tag{2.53}$$

After using Eqs. (2.39) and (2.43), the energy rate is written in the final form as

$$\frac{dT}{dt} = \boldsymbol{F} \cdot \dot{\boldsymbol{R}}_c + \sum_{i=1}^{N} \boldsymbol{F}_i \cdot \dot{\boldsymbol{r}}_i \tag{2.54}$$

If only conservative forces are acting on m_i, then the forces $\boldsymbol{F}_i$ can be written as the gradient of a potential function $V_i(\boldsymbol{r}_i)$.

$$\boldsymbol{F}_i = -\frac{\partial V_i}{\partial \boldsymbol{r}_i} \tag{2.55}$$

Noting that $\frac{\partial V_i \partial \boldsymbol{r}_i}{\dot{\boldsymbol{r}}_i} = \dot{V}_i$ and defining the total conservative potential function to be

$$\frac{d}{dt} V = \sum_{i=1}^{N} \dot{V}_i \tag{2.56}$$

Eq. (2.54) can be written as

$$\frac{dT}{dt} + \frac{dV}{dt} = \boldsymbol{F} \cdot \dot{\boldsymbol{R}}_c \tag{2.57}$$

Studying Eq. (2.57) it is clear that for systems where the total applied force vector $\boldsymbol{F}$ is zero, the total system energy $E = T + V$ is conserved. If the total

resultant force $\boldsymbol{F}$ is itself a conservative force due to a potential function $V_c(\boldsymbol{R}_c)$, then Eq. (2.57) can be written as

$$\frac{dT}{dt} + \frac{dV}{dt} + \frac{dV_c}{dt} = 0 \tag{2.58}$$

and the total system energy $E = T + V + V_c$ is also conserved.

After integrating the kinetic energy rate equation in Eq. (2.54) with respect to time, the change in kinetic energy between two times is given by the work/energy equation

$$T(t_2) - T(t_1) = \int_{t_1}^{t_2} \boldsymbol{F} \cdot \dot{\boldsymbol{R}}_c dt + \sum_{i=1}^{N} \int_{t_1}^{t_2} \boldsymbol{F}_i \cdot \dot{\boldsymbol{r}}_i dt \tag{2.59}$$

which can also be written as the spatial integral

$$T(t_2) - T(t_1) = \int_{\boldsymbol{R}_c(t_1)}^{\boldsymbol{R}_c(t_2)} \boldsymbol{F} \cdot d\boldsymbol{R}_c + \sum_{i=1}^{N} \int_{\boldsymbol{r}(t_1)}^{\boldsymbol{r}(t_2)} \boldsymbol{F}_i \cdot d\boldsymbol{r}_i \tag{2.60}$$

The first term on the right hand side of Eq. (2.60) is the translational work done and the second term is the rotation and deformation work done on the system.

2.3.3 Linear Momentum

In Eq. (2.30) the linear momentum $\boldsymbol{p}_i$ of a single particle is defined. For a system of particles, the total linear momentum of the system is defined as the sum

$$\boldsymbol{p} = \sum_{i=1}^{N} \boldsymbol{p}_i = \sum_{i=1}^{N} \left(m_i \dot{\boldsymbol{R}}_i \right) \tag{2.61}$$

Let r_i be the i-th particle position vector relative to the system center of mass as defined in Eq. (2.43). Taking the derivative of Eq. (2.45), we are able to write the total linear momentum expression in Eq. (2.61) in terms of the total system mass M and the center of mass inertial velocity vector $\dot{\boldsymbol{R}}_c$.

$$\boldsymbol{p} = M \dot{\boldsymbol{R}}_c \tag{2.62}$$

Note that the super particle theorem introduced in Eq. (2.48) also holds for the linear momentum of a system of particles. The linear momentum of the mass center of the system of N particles is the same as the linear momentum of the "superparticle" M.

Let $\boldsymbol{F}_i$ be the force acting on the i-th particle. Note that $\boldsymbol{F}_i$ is composed both of a net external force component $\boldsymbol{F}_{iE}$ and the inertial force component F_{ij} due to interaction with other particles (see Eq. (2.40)). Using the particle

equations of motion in Eq. (2.39), the inertial time rate of change of the total linear momentum of the particle system is expressed as

$$\dot{\boldsymbol{p}} = \sum_{i=1}^{N} \left(m_i \ddot{\boldsymbol{R}}_i \right) = \sum_{i=1}^{N} (\boldsymbol{F}_i) \tag{2.63}$$

Since the inertial forces $\boldsymbol{F}_{ij}$ will cancel each other in this summation due to Newton's third law, the time rate of change of the linear momentum of a particle system is equal to the total external force acting on the system.

$$\boldsymbol{F} = \frac{{}^{\mathcal{N}}d}{dt} (\boldsymbol{p}) \tag{2.64}$$

If no external force $\boldsymbol{F}$ is present, then the total system linear momentum vector $\boldsymbol{p}$ will be constant. This leads to the important law of *conservation of angular momentum*. Unless an external force is acting on a system of N particles, the total linear momentum of the system is conserved. This property is used extensively in collision problems or in the rocket propulsion problem. If two bodies collide, then energy is used to deform the bodies. The total system energy is not conserved during the collision. However, momentum is conserved and can be used to compute the velocities of the bodies after the collision.

Example 2.7: Assume the dynamical system of interest consists of only two particles m_1 and m_2 moving along a one-dimensional, frictionless track at different rates. Before a collision at time t_0 they each have a constant speed of $v_1(t_0^-)$ and $v_2(t_0^-)$ respectively. The total energy before the impact is given by

$$T(t_0^-) = \frac{1}{2} \left(m_1 v_1(t_0^-)^2 + m_2 v_2(t_0^-)^2 \right)$$

The total linear momentum is

$$p(t_0^-) = m_1 v_1(t_0^-) + m_2 v_2(t_0^-)$$

First, Let assume that the collision is perfectly elastic. In this case any energy used to deform the bodies during the collision is regained when the body shapes are restored (i.e. think of two rubber balls colliding). Both total energy $T(t_0^+)$

$$T(t_0^+) = \frac{1}{2} \left(m_1 v_1(t_0^+)^2 + m_2 v_2(t_0^+)^2 \right)$$

and momentum $p(t_0^+)$

$$p(t_0^+) = m_1 v_1(t_0^+) + m_2 v_2(t_0^+)$$

are conserved during the collision. Setting $T(t_0^-) = T(t_0^+)$ and $p(t_0^-) = p(t_0^+)$, we are able to express the particles speeds after the collision as

$$v_1(t_0^+) = \frac{1}{M} \left(v_1(t_0^-)(m_1 - m_2) + 2 v_2(t_0^-) m_2 \right)$$
$$v_2(t_0^+) = \frac{1}{M} \left(v_2(t_0^-)(m_2 - m_1) + 2 v_1(t_0^-) m_1 \right)$$

with $M = m_1 + m_2$ being the total system mass.

Second, we assume that the collision is such that the two particles join and become one (i.e. think of two chunks of clay colliding). In this case the total energy $T(t_0^+)$ after the collision is given by

$$T(t_0^+) = \frac{1}{2}Mv^2$$

where v is the speed of the joined particles after the collision. The linear momentum of the joined particles is

$$p(t_0^+) = Mv$$

Note that this collision in not perfectly elastic and that energy is not conserved. However, linear momentum is conserved and we can set $p(t_0^-) = p(t_0^+)$ to solve for the velocity v of the joined particle after the collision.

$$v = \frac{1}{M}\left(m_1 v_1(t_0^-) + m_2 v_2(t_0^+)\right)$$

The total energy after the collision is given by

$$T(t_0^+) = \frac{1}{2}Mv^2 = \frac{1}{2M}\left(m_1 v_1(t_0^-) + m_2 v_2(t_0^+)\right)^2 = \frac{p^2}{2M}$$

The change in energy $\Delta T = T(t_0^+) - T(t_0^-)$ is given by

$$\Delta T = -\frac{m_1 m_2}{2M}\left(v_1(t_0^-) - v_2(t_0^-)\right)^2$$

The energy lost during this plastic collision is used to permanently deform the two bodies, as well as to radiate heat and produce sound waves.

These two examples are idealized situations. In reality the collisions are never perfectly elastic or plastic. In this case more knowledge is required about the how the bodies will deform to predict the motion after the collision.

2.3.4 Angular Momentum

As was done for the case of a single particle, let's find the angular momentum of the N particle system about an arbitrary point P given by the inertial position vector $\boldsymbol{R}_P$. The relative position of each mass m_i is given though the vector

$$\boldsymbol{\sigma}_i = \boldsymbol{R}_i - \boldsymbol{R}_P \tag{2.65}$$

The total system angular momentum vector $\boldsymbol{H}_P$ about the point P is given as the sum of all the single particle angular momentum vectors about this point.

$$\boldsymbol{H}_P = \sum_{i=1}^{N} \boldsymbol{\sigma}_i \times m_i \dot{\boldsymbol{\sigma}}_i \tag{2.66}$$

Taking the time derivative of $\boldsymbol{H}_P$ we get

$$\dot{\boldsymbol{H}}_P = \sum_{i=1}^{N} \dot{\boldsymbol{\sigma}}_i \times m_i \dot{\boldsymbol{\sigma}}_i + \sum_{i=1}^{N} \boldsymbol{\sigma}_i \times m_i \ddot{\boldsymbol{\sigma}}_i \tag{2.67}$$

After performing similar arguments as in the single particle case this expression is rewritten as

$$\dot{\boldsymbol{H}}_P = \sum_{i=1}^{N} \boldsymbol{\sigma}_i \times m_i \ddot{\boldsymbol{R}}_i - \left(\sum_{i=1}^{N} \boldsymbol{\sigma}_i m_i \right) \times \ddot{\boldsymbol{R}}_P \tag{2.68}$$

Using Eqs. (2.45), (2.65), the following mass center identity is found.

$$\sum_{i=1}^{N} \boldsymbol{\sigma}_i m_i = \sum_{i=1}^{N} \boldsymbol{R}_i m_i - \left(\sum_{i=1}^{N} m_i \right) \boldsymbol{R}_P = M \left(\boldsymbol{R}_c - \boldsymbol{R}_P \right) \tag{2.69}$$

The total external moment $\boldsymbol{L}_P$ applied to the system is defined to be

$$\boldsymbol{L}_P = \sum_{i=1}^{N} \boldsymbol{\sigma}_i \times m_i \ddot{\boldsymbol{R}}_i = \sum_{i=1}^{N} \boldsymbol{\sigma}_i \times \boldsymbol{F}_i \tag{2.70}$$

Using Eqs. (2.69) and (2.70), the system angular momentum derivative $\dot{\boldsymbol{H}}_P$ about a point P is[3]

$$\dot{\boldsymbol{H}}_P = \boldsymbol{L}_P + M \ddot{\boldsymbol{R}}_P \times \left(\boldsymbol{R}_c - \boldsymbol{R}_P \right) \tag{2.71}$$

Note that if either $\boldsymbol{R}_c = \boldsymbol{R}_P$ or $\boldsymbol{R}_P$ is non-accelerating inertially, then Eq. (2.71) reduces to the most familiar Euler equation

$$\dot{\boldsymbol{H}}_P = \boldsymbol{L}_P \tag{2.72}$$

Analogously to the linear momentum development, if no external torque $\boldsymbol{L}_P$ is acting on the system of particles, then the total angular momentum rate vector $\boldsymbol{H}_P$ is constant.

Example 2.8: Two particles are attached on strings and are moving in a planar, circular manner as shown in Figure 2.7. The plane on which the particles are moving is level compared to the gravity field. Thus, given an initial velocity and ignoring frictional effects, the particles will continue to move at a constant rate. Assume that the two circular paths meet tangentially at one point. We would like to investigate how the velocities will change if the particles meet at this point at time t_0. This condition is shown in grey in the figure. The total kinetic energy before the collision is

$$T(t_0^-) = \frac{m_1}{2} m_1 v_1 (t_0^-)^2 + \frac{m_2}{2} m_2 v_2 (t_0^-)^2$$

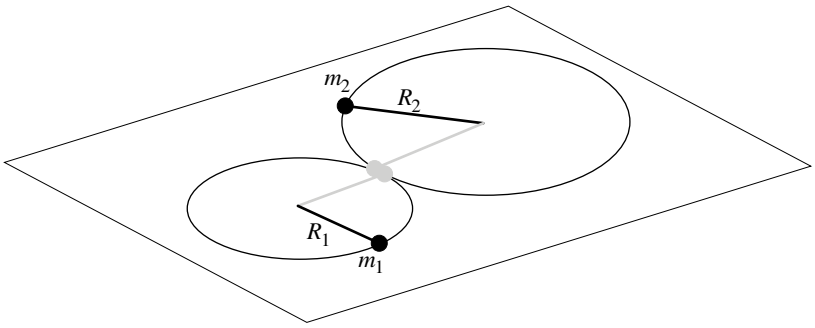


Figure 2.7: Illustration of two Particles Moving in a Circular Manner on a Level Plane

while the angular momentum H along the plane normal direction is

$$H(t_0^-) = R_1 m_1 v_1(t_0^-) + R_2 m_2 v_2(t_0^-)$$

Assuming the collision is perfectly elastic, then both the total energy and angular momentum are conserved. After the collision, we express them as

$$T(t_0^+) = \frac{m_1}{2} m_1 v_1(t_0^+)^2 + \frac{m_2}{2} m_2 v_2(t_0^+)^2$$
$$H(t_0^+) = R_1 m_1 v_1(t_0^+) + R_2 m_2 v_2(t_0^+)$$

Setting $T(t_0^-) = T(t_0^+)$ and $H(t_0^-) = H(t_0^+)$, we are able to solve for the particle velocities after the collision:

$$v_1(t_0^+) = \frac{(m_1 R_1^2 - m_2 R_2^2) v_1(t_0^-) + 2 m_2 R_1 R_2 v_2(t_0^-)}{m_1 R_1^2 + m_2 R_2^2}$$
$$v_2(t_0^+) = \frac{2 m_1 R_1 R_2 v_1(t_0^-) + (m_2 R_1^2 - m_1 R_1^2) v_2(t_0^-)}{m_1 R_1^2 + m_2 R_2^2}$$

2.4 Dynamics of a Continuous System

2.4.1 Equations of Motion

The development of the dynamical equations of motion of a continuous system parallels that of the system of N particles. Any finite sums over all particles are generally replaced with volume integrals over the body $\mathcal{B}$. This allows us to describe any constant mass body, even if it is flexible or does not have a constant shape as in a chunk of jello. However, care must be taken to define a control volume that contains the instantaneous mass of the system when actually carrying out volume integrations.

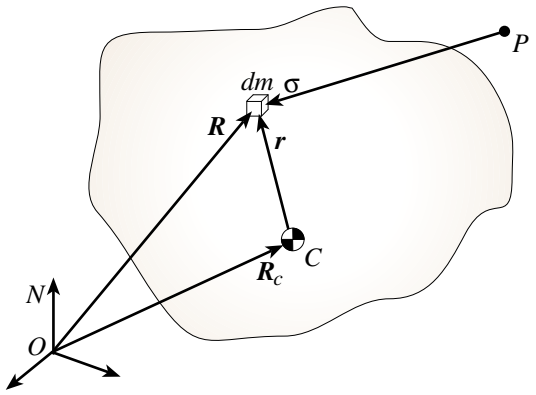


Figure 2.8: Mass Element of Continuous System

Let dm be an infinitesimal body element with the corresponding inertial position vector $\boldsymbol{R}$ as shown in Figure 2.8. Then as such it can be considered to be a particle and abides by Newton's second law. The equations of motion for this infinitesimal element are

$$d\boldsymbol{F} = \ddot{\boldsymbol{R}} dm \tag{2.73}$$

where $d\boldsymbol{F}$ is the total force acting on dm. The force vector $d\boldsymbol{F}$ is broken up into external and internal components as

$$d\boldsymbol{F} = d\boldsymbol{F}_E + d\boldsymbol{F}_I \tag{2.74}$$

To express the volume integral over the body $\mathcal{B}$ let us use the shorthand notation $\int_{\mathcal{B}} = \iiint_{\mathcal{B}}$. The total force $\boldsymbol{F}$ acting on this continuous body is given by

$$\boldsymbol{F} = \int_B d\boldsymbol{F} = \int_B d\boldsymbol{F}_E \tag{2.75}$$

where the internal forces again cancel because of Newton's third law. The total body mass is given by

$$M = \int_B dm \tag{2.76}$$

The system center of mass is defined such that

$$\int_B \boldsymbol{r} dm = 0 \tag{2.77}$$

where $\boldsymbol{r} = \boldsymbol{R} - \boldsymbol{R}_c$ is again the internal position vector of dm relative to $\boldsymbol{R}_c$. Therefore Eq. (2.77) can be rewritten as

$$M\boldsymbol{R}_c = \int_B \boldsymbol{R} dm \tag{2.78}$$

The center of mass vector $\boldsymbol{R}_c$ is then expressed as

$$\boldsymbol{R}_c = \frac{1}{M}\int_B \boldsymbol{R} dm \tag{2.79}$$

After twice differentiating Eq. (2.78) we obtain

$$M\ddot{\boldsymbol{R}}_c = \int_B \ddot{\boldsymbol{R}} dm = \int_B d\boldsymbol{F} \tag{2.80}$$

After substituting Eq. (2.75) we obtain the equivalent *super particle theorem* for a continuous body.

$$M\ddot{\boldsymbol{R}}_c = \boldsymbol{F} \tag{2.81}$$

2.4.2 Kinetic Energy

Let the inertial vector $\boldsymbol{R}$ define the position of the infinitesimal mass element dm. The kinetic energy of the entire continuous body $\mathcal{B}$ is then given as

$$T = \frac{1}{2}\int_B \dot{\boldsymbol{R}} \cdot \dot{\boldsymbol{R}} dm \tag{2.82}$$

After substituting $\dot{\boldsymbol{R}} = \dot{\boldsymbol{R}}_c + \dot{\boldsymbol{r}}$ the kinetic energy is expressed as

$$T = \frac{1}{2}\left(\int_B dm\right)\dot{\boldsymbol{R}}_c \cdot \dot{\boldsymbol{R}}_c + \dot{\boldsymbol{R}}_c \cdot \int_B \dot{\boldsymbol{r}} dm + \frac{1}{2}\int_B \dot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}} dm \tag{2.83}$$

Making use of Eqs. (2.76) and (2.77), the kinetic energy for a continuous body $\mathcal{B}$ is written as

$$T = \frac{1}{2}M\dot{\boldsymbol{R}}_c \cdot \dot{\boldsymbol{R}}_c + \frac{1}{2}\int_B \dot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}} dm \tag{2.84}$$

The first term in Eq. (2.84) represents the translational kinetic energy and the second the rotational and deformational energy.

To find the work done on the continuous body $\mathcal{B}$ the kinetic energy rate is found.

$$\frac{dT}{dt} = M\ddot{\boldsymbol{R}}_c \cdot \dot{\boldsymbol{R}}_c + \int_B \dot{\boldsymbol{r}} \cdot \ddot{\boldsymbol{r}} dm \tag{2.85}$$

After using Eq. (2.81) and the fact that $\ddot{\boldsymbol{r}} = \ddot{\boldsymbol{R}} - \ddot{\boldsymbol{R}}_c$ the kinetic energy rate is given as

$$\frac{dT}{dt} = \boldsymbol{F} \cdot \dot{\boldsymbol{R}}_c + \int_B \left(\ddot{\boldsymbol{R}} dm\right) \cdot \dot{\boldsymbol{r}} - \ddot{\boldsymbol{R}}_c \cdot \int_B \dot{\boldsymbol{r}} dm \tag{2.86}$$

Using Eqs. (2.73) and (2.77) the kinetic energy rate for a continuous, constant mass body $\mathcal{B}$ is given by

$$\frac{dT}{dt} = \boldsymbol{F} \cdot \dot{\boldsymbol{R}}_c + \int_B d\boldsymbol{F} \cdot \dot{\boldsymbol{r}} \tag{2.87}$$

The change in kinetic energy between two times is found by integrating the kinetic energy rate expression with respect to time.

$$T(t_2) - T(t_1) = \int_{t_1}^{t_2} \boldsymbol{F} \cdot \dot{\boldsymbol{R}}_c dt + \int_{t_1}^{t_2} \int_B d\boldsymbol{F} \cdot \dot{\boldsymbol{r}} dt \tag{2.88}$$

This can be also written alternatively as a spatial integration:

$$T(t_2) - T(t_1) = \int_{\boldsymbol{R}(t_1)}^{\boldsymbol{R}(t_2)} \boldsymbol{F} \cdot d\boldsymbol{R}_c + \int_{\boldsymbol{r}(t_1)}^{\boldsymbol{r}(t_2)} \int_B d\boldsymbol{F} \cdot d\boldsymbol{r} \tag{2.89}$$

where the first term expresses the translational work and the second term is the rotational and deformational work done on the system.

2.4.3 Linear Momentum

To determine the total linear momentum of a continuous body $\mathcal{B}$, we express the linear momentum of an infinitesimal body element dm as

$$d\boldsymbol{p} = \dot{\boldsymbol{R}} dm \tag{2.90}$$

Integrating the infinitesimal linear momentum contributions over the entire body, the total linear momentum is given by

$$\boldsymbol{p} = \int_{\mathcal{B}} d\boldsymbol{p} = \int_{\mathcal{B}} \dot{\boldsymbol{R}} dm \tag{2.91}$$

Using the center of mass property in Eq. (2.78), the total linear momentum of the body is written directly in terms of the body mass M and the center of mass motion $\dot{\boldsymbol{R}}_c$.

$$\boldsymbol{p} = M \dot{\boldsymbol{R}}_c \tag{2.92}$$

Again the super particle theorem applies to the continuous body. The sum of the individual infinitesimal linear momenta of the body is the same as the linear momenta of a particle of mass M with the same velocity vector as the body center of mass motion. Note that the body $\mathcal{B}$ is not restricted to be a rigid body in this section. If the body center of mass is inertially stationary (i.e. the body has zero linear momentum), it is still possible for various body components to be moving inertially. For example, consider a heap of jello floating in space. It is possible for the jello to be deforming without moving. While the individual components of jello might have some linear momentum, the total sum of these components cancel each other out to result in a zero net motion of the body center of mass.

Taking the inertial derivative of Eq. (2.92) and making use of the internal/external force properties in Eqs. (2.74) and (2.75), we express the total linear momentum rate as

$$\dot{\boldsymbol{p}} = \int_{\mathcal{B}} \ddot{\boldsymbol{R}} dm = \int_{\mathcal{B}} d\boldsymbol{F} = \boldsymbol{F} \tag{2.93}$$

Thus, the time rate of change of the total linear momentum of a continuous body $\mathcal{B}$ is equal to the total external force vector being applied to this body. If no external is applied, then the total linear momentum is conserved and its rate is zero.

2.4.4 Angular Momentum

To find the angular momentum vector of the continuous body $\mathcal{B}$ about an arbitrary point P, we write the relative position vector $\boldsymbol{\sigma}$ of dm to P as

$$\boldsymbol{\sigma} = \boldsymbol{R} - \boldsymbol{R}_P \tag{2.94}$$

The total system angular momentum vector $\boldsymbol{H}_P$ about P is then given by

$$\boldsymbol{H}_P = \int_B \boldsymbol{\sigma} \times \dot{\boldsymbol{\sigma}} dm \tag{2.95}$$

Taking the derivative of $\boldsymbol{H}_P$ we get

$$\dot{\boldsymbol{H}}_P = \int_B \dot{\boldsymbol{\sigma}} \times \dot{\boldsymbol{\sigma}} dm + \int_B \boldsymbol{\sigma} \times \ddot{\boldsymbol{\sigma}} dm \tag{2.96}$$

which can be rewritten as

$$\dot{\boldsymbol{H}}_P = \int_B \boldsymbol{\sigma} \times \ddot{\boldsymbol{R}} dm - \left(\int_B \boldsymbol{\sigma} dm \right) \times \ddot{\boldsymbol{R}}_P \tag{2.97}$$

The term in the brackets can be expanded to

$$\int_B \boldsymbol{\sigma} dm = \int_B \boldsymbol{R} dm - \left(\int_B dm \right) \boldsymbol{R}_P = M(\boldsymbol{R}_c - \boldsymbol{R}_P) \tag{2.98}$$

The total external moment $\boldsymbol{L}_P$ applied to the system is defined to be

$$\boldsymbol{L}_P = \int_B \boldsymbol{\sigma} \times \ddot{\boldsymbol{R}} dm = \int_B \boldsymbol{\sigma} \times d\boldsymbol{F} \tag{2.99}$$

Using these two identities in Eqs. (2.98) and (2.99) the system angular momentum derivative vector $\dot{\boldsymbol{H}}_P$ about P is

$$\dot{\boldsymbol{H}}_P = \boldsymbol{L}_P + M\ddot{\boldsymbol{R}}_P \times (\boldsymbol{R}_c - \boldsymbol{R}_P) \tag{2.100}$$

As was the case with the system of N particles, if either $\boldsymbol{R}_c = \boldsymbol{R}_P$ or the vector $\boldsymbol{R}_P$ is non-accelerating inertially, then Eq. (2.100) reduces to the Euler equation[2]

$$\dot{\boldsymbol{H}}_P = \boldsymbol{L}_P \tag{2.101}$$

As was the case with the dynamical system of finite particles, the angular momentum of a continuous body is constant if no external torque vector $\boldsymbol{L}_P$ is applied.

2.5 The Rocket Problem

In this section we investigate the thrust that a rocket motor produces by expelling propellant at a high velocity from the spacecraft. Consider the one-stage rocket shown in Figure 2.9. Let m be the mass of the rocket including any propellant that is currently on board. The propellant fuel is being burnt and ejected at a mass flow rate of $\dot{m}$. The current velocity vector of the rocket is $\boldsymbol{v}$, while the exhaust velocity of the ejected propellant particles dm relative to the rocket is $\boldsymbol{v}_e$. Note that the orientation of the exhaust velocity vector $\boldsymbol{v}_e$ does not have to point aftward. If the nozzle would be pointing forward, then the engine would be used to perform a breaking maneuver. The rocket is assumed to be flying through an atmosphere with an ambient pressure P_a. At the point where the exhaust gases escape the engine nozzle the exhaust pressure is given by P_e.

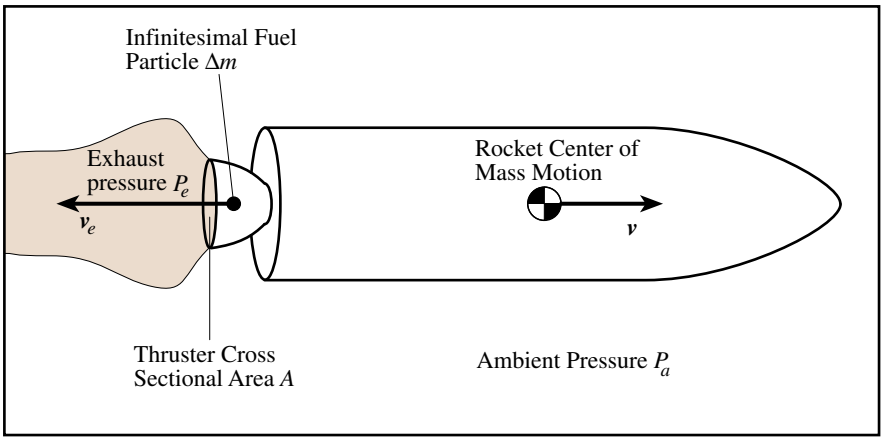


Figure 2.9: A One-Stage Rocket Expelling a Propellant Particle Δm with an Ambient Atmosphere p_a.

We would like to develop the thrust vector that the rocket engine is exerting onto the spacecraft. To do so, we utilize Eq. (2.72) or (2.101) which state that the external force $\boldsymbol{F}$ exerted onto a system of particles or a continuous body is equal time rate of change in linear momentum. Let us treat the rocket mass m and the expelled propellant particle Δm as a two particle system and track their linear momentum change over a small time interval Δt. Using Eq. (2.72) we can write the momentum equation as

$$\boldsymbol{F}\Delta t = \boldsymbol{p}(t + \Delta t) - \boldsymbol{p}(t) \tag{2.102}$$

The quantity $\boldsymbol{F}\Delta t$ is the impulse being applied to the system over the time interval dt. At time t is the rocket and propellant mass is still m. At time $t + \Delta t$, the rocket mass has been reduced to $m - dm$ and the propellant particle Δm is about to leave the engine nozzle. Assume that the only external force acting on this two-particle system is due to pressure differential at the engine

nozzle. Let A be the nozzle cross sectional area, then the external force $\boldsymbol{F}$ is expressed as

$$\boldsymbol{F} = -\frac{\boldsymbol{v}_e}{v_e} A\,(P_e - P_a) \tag{2.103}$$

More generally, however, we write the external force vector $\boldsymbol{F}$ as

$$\boldsymbol{F} = -\frac{\boldsymbol{v}_e}{v_e} A\,(P_e - P_a) + \boldsymbol{F}_e \tag{2.104}$$

where $\boldsymbol{F}_e$ is the net sum of non-pressure related external forces such as gravitational forces acting on the system. The pressure induced force is assumed to be collinear with the exhaust velocity vector $\boldsymbol{v}_e$. Note that if $P_a = P_e$ (exhaust expands to ambient pressure) or $P_a = P_e = 0$ (operating in a vacuum and exhaust expanding to zero pressure), then the net external force on the system is zero. Further, if the direction of the exhaust velocity vector $\boldsymbol{v}_e$ is in the opposite direction to the rocket velocity vector $\boldsymbol{v}$, then a positive pressure differential $P_e - P_a > 0$ results in an acceleration in the rocket velocity direction.

The linear momentum p of the system at time t is

$$\boldsymbol{p}(t) = m\boldsymbol{v} \tag{2.105}$$

since the propellant particle dm is still joined with the rocket. At time $t + \Delta t$ the small propellant mass Δm is being ejected from the rocket with a relative velocity vector $\boldsymbol{v}_e$. Since the rocket is loosing mass, the mass difference Δm over time dt is a negative quantity. The linear momentum at time $t + \Delta t$ is

$$\boldsymbol{p}(t + \Delta t) = (m + \Delta m)(\boldsymbol{v} + \Delta \boldsymbol{v}) - \Delta m(\boldsymbol{v} + \boldsymbol{v}_e) \tag{2.106}$$

where $(m + \Delta m)$ is the rocket mass without the escaping fuel particle and $\Delta \boldsymbol{v}$ is the change in rocket velocity vector over the time interval Δt. Dropping higher order differential terms in Eq. (2.106) and substituting the $\boldsymbol{F}$, $\boldsymbol{p}(t)$ and $\boldsymbol{p}(t+\Delta t)$ expressions into Eq. (2.102) leads to

$$-\frac{\boldsymbol{v}_e}{v_e} A\,(P_e - P_a)\,\Delta t + \boldsymbol{F}_e \Delta t = m\Delta \boldsymbol{v} - \Delta m \boldsymbol{v}_e \tag{2.107}$$

Dividing both sides by Δt and solving for the acceleration term we find

$$m\frac{\Delta \boldsymbol{v}}{\Delta t} = -\frac{\boldsymbol{v}_e}{v_e} A\,(P_e - P_a) + \frac{\Delta m}{\Delta t}\boldsymbol{v}_e + \boldsymbol{F}_e \tag{2.108}$$

Allowing the time step Δt to become infinitesimally small, we arrive at the rocket equations of motion:

$$m\frac{d\boldsymbol{v}}{dt} = \underbrace{-\boldsymbol{v}_e \left(\frac{A}{v_e}\,(P_e - P_a) - \frac{dm}{dt} \right)}_{\boldsymbol{F}_s} + \boldsymbol{F}_e = \boldsymbol{F}_s + \boldsymbol{F}_e \tag{2.109}$$

The $\boldsymbol{F}_s$ force component is called the *static thrust* of the rocket engine. If the rocket were attached to a test stand, then it would require a force $\boldsymbol{F}_s$ to keep the rocket immobile during the engine test firing.

If the exhaust velocity vector is in the opposite direction to the rocket velocity vector $\boldsymbol{v}$ as shown in Figure 2.9, and the rocket is operating in a weightless environment with $\boldsymbol{F}_e = 0$, then the rocket equations of motion simplify to the famous one-dimensional form

$$m\frac{dv}{dt} = A\left(P_e - P_a\right) - \frac{dm}{dt}v_e = F_s \tag{2.110}$$

with the parameter F_s being the scalar static rocket thrust. Let us assume that over a time interval from t_0 to t_f that the negative mass flow rate $\dot{m}$ is constant. Eq. (2.110) can be rewritten as

$$m\frac{dv}{dm}\dot{m} = F_s \tag{2.111}$$

Rearranging this equation by separating the dv and dm terms, and integrating from t_0 to t_f, we find

$$\int_{v_0}^{v_f} dv = v_f - v_0 = \frac{F_s}{\dot{m}}\int_{m_0}^{m_f}\frac{dm}{m} = -\frac{F_s}{\dot{m}}\ln\left(\frac{m_0}{m_f}\right) \tag{2.112}$$

Note that F_s and $\dot{m}$ can be taken outside the integral sign since they are constants in this investigation. The scalar velocity v_0 is the velocity that the rocket possessed at t_0, while v_f is the rocket velocity at the thruster burnout at t_f. The initial rocket mass is m_0 and the smaller, final rocket mass is m_f. The burn-out velocity v_f can be solve for in terms of the initial rocket velocity and mass, as well as the final burnout mass m_f.

$$v_f = v_0 - \frac{F_s}{\dot{m}}\ln\left(\frac{m_0}{m_f}\right) \tag{2.113}$$

The second term in Eq. (2.113) is a positive quantity since $m_0 > m_f$ and the mass flow rate $\dot{m}$ is a negative quantity. Let $\Delta m < 0$ be the amount of fuel mass lost over the given time interval. Then $m_f = m_0 + \Delta m$. The change in velocity $\Delta v = v_f - v_0$ that results from ejecting Δm of fuel is given by

$$\Delta v = -\frac{F_s}{\dot{m}}\ln\left(\frac{1}{1+\epsilon}\right) \tag{2.114}$$

where $\epsilon = \Delta m/m_0$ is the ratio of fuel spent over the time interval over the initial rocket mass. Note that this change in velocity only depends on the amount of fuel spent and F_s, not on the length of the burning time. Thus, if a thruster produces half the mass flow rate $\dot{m}$ as another thruster, but burns for twice as long, then both thrusters will produce the same velocity change Δv. However, this result is only true if no other external forces are acting on the body. If gravity is pulling on the rocket, then the amount of time spent trying

to accelerate the rocket will have a drastic effect on the rocket velocity at burn out time.

A common measure of rocket thruster efficiency is the *specific impulse* I_{sp} defined as[3, 5]

$$I_{sp} = \frac{F_s}{(-\dot{m})g} \tag{2.115}$$

and has units of seconds. The gravitational acceleration g used here is that experienced on the Earth's surface. The higher this I_{sp} value is, the more force the rocket thruster is able to produce for a given mass flow rate. If the exhaust pressure P_e is close to the ambient pressure P_a, the pressure contribution to the static thrust F_s in Eq. (2.110) is negligible. In this case $F_s \approx -\dot{m}v_e$ and the specific impulse simplifies to

$$I_{sp} \approx \frac{v_e}{g} \tag{2.116}$$

From this simplification it is evident that to achieve higher thruster efficiencies, the exhaust velocity v_e should be as high as possible. The faster a given fuel particle is ejected from the rocket, the larger a momentum change (i.e. rocket speed up) it will cause. Using the specific impulse definition, the rocket velocity change Δv for a given fuel ratio ϵ burned is given by

$$\Delta v = I_{sp} g \ln\left(\frac{1}{1+\epsilon}\right) \tag{2.117}$$

The specific impulse ranges for different rocket thruster systems are shown in Table 2.1.[5] Note that the higher specific impulse propulsion methods, such as the ion or arcjet thrusters, typically produce only a very small thrust. Such modes of propulsion are able to achieve a desired Δv with a much smaller amount of fuel mass Δm than a propulsion method with a lower I_{sp}. However, due to the small amount of thrust produced, these efficient propulsion methods will take a much longer time to produce this desired velocity change.

Example 2.9: Assume we are trying to launch an initially at rest sounding rocket vertically from the Earth's surface and it is to only fly several miles high. For these small altitudes, we are still able to assume that the gravitational attraction g is constant during the flight. The solid rocket motor produces a constant I_{sp} for the duration of it burn. Since the only external force acting on the rocket is the constant gravitational acceleration, the rocket equations of motion in the vertical direction are given by Eq. (2.109):

$$m\dot{v} = F_s - mg = g\,(m - I_{sp}\dot{m}) \tag{2.118}$$

This equation illustrates the challenge that a highly efficient ion propulsion system would have in attempting to launch this sounding rocket. The change in velocity expression given in Eq. (2.114) assumes that no external forces are acting on the rocket except for the ambient and exhaust pressure. With

Table 2.1: Specific Impulse and Thrust Ranges for Different Rocket Thruster Designs

Thruster Type	Vacuum I_{sp} [sec]	Thrust Range [N]	Comments
Solid Motor	280 – 300	$50 - 5 \cdot 10^6$	Simple, reliable low-cost design with a low performance, but a high thrust
Cold Gas	50 – 75	0.05 – 200	Extremely simple and reliable design with a very low performance and heavy weight for the small thrust produced
Liquid Motor	150 – 450	$5 - 5 \cdot 10^6$	Higher performance thruster at the cost of a more complicated mechanical and cryogenic design
Electrothermal Arcjet	450 – 1500	0.05 – 5	Higher performance, low thrust system with a complicated thermal interface
Ion	2000 – 6000	$5 \cdot 10^{-6} - 0.5$	Very high performance system with typically a very low thrust

the gravity force acting on our sounding rocket, the thruster is constantly battling the gravitational acceleration. In fact, if the rocket thrust is less than the weight mg of the rocket, then the propulsion system will not be able to lift the rocket off the launch pad. Thus, while a ion high performance propulsion system is very effective in accelerating a spacecraft in a weightless or free-falling environment, it would be an inappropriate propulsion choice to launch a rocket of a planet's surface. The rocket velocity at burn out time t_f is then given by

$$v_f = g\left(-t_f + I_{sp}\ln\left(\frac{m_0}{m_f}\right)\right) \tag{2.119}$$

The longer the thruster takes to accelerate the rocket to the desired velocity, the longer the thruster must combat the gravitational acceleration. Since the efficient ion propulsion system requires a large time t_f to achieve a desired Δv, the gravity is also given a large amount of time to counter the achievements of the ion thruster. This is why it is common to use solid or liquid chemical propulsion systems to launch a rocket from the planet's surface to a low-Earth orbit. While these propulsion choices are less efficient, they provide a thrust which is much larger than the rocket weight. With this large static thrust the rocket is propelled to the desired velocity quickly and the gravity field has

less time to decelerate the craft.

Problems

2.1 Plot the magnitude of the gravity force as it varies from Earth's surface to a height of 300 km.

2.2 Given a spring with a spring stiffness constant of $k = 5$ kg/s^2 and a stored potential energy of 100 Nm, find the spring deflection x and the force F required to keep the spring at this deflection.

2.3 A mass m is sliding down a constant slope of 10 degrees with an initial velocity of $v(t_0) = 1$ m/s. How long will it take this mass to accelerate to a velocity of $v(t_f) = 10$ m/s and what distance will it have traveled?

2.4 ♣ A skydiver exits an aircraft at an altitude of 3000 m. The aircraft is flying horizontally at 36 m/s. The skydiver has a mass m of 80 kg, a forward projected surface area A of 0.75 m^2 and a coefficient of drag c_d of 0.555. Assume a uniform gravitation field with a gravitational acceleration of 9.81 m/s^2. The air density ρ is 1.293 kg/m^3. Recall the relationship Drag $= \frac{1}{2}\rho v^2 c_d A$, and this force is opposite to the velocity vector.[6]

a) What is the theoretical terminal velocity of this skydiver?

b) Find the skydiver equations of motion and solve them numerically for a 45 second freefall. Plot the altitude versus horizontal position, the skydiver speed versus time and the horizontal / vertical velocity versus time.

c) Taking into account that the more air speed a skydiver has, the better and faster the parachute will open, what is the "worst" time for a skydiver to try to open the parachute?

d) How long does it take for the skydiver to reach 95% of the terminal velocity?

e) What acceleration does the skydiver experience at terminal velocity?

f) How far forward does the skydiver get thrown on exit before he or she essentially descends vertically?

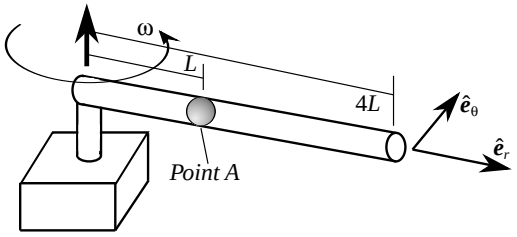


Figure 2.10: Ball in Rotating Tube

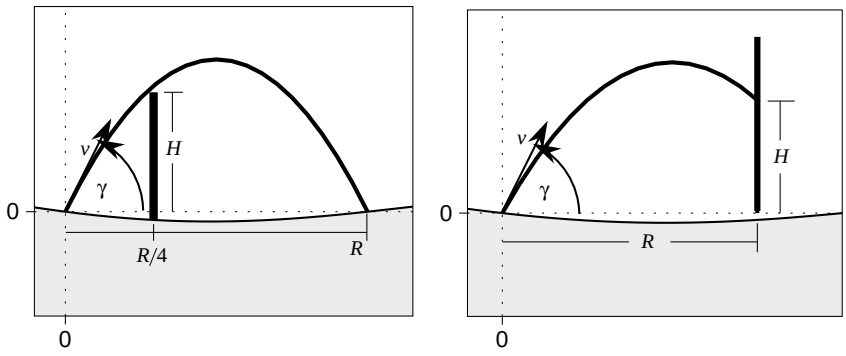
2.5 A ball of mass m is sliding in a frictionless tube as shown in Figure 2.10. The tube is rotating at a constant angular velocity $\boldsymbol{\omega}$. Initially the ball is at rest relative to the tube at Point A at $\boldsymbol{r} = L\hat{\boldsymbol{e}}_r$.

a) What is the velocity vector when the ball exits the tube?

b) Up to the point where the ball exists the tube, how much work has been performed onto the ball?

c) Find an expression for the angular momentum vector $\boldsymbol{H}_A$ of the mass m about point A.

2.6 A cannon tries to hit a target which is a distance R away with a projectile of mass m as shown in Figure 2.11(i). However, at a distance $R/4$ there is an obstacle of height H present. What is the smallest elevation angle γ_0 and corresponding initial speed v_0 the projectile m must possess initially to hit the target and miss the obstacle. Assume a constant gravity field is present.



(i) Clearing an Obstacle **(ii)** Hitting Elevated Target

Figure 2.11: Ballistic Trajectory Problems

2.7 A cannon tries to hit a target which is a distance R away and elevated of the ground by a height H with a projectile m as shown in Figure 2.11(ii). What is the smallest initial velocity v and corresponding heading angle γ the particle may have to hit this target. Assume a constant gravity field is present.

2.8 As shown in Figure 2.12, a ball with mass m is propelled by a spring with a spring stiffness k to roll without friction on a surface until it is launched into the air by a ramp of height h. The departure angle γ is fixed by the ramp *and is not a variable*. The goal is to hit a target on ground level a distance d away from the ramp.

a) Find the initial velocity v_0 the ball must have when leaving the ramp to hit the target.

b) What is the initial compression z the spring must have such that the mass will have the necessary velocity v_0 when leaving the ramp?

2.9 Consider the two-particle system studied in Example 2.7. Verify the results shown by providing all the algebra required to complete the steps outlined.

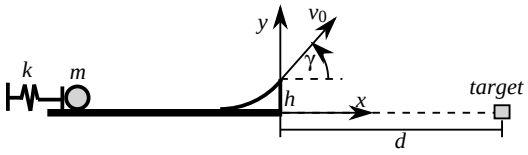


Figure 2.12: Spring Propelled Mass

2.10 A massless cylinder is rolling down a slope with an inclination angle α under the influence of a constant gravity field. A mass m is attached to the cylinder and is offset from the cylinder center by $R/2$ as shown in Figure 2.13.

a) Find the equations of motion of the the mass m in terms of the angle θ.

b) What is the normal force $\boldsymbol{N} = N\hat{\boldsymbol{n}}_2$ that the ground is exerting against the cylinder.

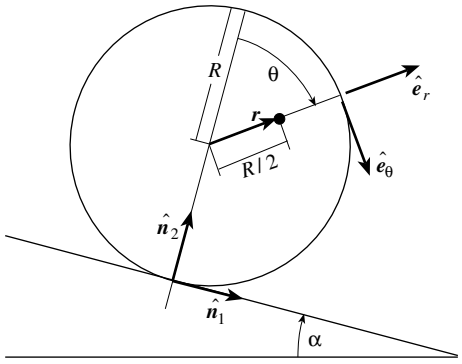


Figure 2.13: Rolling Cylinder with Offset mass

2.11 ♣ A ball m is freely rolling in the lower half of a sphere under the influence of a constant gravity field as shown in Figure 2.14. The sphere has a constant radius r. Assume that $\dot{\phi}(t_0)$ is zero and that $\theta(t_0)$, $\dot{\theta}(t_0)$ and $\phi(t_0)$ are given.

a) Find the equation of motion of the ball rolling without slip inside the sphere in terms of the spherical angle ϕ. *Hint: The angular momentum about the $\hat{\boldsymbol{n}}_3$ axis is conserved.*

b) What is the normal force that the wall of the sphere exerts onto the ball at any point in time?

c) Since $\dot{\phi}(t_0) = 0$, the ball is starting out on an extrema. Find an expression in terms of θ_0, $\dot{\theta}_0$ and $\dot{\phi}_0$ that determines the other motion extrema where $\dot{\phi} = 0$. *Hint: Use conservation of energy.*

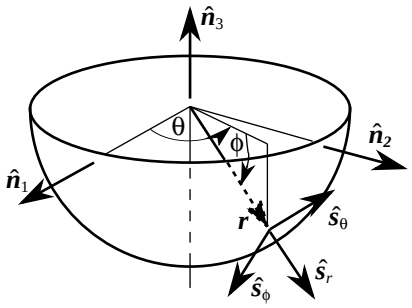


Figure 2.14: Ball rolling inside a Sphere

2.12 A cloud contains four particles with masses $m_1 = m_2 = 1$ and $m_3 = m_4 = 2$. The position vector of each particle is

$$\boldsymbol{R}_1 = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \quad \boldsymbol{R}_2 = \begin{pmatrix} -1 \\ -3 \\ 2 \end{pmatrix} \quad \boldsymbol{R}_3 = \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix} \quad \boldsymbol{R}_4 = \begin{pmatrix} 3 \\ -1 \\ -2 \end{pmatrix}$$

and their respective velocity vectors are

$$\dot{\boldsymbol{R}}_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \quad \dot{\boldsymbol{R}}_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \quad \dot{\boldsymbol{R}}_3 = \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix} \quad \dot{\boldsymbol{R}}_4 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

a) How much of the total cloud kinetic energy is translational kinetic energy and how much is rotation and deformation energy?

b) What is the cloud angular momentum vector about the origin and about the center of mass?

2.13 Two particles with mass $m/2$ are attached by a linear spring with a spring constant k as shown in Figure 2.15. Consider arbitrary initial position and velocity of each mass on the plane. For simplicity however, assume that the initial separation $2r_0$ is the unstretched length of the spring, and that the mass center has zero inertial velocity initially.

a) Determine the differential equations of motion whose solution would give $r(t)$ and $\theta(t)$ as functions of time and initial conditions; it is not necessary to solve these differential equations.

b) Determine an expression that relates the radial velocity $\dot{r}$ and the angular velocity $\dot{\theta}$ as functions of r, θ and initial conditions.

2.14 A particle of mass m is free to sling along a vertical ring as shown in Figure 2.16. The ring itself is rotating at a constant rate $\dot{\phi}$.

a) Determine the equations of motion of the particle in terms of θ.

b) What are the normal forces produced by the ring onto the particle?

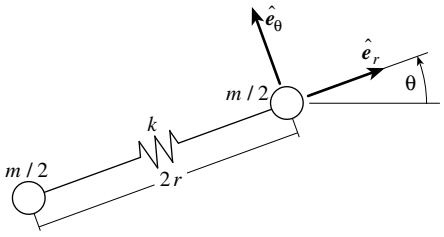


Figure 2.15: Two Masses Moving in a Plane

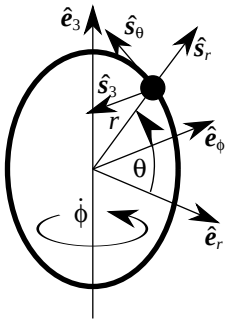


Figure 2.16: Particle Sliding Along a Rotating Ring

2.15 Newton's second Law for a particle of mass m states that $\boldsymbol{F} = d/dt(m\boldsymbol{v})$. If m is time varying, then one might expect $\boldsymbol{F} = \dot{m}\boldsymbol{v} + m\dot{\boldsymbol{v}}$ to be true. Explain why this logic is incorrect and does not lead to the correct rocket thrust equation.

2.16 The static thrust $\boldsymbol{F}_s$ of a rocket is given in Eq. (2.109). Draw a freebody diagram of a rocket engine test stand and verify that this is indeed that static force required to keep the rocket in place.

Bibliography

[1] Wiesel, W. E., *Spaceflight Dynamics*, McGraw-Hill, Inc., New York, 1989.

[2] Junkins, J. L. and Turner, J. D., *Optimal Spacecraft Rotational Maneuvers*, Elsevier Science Publishers, Amsterdam, Netherlands, 1986.

[3] Greenwood, D. T., *Principles of Dynamics*, Prentice-Hall, Inc, Englewood Cliffs, New Jersey, 2nd ed., 1988.

[4] Craig, R. R., *Structural Dynamics*, John Wiley & Sons, New York, 1981.

[5] Wertz, J. R. and Larson, W. J., *Space Mission Analysis and Design*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1991.

[6] Nelson, R. C., *Flight Stability and Automatic Control*, McGraw-Hill, Inc., New York, 1989.

Chapter Three

Rigid Body Kinematics

ATTITUDE coordinates (sometimes also referred to as attitude parameters) are sets of coordinates $\{x_1, x_2, \ldots, x_n\}$ that completely describe the orientation of a rigid body relative to some reference coordinate frame. There is an infinite number of attitude coordinates to choose from. Each set has strengths and weaknesses compared to other sets. This is analogous to choosing among the infinite sets of translational coordinates such as cartesian, polar or spherical coordinates to describe a spatial position of a point. However, describing the attitude of an object relative to some reference frame does differ in a fundamental way from describing the corresponding relative spatial position of a point. In cartesian space, the linear displacement between two spatial positions can grow arbitrarily large. On the other hand two rigid body (or coordinate frame) orientations can differ at most by a 180^o rotation, a finite rotational displacement. If an object rotates past 180^o, then its orientation actually starts to approach the starting angular position again. This concept of two orientations only being able to differ by finite rotations is important when designing control laws. A smart choice in attitude coordinates will be able to exploit this fact and produce a control law that is able to intelligently handle very large orientation errors.

The quest for "*the best rigid body orientation description*" is a very fundamental and important one. It has been studied by such great scholars as Euler, Jacobi, Hamilton, Cayley, Klein, Rodrigues and Gibbs and has led to a rich collection of elegant results. A good choice for attitude coordinates can greatly simplify the mathematics and avoid such pitfalls as mathematical and geometrical singularities or highly nonlinear kinematic differential equations. Among other things, a bad choice of attitude coordinates can artificially limit the operational range of a controlled system by requiring it to operate within the non-singular range of the chosen attitude parameters.

The following list contains four truths about rigid body attitude coordinates that are listed without proof.[1]

1. A minimum of three coordinates is required to describe the relative angular displacement between two reference frames $\mathcal{F}_1$ and $\mathcal{F}_2$.

2. Any minimal set of three attitude coordinates will contain at least one geometrical orientation where the coordinates are singular, namely at least two coordinates are undefined or not unique.

3. At or near such a geometric singularity, the corresponding kinematic differential equations are also singular.

4. The geometric singularities and associated numerical difficulties can be avoided altogether through a regularization.[2] Redundant sets of four or more coordinates exist which are universally determined and contain no geometric singularities.

3.1 Direction Cosine Matrix

Rigid body orientations are described using displacements of body-fixed referenced frames. The reference frame itself is usually defined using a set of three orthogonal, right-handed unit vectors. For notational purposes, a reference frame (or rigid body) is labeled through a script capital letter such as $\mathcal{F}$ and its associated unit base vectors are labeled with lower case letters such as $\hat{\boldsymbol{f}}_i$. There is always an infinity of ways to attach a reference frame to a rigid body. However, typically the reference frame base vectors are chosen such that they are aligned with the principal body axes.

Let the two reference frames $\mathcal{N}$ and $\mathcal{B}$ each be defined through sets of orthonormal right-handed sets of vectors $\{\hat{\boldsymbol{n}}\}$ and $\{\hat{\boldsymbol{b}}\}$ where we use the shorthand *vectrix* notation

$$\{\hat{\boldsymbol{n}}\} \equiv \begin{Bmatrix} \hat{\boldsymbol{n}}_1 \\ \hat{\boldsymbol{n}}_2 \\ \hat{\boldsymbol{n}}_2 \end{Bmatrix} \qquad \{\hat{\boldsymbol{b}}\} \equiv \begin{Bmatrix} \hat{\boldsymbol{b}}_1 \\ \hat{\boldsymbol{b}}_2 \\ \hat{\boldsymbol{b}}_2 \end{Bmatrix} \tag{3.1}$$

The sets of unit vectors are shown in Figure 3.1. The reference frame $\mathcal{B}$ can be thought of being a generic rigid body and the reference frame $\mathcal{N}$ could be associated with some particular inertial coordinate system. Let the three angles α_{1i} be the angles formed between the first body vector $\hat{\boldsymbol{b}}_1$ and the three inertial axes. The cosines of these angles are called the direction cosines of $\hat{\boldsymbol{b}}_1$ relative to the $\mathcal{N}$ frame. The unit vector $\hat{\boldsymbol{b}}_1$ can be projected onto $\{\hat{\boldsymbol{n}}\}$ as

$$\hat{\boldsymbol{b}}_1 = \cos\alpha_{11}\hat{\boldsymbol{n}}_1 + \cos\alpha_{12}\hat{\boldsymbol{n}}_2 + \cos\alpha_{13}\hat{\boldsymbol{n}}_3 \tag{3.2}$$

Clearly the direction cosines $\cos\alpha_{1j}$ are the three orthogonal components of $\hat{\boldsymbol{b}}_j$. Analogously, the direction angles α_{2i} and α_{3i} between the unit vectors $\hat{\boldsymbol{b}}_2$ and $\hat{\boldsymbol{b}}_3$ and the reference frame $\mathcal{N}$ base vectors can be found. These vectors are then expressed as

$$\hat{\boldsymbol{b}}_2 = \cos\alpha_{21}\hat{\boldsymbol{n}}_1 + \cos\alpha_{22}\hat{\boldsymbol{n}}_2 + \cos\alpha_{23}\hat{\boldsymbol{n}}_3 \tag{3.3}$$

$$\hat{\boldsymbol{b}}_3 = \cos\alpha_{31}\hat{\boldsymbol{n}}_1 + \cos\alpha_{32}\hat{\boldsymbol{n}}_2 + \cos\alpha_{33}\hat{\boldsymbol{n}}_3 \tag{3.4}$$

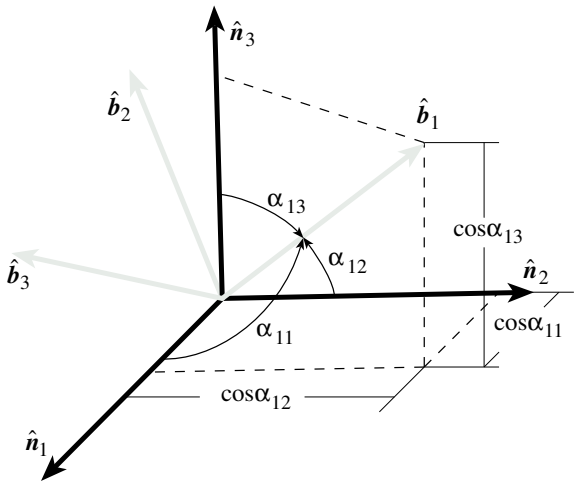


Figure 3.1: Direction Cosines

The set of orthonormal base vectors $\{\hat{\boldsymbol{b}}\}$ can be compactly expressed in terms of the base vectors $\{\hat{\boldsymbol{n}}\}$ as

$$\{\hat{\boldsymbol{b}}\} = \begin{bmatrix} \cos\alpha_{11} & \cos\alpha_{12} & \cos\alpha_{13} \\ \cos\alpha_{21} & \cos\alpha_{22} & \cos\alpha_{23} \\ \cos\alpha_{31} & \cos\alpha_{32} & \cos\alpha_{33} \end{bmatrix} \{\hat{\boldsymbol{n}}\} = [C]\{\hat{\boldsymbol{n}}\} \tag{3.5}$$

where the matrix $[C]$ is called the *direction cosine matrix*. Note that each entry of $[C]$ can be computed through

$$C_{ij} = \cos(\angle\ \hat{\boldsymbol{b}}_i, \hat{\boldsymbol{n}}_j) = \hat{\boldsymbol{b}}_i \cdot \hat{\boldsymbol{n}}_j \tag{3.6}$$

Analogously to Eq. (3.5), the set of $\{\hat{\boldsymbol{n}}\}$ vectors can be projected onto $\{\hat{\boldsymbol{b}}\}$ vectors as

$$\{\hat{\boldsymbol{n}}\} = \begin{bmatrix} \cos\alpha_{11} & \cos\alpha_{21} & \cos\alpha_{31} \\ \cos\alpha_{12} & \cos\alpha_{22} & \cos\alpha_{32} \\ \cos\alpha_{13} & \cos\alpha_{23} & \cos\alpha_{33} \end{bmatrix} \{\hat{\boldsymbol{b}}\} = [C]^T\{\hat{\boldsymbol{b}}\} \tag{3.7}$$

Substituting Eq. (3.7) into (3.5) yields

$$\{\hat{\boldsymbol{b}}\} = [C][C]^T\{\hat{\boldsymbol{b}}\} \tag{3.8}$$

which requires that

$$[C][C]^T = [I_{3\times 3}] \tag{3.9}$$

Similarly substituting Eq. (3.5) into (3.7) yields

$$[C]^T[C] = [I_{3\times 3}] \tag{3.10}$$

Eqs. (3.9) and (3.10) show that the direction cosine matrix $[C]$ is orthogonal.[1, 3–5] Therefore the inverse of $[C]$ is the transpose of $[C]$.

$$[C]^{-1} = [C]^T \tag{3.11}$$

Thanks to the orthogonality of the direction cosine matrix $[C]$, we will see below that the forward and inverse transformation (projection) of vectors between rotationally displaced reference frames can be accomplished without arithmetic.

Another important property of the direction cosine matrix is that its determinant is ± 1. This can be shown as follows.[5] From Eq. (3.9) it is evident that

$$\det\left(CC^T\right) = \det\left([I_{3\times 3}]\right) = 1 \tag{3.12}$$

Since $[C]$ is a square matrix this can be written as[6]

$$\det(C)\det(C^T) = 1 \tag{3.13}$$

Since $\det(C)$ is the same as $\det(C^T)$, this is further reduced to[7]

$$\left(\det(C)\right)^2 = 1 \Longleftrightarrow \det(C) = \pm 1 \tag{3.14}$$

As is shown by Goldstein in Ref. 8, if the reference frame base vectors $\{\hat{\boldsymbol{b}}\}$ and $\{\hat{\boldsymbol{n}}\}$ are right-handed, then $\det(C) = +1$. Goldstein also shows that the 3x3 direction cosine matrix $[C]$ will only have one real eigenvalue of ± 1. Again it will be $+1$ if the reference frame base vectors are right-handed.

In a standard coordinate transformation setting, the $[C]$ matrix is typically not restricted to projecting one set of base vectors from one reference frame onto another. Rather, the direction cosine's most powerful feature is the ability to directly project (or transform) an arbitrary vector, with components written in one reference frame, into a vector with components written in another reference frame. To show this let $\boldsymbol{v}$ be an arbitrary vector and let the reference frames $\mathcal{B}$ and $\mathcal{N}$ be defined as earlier. Let the scalars v_{b_i} be the vector components of $\boldsymbol{v}$ in the $\mathcal{B}$ reference frame.

$$\boldsymbol{v} = v_{b_1}\hat{\boldsymbol{b}}_1 + v_{b_2}\hat{\boldsymbol{b}}_2 + v_{b_3}\hat{\boldsymbol{b}}_3 = \{v_b\}^T\{\hat{\boldsymbol{b}}\} \tag{3.15}$$

Similarly $\boldsymbol{v}$ can be written in terms of $\mathcal{N}$ frame components v_{n_i} as

$$\boldsymbol{v} = v_{n_1}\hat{\boldsymbol{n}}_1 + v_{n_2}\hat{\boldsymbol{n}}_2 + v_{n_3}\hat{\boldsymbol{n}}_3 = \{v_n\}^T\{\hat{\boldsymbol{n}}\} \tag{3.16}$$

Substituting Eq. (3.7) into Eq. (3.16) the $\boldsymbol{v}$ vector components in the $\mathcal{N}$ frame can be directly projected into the $\mathcal{B}$ frame.

$$\boldsymbol{v}_b = [C]\boldsymbol{v}_n \tag{3.17}$$

Since the inverse of $[C]$ is simply $[C]^T$, the inverse transformation is

$$\boldsymbol{v}_n = [C]^T\boldsymbol{v}_b \tag{3.18}$$

The fact that Eqs. (3.17) and (3.18) are exactly analogous to Eqs. (3.5) and (3.7) is a fundamental property of Gibbsian vectors, and more generally, cartesian tensors.

Another common problem is that several cascading reference frames are present where each reference frame orientation is defined relative to the previous one, and it is desired to replace the sequence of projections by a single projection. Let $\{\hat{\boldsymbol{r}}\}$ contain the base vectors of the reference frame $\mathcal{R}$ whose relative orientation to the $\mathcal{B}$ frame is given through $[C']$.

$$\{\hat{\boldsymbol{r}}\} = [C']\{\hat{\boldsymbol{b}}\} \tag{3.19}$$

The basis vectors $\{\hat{\boldsymbol{n}}\}$ in the $\mathcal{N}$ frame can be projected directly into the $\mathcal{R}$ frame through

$$\{\hat{\boldsymbol{r}}\} = [C']\{C\}\{\hat{\boldsymbol{n}}\} = [C'']\{\hat{\boldsymbol{n}}\} \tag{3.20}$$

where the direction cosine matrix $[C''] = [C'][C]$ projects vectors in the $\mathcal{N}$ frame to vectors in the $\mathcal{R}$ frame. The direct transformation matrix from the first to the last cascading reference frame is clearly found by successive matrix-multiplications of each relative transformation matrix in reverse order as shown above. This property $[C''] = [C'][C]$ for composition of successive rotations is very important. When rotational coordinates are introduced to parameterize the $[C]$ matrix, the corresponding "composition" relationship among the three sets of coordinates is also of fundamental importance.

The direction cosine matrix is the most fundamental, but highly redundant, method of describing a relative orientation. As was mentioned earlier, the minimum number of parameter required to describe a reference frame orientation is three. The direction cosine matrix has nine entries. The six extra parameters in the matrix are made redundant through the orthogonality condition $[C][C]^T = [I_{3\times 3}]$. This is why in practice the elements of the direction cosine matrix are rarely used as coordinates to keep track of an orientation; instead less redundant attitude parameters are used. The biggest asset of the direction cosine matrix is the ability to easily transform vectors from one reference frame to another.

Example 3.1: Let the two reference frames $\mathcal{B}$ and $\mathcal{F}$ be defined relative to the inertial reference frame $\mathcal{N}$ by the orthonormal unit base vectors

$$\hat{\boldsymbol{b}}_1 = \big(0,1,0\big)^T \qquad \hat{\boldsymbol{b}}_2 = \big(1,0,0\big)^T \qquad \hat{\boldsymbol{b}}_3 = \big(0,0,-1\big)^T$$
$$\hat{\boldsymbol{f}}_1 = \Big(\tfrac{1}{2},\tfrac{\sqrt{3}}{2},0\Big)^T \qquad \hat{\boldsymbol{f}}_2 = \big(0,0,1\big)^T \qquad \hat{\boldsymbol{f}}_3 = \Big(\tfrac{\sqrt{3}}{2},-\tfrac{1}{2},0\Big)^T$$

where the $\hat{\boldsymbol{b}}_i$ and $\hat{\boldsymbol{f}}_i$ vector components are written in the inertial $\mathcal{N}$ frame. Let us use the following notation to label the various direction cosine matrices. The matrix $[BN]$ maps vectors written in the $\mathcal{N}$ frame into vectors written in the $\mathcal{B}$ frame. Analogously, the matrix $[FB]$ maps vectors in the $\mathcal{B}$ frame

into $\mathcal{F}$ frame vectors and so on. To find the entries of the various relative rotation matrices, note the following useful identity.

$$[FB]_{ij} = \cos\alpha_{ij} = \hat{\boldsymbol{f}}_i \cdot \hat{\boldsymbol{b}}_j$$

Given the base vectors of each frame, it is not necessary to find the angles between each set of vectors to find the appropriate direction cosine matrix. Since all base vectors have unit length, the inner product of the corresponding vectors will provide the needed direction cosines. The rotation matrices $[BN]_{ij} = \hat{\boldsymbol{b}}_i \cdot \hat{\boldsymbol{n}}_j$, $[FN]_{ij} = \hat{\boldsymbol{f}}_i \cdot \hat{\boldsymbol{n}}_j$ and $[FB]_{ij} = \hat{\boldsymbol{f}}_i \cdot \hat{\boldsymbol{b}}_j$ are

$$[BN] = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad [FN] = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 1 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \end{bmatrix}$$

$$[FB] = \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ 0 & 0 & -1 \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \end{bmatrix}$$

Instead of calculating the rotation matrix $[FB]$ from dot products of the respective base vectors, it could also be calculated using Eq. (3.20).

$$[FB] = [FN][BN]^T = \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ 0 & 0 & -1 \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \end{bmatrix} \quad \surd$$

To find the kinematic differential equation in terms of the direction cosine matrix $[C]$, let us write the instantaneous angular velocity vector $\boldsymbol{\omega}$ of the $\mathcal{B}$ frame relative to the $\mathcal{N}$ frame in $\mathcal{B}$ frame orthogonal components as

$$\boldsymbol{\omega} = \omega_1 \hat{\boldsymbol{b}}_1 + \omega_2 \hat{\boldsymbol{b}}_2 + \omega_3 \hat{\boldsymbol{b}}_3 \tag{3.21}$$

Let ${}^{\mathcal{N}}d/dt\{\hat{\boldsymbol{b}}\}$ be the derivative of the $\mathcal{B}$ frame base vectors taken in the $\mathcal{N}$ frame. Using the transport theorem we find[9]

$$\frac{{}^{\mathcal{N}}d}{dt}\{\hat{\boldsymbol{b}}_i\} = \frac{{}^{\mathcal{B}}d}{dt}\{\hat{\boldsymbol{b}}_i\} + \boldsymbol{\omega} \times \{\hat{\boldsymbol{b}}_i\} \tag{3.22}$$

Since the $\mathcal{B}$ frame base vectors are fixed within their frame the expression ${}^{\mathcal{B}}d/dt\{\hat{\boldsymbol{b}}\}$ is zero. After introducing the skew-symmetric tilde matrix operator

$$[\tilde{\boldsymbol{x}}] = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix} \tag{3.23}$$

Eq. (3.22) leads to the vectrix equation

$$\frac{{}^{\mathcal{N}}d}{dt}\{\hat{\boldsymbol{b}}\} = -[\tilde{\boldsymbol{\omega}}]\{\hat{\boldsymbol{b}}\} \tag{3.24}$$

Taking the time derivative of the right hand side of Eq. (3.5) we find

$$\frac{\mathcal{N}d}{dt}\left([C]\{\hat{\boldsymbol{n}}\}\right) = \frac{d}{dt}\left([C]\right)\{\hat{\boldsymbol{n}}\} + [C]\frac{\mathcal{N}d}{dt}\left(\{\hat{\boldsymbol{n}}\}\right) = [\dot{C}]\{\hat{\boldsymbol{n}}\} \tag{3.25}$$

where the short hand notation $d/dt([C]) = [\dot{C}]$ is used. Using Eq. (3.5), Eqs. (3.24) and (3.25) are combined to

$$\left([\dot{C}] + [\tilde{\boldsymbol{\omega}}][C]\right)\{\hat{\boldsymbol{n}}\} = 0 \tag{3.26}$$

Since Eq. (3.26) must hold for any set of $\{\hat{\boldsymbol{n}}\}$, the kinematic differential equation satisfied by the direction cosine matrix $[C]$ is found to be[1, 10]

$$[\dot{C}] = -[\tilde{\boldsymbol{\omega}}][C] \tag{3.27}$$

It can easily be verified that Eq. (3.9) is indeed an exact solution of above differential equation. Take the derivative of $[C][C]^T$

$$\frac{d}{dt}\left([C][C]^T\right) = [\dot{C}][C]^T + [C][\dot{C}]^T \tag{3.28}$$

and then substitute Eq. (3.27) to obtain

$$\frac{d}{dt}\left([C][C]^T\right) = -[\tilde{\boldsymbol{\omega}}][C][C]^T - [C][C]^T[\tilde{\boldsymbol{\omega}}]^T \tag{3.29}$$

Making use of the orthogonality of $[C]$ and since $[\tilde{\boldsymbol{\omega}}] = -[\tilde{\boldsymbol{\omega}}]^T$ is skew-symmetric this simplifies to

$$\frac{d}{dt}\left([C][C]^T\right) = -[\tilde{\boldsymbol{\omega}}] + [\tilde{\boldsymbol{\omega}}] = 0 \tag{3.30}$$

Since $[C][C]^T$ is a constant solution of the differential equation in Eq. (3.27), and Eq. (3.9) is satisfied initially, the solution of Eq. (3.27) will theoretically satisfy the orthogonality condition for all time. In practice, numerical solutions of Eq. (3.27) will slowly accumulate arithmetic errors so that the orthogonality condition $[C][C]^T - [I_{3\times 3}] = 0$ is slightly in error. There are several ways to resolve this minor difficulty.

Given an arbitrary time history of $\boldsymbol{\omega}(t)$, Eq. (3.27) represents a rigorously linear differential equation which can be integrated to yield the instantaneous direction cosine matrix $[C]$. A major advantage of the kinematic differential equation for $[C]$ is that it is linear and universally applicable. There are no geometric singularities present in the attitude description or its kinematic differential equations. However, this advantage comes at the cost of having a highly redundant formulation. Several other attitude parameters will be presented in the following sections which include a minimal number (3) of attitude parameters. However, all minimal sets of attitude coordinates have kinematic differential equations which contain some degree of nonlinearity and also embody geometric and/or mathematical singularities. Only the once redundant Euler parameters (quaternions) will be found to retain a singularity free description and possess linear kinematic differential equations analogous to the direction cosine matrix.

3.2 Euler Angles

The most commonly used sets of attitude parameters are the Euler angles. They describe the attitude of a reference frame $\mathcal{B}$ relative to the frame $\mathcal{N}$ through three successive rotation angles $(\theta_1, \theta_2, \theta_3)$ about the sequentially displaced body fixed axes $\{\hat{\boldsymbol{b}}\}$. Note that the order of the axes about which the reference frame is rotated is important here. Performing three successive rotations about the 3rd, 2nd and 1st body axis, labeled (3-2-1) for short, does not yield the same orientation as if instead the rotation order is (1-2-3). Note these sequential rotations provide an instantaneous geometrical recipe for $\mathcal{N}$. Clearly, for $\mathcal{B}$ undergoing general motion, the $\theta_i(t)$ are time varying in a general way.

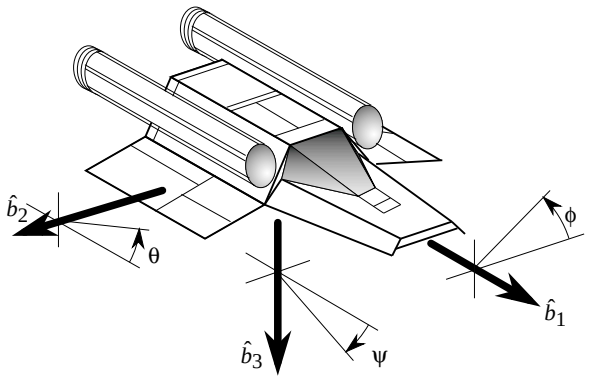


Figure 3.2: Yaw, Pitch and Roll Euler Angles

Aircraft and spacecraft orientations are commonly described through the Euler angles yaw, pitch and roll (ψ,θ,ϕ) as shown in Figure 3.2. They are usually measured relative to axes associated with a nominal flight path. The position of $\{\hat{\boldsymbol{b}}\}$ relative to $\{\hat{\boldsymbol{n}}\}$ is *described* by a sequence of three rigid rotations about prescribed body fixed axes. While the conceptual description is a sequence of rotations, we can consider the *instantaneous* values of these three angles and thereby establish a means for describing general, non-sequential rotations. The popularity of Euler angles stems from the fact that the relative attitude is easy to visualize for small angles. To transform components of a vector in the $\mathcal{N}$ frame into the $\mathcal{B}$ frame through a sequence of Euler angle rotations, the reference axes are first rotated about the $\hat{\boldsymbol{b}}_3$ axis by the yaw angle ψ, then about the $\hat{\boldsymbol{b}}_2$ axis by the pitch angle θ and finally about the $\hat{\boldsymbol{b}}_1$ axis by the roll angle ϕ as is shown in Figure 3.3. Thus the standard yaw-pitch-roll (ψ,θ,ϕ) angles are the (3-2-1) set of Euler angles.[11]

Another very popular set of Euler angles is the (3-1-3) set of Euler angles. These angles are commonly used by astronomers to define the orientation of orbit planes of the planets relative to the Earth's orbit plane.[1] While the (3-2-1) Euler angles are considered an *asymmetric* set, the (3-1-3) Euler angles are

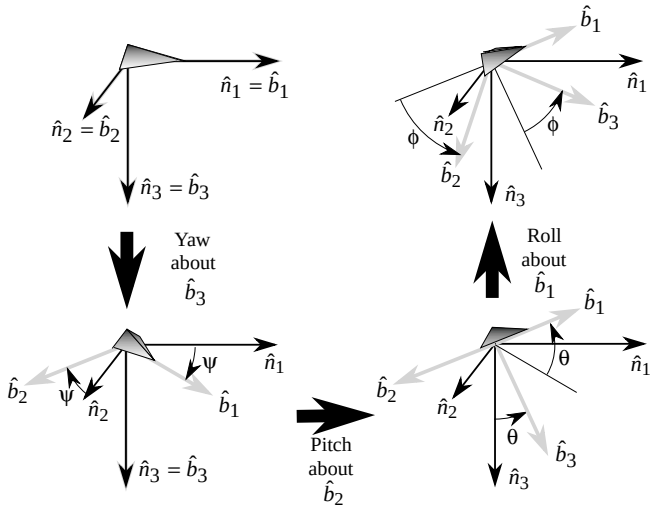


Figure 3.3: Successive Yaw, Pitch and Roll Rotations

a *symmetric* set since two rotations about the third body axis are performed. Instead of being called yaw, pitch and roll angles, the (3-1-3) Euler angles are called longitude of the ascending node Ω, inclination i and argument of the perihelion ω and are illustrated in Figure 3.4 below.[1, 12]

The direction cosine matrix introduced in section 3.1 can be parameterized in terms of the Euler angles. Since each Euler angle defines a successive rotation about the i-th body axis, let the three single-axis rotation matrices $[M_i(\theta)]$ be defined as

$$[M_1(\theta)] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix} \tag{3.31a}$$

$$[M_2(\theta)] = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix} \tag{3.31b}$$

$$[M_3(\theta)] = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.31c}$$

Let the (α, β, γ) Euler angle sequence be $(\theta_1, \theta_2, \theta_3)$. Using Eq. (3.20) to combine successive rotations, the direction cosine matrix in terms of the (α, β, γ) Euler angles can be written as[1]

$$[C(\theta_1, \theta_2, \theta_3)] = [M_\gamma(\theta_3)][M_\beta(\theta_2)][M_\alpha(\theta_1)] \tag{3.32}$$

In particular, the direction cosine matrix in terms of the (3-2-1) Euler angles $(\theta_1, \theta_2, \theta_3) = (\psi, \theta, \phi)$ is[11]

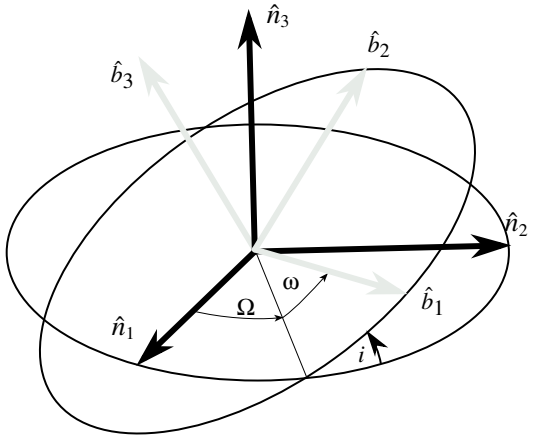


Figure 3.4: (3-1-3) Euler Angle Illustration

n particular, the direction cosine matrix in terms of the (3-2-1) Euler n particular, the direction cosine matrix in terms of the (3-2-1) Euler n particular, the direction cosine matrix in terms of the (3-2-1) Euler n particular, the direction cosine matrix in terms of the (3-2-1) Euler n particular, the direction cosine matrix in terms of the (3-2-1) Euler n particular, the direction cosine matrix in terms of the (3-2-1) Euler n particular, the direction cosine matrix in terms of the (3-2-1) Euler n particular, the direction cosine matrix in terms of the (3-2-1) Euler

$$[C] = \begin{bmatrix} c\theta_2 c\theta_1 & c\theta_2 s\theta_1 & -s\theta_2 \\ s\theta_3 s\theta_2 c\theta_1 - c\theta_3 s\theta_1 & s\theta_3 s\theta_2 s\theta_1 + c\theta_3 c\theta_1 & s\theta_3 c\theta_2 \\ c\theta_3 s\theta_2 c\theta_1 + s\theta_3 s\theta_1 & c\theta_3 s\theta_2 s\theta_1 - s\theta_3 c\theta_1 & c\theta_3 c\theta_2 \end{bmatrix} \tag{3.33}$$

where the short hand notation $c\xi = \cos\xi$ and $s\xi = \sin\xi$ is used. The inverse transformations from the direction cosine matrix $[C]$ to the (ψ, θ, ϕ) angles are

$$\psi = \theta_1 = \tan^{-1}\left(\frac{C_{12}}{C_{11}}\right) \tag{3.34a}$$

$$\theta = \theta_2 = -\sin^{-1}(C_{13}) \tag{3.34b}$$

$$\phi = \theta_3 = \tan^{-1}\left(\frac{C_{23}}{C_{33}}\right) \tag{3.34c}$$

In terms of the (3-1-3) Euler angles $(\theta_1, \theta_2, \theta_3) = (\Omega, i, \omega)$ the direction cosine matrix $[C]$ is written as[1]

$$[C] = \begin{bmatrix} c\theta_3 c\theta_1 - s\theta_3 c\theta_2 s\theta_1 & c\theta_3 s\theta_1 + s\theta_3 c\theta_2 c\theta_1 & s\theta_3 s\theta_2 \\ -s\theta_3 c\theta_1 - c\theta_3 c\theta_2 s\theta_1 & -s\theta_3 s\theta_1 + c\theta_3 c\theta_2 c\theta_1 & c\theta_3 s\theta_2 \\ s\theta_2 s\theta_1 & -s\theta_2 c\theta_1 & c\theta_2 \end{bmatrix} \tag{3.35}$$

The inverse transformations from the direction cosine matrix $[C]$ to the (3-1-3) Euler angles (Ω, i, ω) are

$$\Omega = \theta_1 = \tan^{-1}\left(\frac{C_{31}}{-C_{32}}\right) \tag{3.36a}$$

$$i = \theta_2 = \cos^{-1}(C_{33}) \tag{3.36b}$$

$$\omega = \theta_3 = \tan^{-1}\left(\frac{C_{13}}{C_{23}}\right) \tag{3.36c}$$

The complete set of 12 transformations between the various Euler angle sets and the direction cosine matrix can be found in the Appendix C. We emphasize that while Eqs. (3.32)-(3.36) are easily established by sequential angular displacements, we consider the inverse situation; given a generally varying $[C]$ matrix, we can consider equations such as Eqs. (3.32)-(3.36) to hold at any instant in the motion, and thus $\{\psi(t), \theta(t), \phi(t)\}$ or $\{\Omega(t), i(t), \omega(t)\}$ can be considered as candidate coordinates for general rotational motion.

Note that each of the 12 possible sets of Euler angles has a geometric singularity where two angles are not uniquely defined. For the (3-2-1) Euler angles pitching up or down 90 degrees results in a geometric singularity. If the pitch angle is $\pm$90 degrees, then it does not matter if ψ=0 and ϕ=10 degrees or ψ = 10 and $\phi = 0$ degrees. Only the sum $\psi + \phi$ is unique in this case. For the (3-1-3) Euler angles the geometric singularity occurs for an inclination angle of zero or 180 degrees. This geometric singularity also manifests itself in a mathematical singularity of the corresponding Euler angle kinematic differential equation.

Let $\boldsymbol{\theta} = \{\theta_1, \theta_2, \theta_3\}$ and $\boldsymbol{\phi} = \{\phi_1, \phi_2, \phi_3\}$ be two Euler angle vectors with identical rotation sequences. Often it is necessary to find the attitude that corresponds to performing two *successive* rotations, i.e. "adding" the two rotations. If a rigid body first performs the rotation $\boldsymbol{\theta}$ and then the rotation $\boldsymbol{\phi}$, then the final attitude is expressed relative to the original attitude through the vector $\boldsymbol{\varphi} = \{\varphi_1, \varphi_2, \varphi_3\}$ defined through

$$[FN(\boldsymbol{\varphi})] = [FB(\boldsymbol{\phi})][BN(\boldsymbol{\theta})] \tag{3.37}$$

Eq. (3.37) could be used to solve for $\boldsymbol{\varphi}$ in terms of the vector components of $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$. This process is very tedious and typically does not provide any simple, compact final expressions. However, for the case where $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ are vectors of symmetric Euler angles, then it is possible to obtain relatively compact transformations from the first two vectors into the overall vector using spherical geometry relationships.[5, 13]

A sample spherical triangle is shown in Figure 3.5. The following two spherical triangle laws are the only two required in deriving the symmetrical Euler angle successive rotation property. The *spherical law of sines* states that

$$\frac{\sin A}{\sin a} = \frac{\sin B}{\sin b} = \frac{\sin C}{\sin c} \tag{3.38}$$

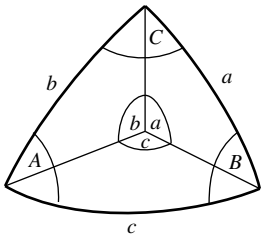
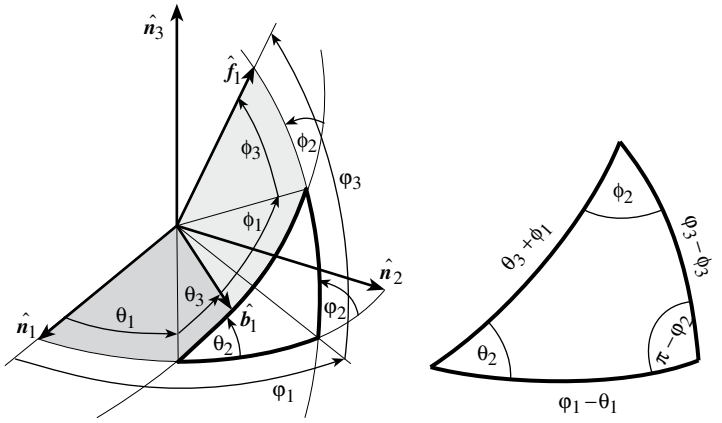


Figure 3.5: Spherical Triangle Labels



(i) Successive (3-1-3) Euler Angles **(ii)** Spherical Triangle

Figure 3.6: Illustration of Successive (3-1-3) Euler Angle Rotations

and the *spherical law of cosines* states that

$$\cos A = -\cos B \cos C + \sin B \sin C \cos a \tag{3.39a}$$
$$\cos B = -\cos A \cos C + \sin A \sin C \cos b \tag{3.39b}$$
$$\cos C = -\cos A \cos B + \sin A \sin B \cos c \tag{3.39c}$$

Figure 3.6(i) illustrates the orientation of the first body axis as it is first rotated from $\mathcal{N}$ to $\mathcal{B}$ with the (3-1-3) Euler angle vector $\boldsymbol{\theta}$ and then from $\mathcal{B}$ to $\mathcal{F}$ with the (3-1-3) vector $\boldsymbol{\phi}$. The (3-1-3) Euler angle description of the direct rotation from $\mathcal{N}$ to $\mathcal{F}$ is clearly given by the angles φ_1, φ_2 and φ_3. To obtain direct transformations from $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ to $\boldsymbol{\varphi}$, the bold spherical triangle in Figure 3.6(i) is used. The spherical arc lengths and angles of this triangle are labeled in Figure 3.6(ii). Using the spherical law of cosines we find that

$$\cos(\pi - \varphi_2) = -\cos\theta_2 \cos\phi_2 + \sin\theta_2 \sin\phi_2 \cos(\theta_3 + \phi_1) \tag{3.40}$$

This is trivially solved for the angle φ_2 as

$$\varphi_2 = \cos^{-1}\left(\cos\theta_2\cos\phi_2 - \sin\theta_2\sin\phi_2\cos(\theta_3+\phi_1)\right) \tag{3.41}$$

Using the spherical laws of sines, we are able to find the following expressions for φ_1 and φ_3:

$$\sin(\varphi_1-\theta_1) = \frac{\sin\phi_2}{\sin\varphi_2}\sin(\theta_3+\phi_1) \tag{3.42}$$

$$\sin(\varphi_3-\phi_3) = \frac{\sin\theta_2}{\sin\varphi_2}\sin(\theta_3+\phi_1) \tag{3.43}$$

To avoid quadrant problems, we prefer to find expressions of φ_1 and φ_3 that involve the tan function instead of the sin function. To accomplish this, using the spherical law of cosines we find the following two relationships:

$$\cos(\varphi_1-\theta_1) = \frac{\cos\phi_2 - \cos\theta_2\cos\varphi_2}{\sin\theta_2\sin\varphi_2} \tag{3.44}$$

$$\cos(\varphi_3-\phi_3) = \frac{\cos\theta_2 - \cos\phi_2\cos\varphi_2}{\sin\phi_2\sin\varphi_2} \tag{3.45}$$

Combining Eqs. (3.42) through (3.45), we are able to solve for φ_1 and φ_3 using the inverse tan function.

$$\varphi_1 = \theta_1 + \tan^{-1}\left(\frac{\sin\theta_2\sin\phi_2\sin(\theta_3+\phi_1)}{\cos\phi_2 - \cos\theta_2\cos\varphi_2}\right) \tag{3.46}$$

$$\varphi_3 = \phi_3 + \tan^{-1}\left(\frac{\sin\theta_2\sin\phi_2\sin(\theta_3+\phi_1)}{\cos\theta_2 - \cos\phi_2\cos\varphi_2}\right) \tag{3.47}$$

Using Eqs. (3.41), (3.46) and (3.47) to solve for $\boldsymbol{\phi}$ instead of back-solving $\boldsymbol{\varphi}$ out of the direction cosine matrix in Eq. (3.37) is numerically more efficient. While the Euler angle successive or composite rotation was developed for the (3-1-3) special case, the transformations in Eqs. (3.41), (3.46) and (3.47) actually hold for *any* symmetric rotation sequence.[5, 13] Asymmetric sets, however, will have to be composited using the corresponding direction cosine matrices.

On occasion it is required to find the *relative* attitude vector between two reference frame. For example, given the symmetric Euler angle vectors $\boldsymbol{\theta}$ and $\boldsymbol{\varphi}$, find the corresponding vector $\boldsymbol{\phi}$ which relates $\mathcal{B}$ to $\mathcal{F}$. Using the same spherical triangle in Figure 3.6(ii), we find the following closed form expressions for $\boldsymbol{\phi}$.

$$\phi_1 = -\theta_3 + \tan^{-1}\left(\frac{\sin\theta_2\sin\varphi_2\sin(\varphi_1-\theta_1)}{\cos\theta_2\cos\phi_2 - \cos\varphi_2}\right) \tag{3.48}$$

$$\phi_2 = \cos^{-1}\left(\cos\theta_2\cos\varphi_2 + \sin\theta_2\sin\varphi_2\cos(\varphi_1-\theta_1)\right) \tag{3.49}$$

$$\phi_3 = \varphi_3 - \tan^{-1}\left(\frac{\sin\theta_2\sin\varphi_2\sin(\varphi_1-\theta_1)}{\cos\theta_2 - \cos\phi_2\cos\varphi_2}\right) \tag{3.50}$$

Similar expressions can be found to express $\boldsymbol{\theta}$ in terms of $\boldsymbol{\phi}$ and $\boldsymbol{\varphi}$.

Example 3.2: Let the orientations of two spacecraft $\mathcal{B}$ and $\mathcal{F}$ relative to an inertial frame $\mathcal{N}$ be given through the asymmetric (3-2-1) Euler angles $\boldsymbol{\theta}_{\mathcal{B}} = (30, -45, 60)^T$ and $\boldsymbol{\theta}_{\mathcal{F}} = (10, 25, -15)^T$ degrees. What is the relative orientation of spacecraft $\mathcal{B}$ relative to $\mathcal{F}$ in terms of (3-2-1) Euler angles.

The orientation matrices $[BN]$ and $[FN]$ are found using Eq. (3.33).

$$[BN] = \begin{bmatrix} 0.612372 & 0.353553 & 0.707107 \\ -0.78033 & 0.126826 & 0.612372 \\ 0.126826 & -0.926777 & 0.353553 \end{bmatrix}$$

$$[FN] = \begin{bmatrix} 0.892539 & 0.157379 & -0.422618 \\ -0.275451 & 0.932257 & -0.234570 \\ 0.357073 & 0.325773 & 0.875426 \end{bmatrix}$$

The direction cosine matrix $[BF]$ which describes the attitude of $\mathcal{B}$ relative to $\mathcal{F}$ is computed by using Eq. (3.20).

$$[BF] = [BN][FN]^T = \begin{bmatrix} 0.303372 & -0.0049418 & 0.952859 \\ -0.935315 & 0.1895340 & 0.298769 \\ -0.182075 & -0.9818620 & 0.052877 \end{bmatrix}$$

Using the transformations in Eq. (3.34) the relative (3-2-1) Euler angles are

$$\psi = \tan^{-1}\left(\frac{-0.0049418}{0.303372}\right) = 0.933242 \ deg$$

$$\theta = -\sin^{-1}(0.952859) = -1.26252 \ deg$$

$$\phi = \tan^{-1}\left(\frac{0.298769}{0.052877}\right) = -57.6097 \ deg$$

Since ϕ is much larger than ψ and θ, the attitude of $\mathcal{B}$ could be described qualitatively to differ from $\mathcal{F}$ by a -57.6 degree roll. This result was not immediately obvious studying the original Euler angle vectors $\boldsymbol{\theta}_{\mathcal{B}}$ and $\boldsymbol{\theta}_{\mathcal{F}}$.

Let the vector $\boldsymbol{\omega}$ define the instantaneous rotational velocity of the $\mathcal{B}$ frame relative to the $\mathcal{N}$ frame. To avoid having to integrate the direction cosine matrix directly given an $\boldsymbol{\omega}$ time history, the Euler angle kinematic differential equations are needed. The (3-2-1) Euler kinematic differential equation is derived below. The methodology can be used for any set of Euler angles. The vector $\boldsymbol{\omega}$ is written in body frame components as

$$\boldsymbol{\omega} = \omega_1 \hat{\boldsymbol{b}}_1 + \omega_2 \hat{\boldsymbol{b}}_2 + \omega_3 \hat{\boldsymbol{b}}_3 \tag{3.51}$$

From Figure 3.3 it is evident that the $\mathcal{B}$ frame rotation can also be written in terms of the Euler angle rates $(\dot{\psi}, \dot{\theta}, \dot{\phi})$ as

$$\boldsymbol{\omega} = \dot{\psi}\hat{\boldsymbol{n}}_3 + \dot{\theta}\hat{\boldsymbol{b}}'_2 + \dot{\phi}\hat{\boldsymbol{b}}_1 \tag{3.52}$$

The unit vector $\hat{\boldsymbol{b}}_2'$ is the direction of the body fixed axis $\hat{\boldsymbol{b}}_2$ before performing a roll ϕ about $\hat{\boldsymbol{b}}_1$ as is shown in Figure 3.3. It can be written in terms of $\{\hat{\boldsymbol{b}}\}$ as

$$\hat{\boldsymbol{b}}_2' = \cos\phi\hat{\boldsymbol{b}}_2 - \sin\phi\hat{\boldsymbol{b}}_3 \tag{3.53}$$

The direction cosine matrix in terms of the (3-2-1) Euler angles in Eq. (3.33) is used to express $\hat{\boldsymbol{n}}_3$ in terms of $\{\hat{\boldsymbol{b}}\}$.

$$\hat{\boldsymbol{n}}_3 = -\sin\theta\hat{\boldsymbol{b}}_1 + \sin\phi\cos\theta\hat{\boldsymbol{b}}_2 + \cos\phi\cos\theta\hat{\boldsymbol{b}}_3 \tag{3.54}$$

After substituting Eqs. (3.53) and (3.54) into Eq. (3.52) and then comparing terms with Eq. (3.51), the following kinematic equation is found.

$$\begin{pmatrix}\omega_1\\ \omega_2\\ \omega_3\end{pmatrix} = \begin{bmatrix}-\sin\theta & 0 & 1\\ \sin\phi\cos\theta & \cos\phi & 0\\ \cos\phi\cos\theta & -\sin\phi & 0\end{bmatrix}\begin{pmatrix}\dot{\psi}\\ \dot{\theta}\\ \dot{\phi}\end{pmatrix} \tag{3.55}$$

The kinematic differential equation of the (3-2-1) Euler angles is the inverse of Eq. (3.55).

$$\begin{pmatrix}\dot{\psi}\\ \dot{\theta}\\ \dot{\phi}\end{pmatrix} = \frac{1}{\cos\theta}\begin{bmatrix}0 & \sin\phi & \cos\phi\\ 0 & \cos\phi\cos\theta & -\sin\phi\cos\theta\\ \cos\theta & \sin\phi\sin\theta & \cos\phi\sin\theta\end{bmatrix}\begin{pmatrix}\omega_1\\ \omega_2\\ \omega_3\end{pmatrix} = [B(\psi,\theta,\phi)]\boldsymbol{\omega} \tag{3.56}$$

Similarly, the kinematic differential equations for the (3-1-3) Euler angles are found be

$$\boldsymbol{\omega} = \begin{bmatrix}\sin\theta_3\sin\theta_2 & \cos\theta_3 & 0\\ \cos\theta_3\sin\theta_2 & -\sin\theta_3 & 0\\ \cos\theta_2 & 0 & 1\end{bmatrix}\begin{pmatrix}\dot{\theta}_1\\ \dot{\theta}_2\\ \dot{\theta}_3\end{pmatrix} \tag{3.57}$$

with the inverse relationship

$$\begin{pmatrix}\dot{\theta}_1\\ \dot{\theta}_2\\ \dot{\theta}_3\end{pmatrix} = \frac{1}{\sin\theta_2}\begin{bmatrix}\sin\theta_3 & \cos\theta_3 & 0\\ \cos\theta_3\sin\theta_2 & -\sin\theta_3\sin\theta_2 & 0\\ -\sin\theta_3\cos\theta_2 & -\cos\theta_3\cos\theta_2 & \sin\theta_2\end{bmatrix}\boldsymbol{\omega} = [B(\boldsymbol{\theta})]\boldsymbol{\omega} \tag{3.58}$$

The complete set of 12 transformations between the various Euler angle rates and the body angular velocity vector can be found in Appendix C. Note that the Euler angle kinematic differential equations encounter a singularity either at $\theta_2 = \pm 90$ degrees for the (3-2-1) set or at $\theta_2 = 0$ or 180 degrees for the (3-1-3) set. It turns out that all Euler angles sets encounter a singularity at specific second rotation angle θ_2 only. The first and third rotation angles θ_1 and θ_3 never lead to a singularity. In all cases, it can be verified that the singularity occurs for those θ_2 values that result in θ_1 and θ_3 being measured in the same plane. If the Euler angle set is symmetric, then the singular orientation is at $\theta_2 = 0$ or 180 degrees. If the Euler angle set is asymmetric, then the singular

orientation is $\theta_2 = \pm 90$ degrees. Therefore asymmetric sets such as the (3-2-1) Euler reference frame. Symmetric sets as the (3-1-3) Euler angles would not be convenient to describe small departure rotations of $\{\hat{\boldsymbol{b}}\}$ from the $\{\hat{\boldsymbol{n}}\}$ axes since for small angles one would always operate very close to the singular attitude at $\theta_2 = 0$.

The Euler angles provide a compact, three parameter attitude description whose coordinates are easy to visualize. One main drawback of these angles is that a rigid body or reference frame is never further than a 90 degree rotation away from a singular orientation. Therefore their use in describing large, arbitrary and especially arbitrary rotations is limited. Also, their kinematic differential equations are fairly nonlinear, containing computationally intensive trigonometric functions. The linearized Euler angle kinematic differential equations are only valid for a relatively small domain of rotations.

3.3 Principal Rotation Vector

The following theorem has been very fundamental in the development of several types of attitude coordinates and is generally referenced to Euler.[14–16]

Theorem 3.1 (Euler's Principal Rotation) *A rigid body or coordinate reference frame can be brought from an arbitrary initial orientation to an arbitrary final orientation by a single rigid rotation through a principal angle Φ about the principal axis $\hat{\boldsymbol{e}}$; the principal axis being a judicious axis fixed in both the initial and final orientation.*

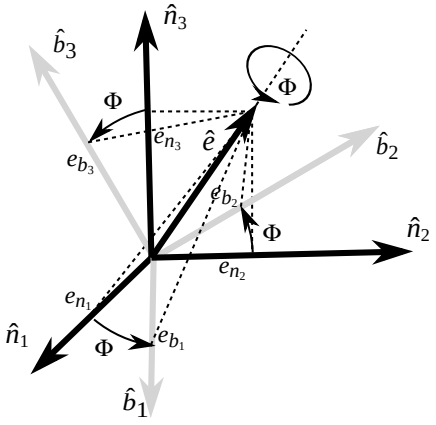


Figure 3.7: Illustration of Euler's Principal Rotation Theorem

This theorem can be visualized using Figure 3.7. Let the principal axis unit vector $\hat{\boldsymbol{e}}$ be written in $\mathcal{B}$ and $\mathcal{N}$ frame components as

$$\hat{\boldsymbol{e}} = e_{b_1}\hat{\boldsymbol{b}}_1 + e_{b_2}\hat{\boldsymbol{b}}_2 + e_{b_3}\hat{\boldsymbol{b}}_3 \tag{3.59a}$$

$$\hat{\boldsymbol{e}} = e_{n_1}\hat{\boldsymbol{n}}_1 + e_{n_2}\hat{\boldsymbol{n}}_2 + e_{n_3}\hat{\boldsymbol{n}}_3 \tag{3.59b}$$

Implicit in the theorem we see that $\hat{\boldsymbol{e}}$ will have the same vector components in the $\mathcal{B}$ as in the $\mathcal{N}$ reference frame; i.e. $e_{b_i} = e_{n_i} = e_i$. Eq. (3.5) shows that

$$\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} = [C] \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} \tag{3.60}$$

must be true. Therefore the principal axis unit vector $\hat{\boldsymbol{e}}$ is the unit eigenvector of $[C]$ corresponding to the eigenvalue $+1$. Thus the proof of the Principal Rotation Theorem reduces to proving the $[C]$ has an eigenvalue of $+1$. This proof is given in Goldstein in Ref. 8. The eigenvalue $+1$ is unique and the corresponding eigenvector is unique to within a sign of Φ and $\hat{\boldsymbol{e}}$, except for the case of a zero rotation. In this case $[C] = [I_{3\times 3}]$ and Φ would be zero, but there would be an infinity of unit axes $\hat{\boldsymbol{e}}$ such that $\hat{\boldsymbol{e}} = [I_{3\times 3}]\hat{\boldsymbol{e}}$. For the general case, the lack of sign uniqueness of Φ and $\hat{\boldsymbol{e}}$ will not cause any practical problems. The sets $(\hat{\boldsymbol{e}}, \Phi)$ and $(-\hat{\boldsymbol{e}}, -\Phi)$ both describe the same orientation.

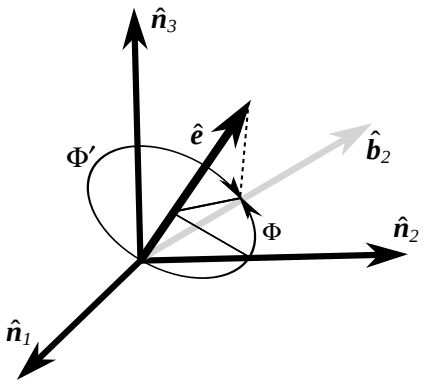


Figure 3.8: Illustration of Both Principal Rotation Angles

The principal rotation angle Φ is also not unique. Figure 3.7 shows the direction of the angle Φ labeled such that the shortest rotation about $\hat{\boldsymbol{e}}$ will be performed to move from $\mathcal{N}$ to $\mathcal{B}$. However, this is not necessary. If so desired, one can also rotate in the opposite direction by the angle Φ' and achieve the exact same orientation as shown in Figure 3.8. The difference between Φ and Φ' will always be 360 degrees. In most cases the magnitude of Φ is simply chosen to be less than or equal to 180 degrees.

To find the direction cosine matrix $[C]$ in terms of the principal rotation components $\hat{\boldsymbol{e}}$ and Φ, the fact is used that each reference frame base vector $\hat{\boldsymbol{n}}_i$

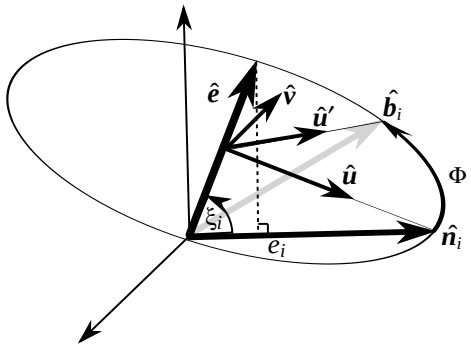


Figure 3.9: Mapping $\hat{\boldsymbol{n}}_i$ into $\hat{\boldsymbol{b}}_i$ Base Vectors

is related to $\hat{\boldsymbol{b}}_i$ through a single axis rotation about $\hat{\boldsymbol{e}}$. Let the unit principal axis vector be written as

$$\hat{\boldsymbol{e}} = e_1\hat{\boldsymbol{n}}_1 + e_2\hat{\boldsymbol{n}}_2 + e_3\hat{\boldsymbol{n}}_3 \tag{3.61}$$

and let ξ_i be the angle between $\hat{\boldsymbol{n}}_i$ and $\hat{\boldsymbol{e}}$ as shown in Figure 3.9. Let's note the following useful identity

$$\hat{\boldsymbol{e}} \cdot \hat{\boldsymbol{n}}_i = \cos\xi_i = e_i \tag{3.62}$$

Studying Figure 3.9 the base vector $\hat{\boldsymbol{b}}_i$ can be written as

$$\hat{\boldsymbol{b}}_i = \cos\xi_i\hat{\boldsymbol{e}} + \sin\xi_i\hat{\boldsymbol{u}}' = e_i\hat{\boldsymbol{e}} + \sin\xi_i\hat{\boldsymbol{u}}' \tag{3.63}$$

The unit vector $\hat{\boldsymbol{u}}'$ is given by

$$\hat{\boldsymbol{u}}' = \cos\Phi\hat{\boldsymbol{u}} + \sin\Phi\hat{\boldsymbol{v}} \tag{3.64}$$

It follows from the geometry of the single axis rotation that

$$\hat{\boldsymbol{v}} = \frac{\hat{\boldsymbol{e}} \times \hat{\boldsymbol{n}}_i}{|\hat{\boldsymbol{e}} \times \hat{\boldsymbol{n}}_i|} = \frac{1}{\sin\xi_i}(\hat{\boldsymbol{e}} \times \hat{\boldsymbol{n}}_i) \tag{3.65}$$

$$\hat{\boldsymbol{u}} = \hat{\boldsymbol{v}} \times \hat{\boldsymbol{e}} = \frac{1}{\sin\xi_i}(\hat{\boldsymbol{e}} \times \hat{\boldsymbol{n}}_i) \times \hat{\boldsymbol{e}} \tag{3.66}$$

The expression for $\hat{\boldsymbol{u}}$ can be further reduced by making use of the triple cross product identity

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{c} \tag{3.67}$$

to the simpler form

$$\hat{\boldsymbol{u}} = \frac{1}{\sin\xi_i}(\hat{\boldsymbol{n}}_i - e_i\hat{\boldsymbol{e}}) \tag{3.68}$$

After substituting Eqs. (3.64), (3.65) and (3.68) into Eq. (3.63), each base vector $\hat{\boldsymbol{b}}_i$ is expressed in terms of reference frame N base vectors.

$$\hat{\boldsymbol{b}}_i = \cos\Phi\hat{\boldsymbol{n}}_i + (1-\cos\Phi)\,\hat{\boldsymbol{e}}\hat{\boldsymbol{e}}^T\hat{\boldsymbol{n}}_i + \sin\Phi\,(\hat{\boldsymbol{e}}\times\hat{\boldsymbol{n}}_i) \tag{3.69}$$

where $\hat{\boldsymbol{e}}\hat{\boldsymbol{e}}^T$ is the outer vector dot product of the vector $\hat{\boldsymbol{e}}$. Making use of the definition of $[\tilde{\boldsymbol{e}}]$ in Eq. (3.23) the set of base vectors $\{\hat{\boldsymbol{b}}\}$ can be expressed as

$$\{\hat{\boldsymbol{b}}\} = \big(\cos\Phi[I_{3\times3}] + (1-\cos\Phi)\,\hat{\boldsymbol{e}}\hat{\boldsymbol{e}}^T - \sin\Phi[\tilde{\boldsymbol{e}}]\big)\,\{\hat{\boldsymbol{n}}\} \tag{3.70}$$

Using the relationship $\{\hat{\boldsymbol{b}}\} = [C]\{\hat{\boldsymbol{n}}\}$, the direction cosine matrix can be directly extracted from Eq. (3.69) to be

$$[C] = \begin{bmatrix} e_1^2\Sigma + c\Phi & e_1e_2\Sigma + e_3s\Phi & e_1e_3\Sigma - e_2s\Phi \\ e_2e_1\Sigma - e_3s\Phi & e_2^2\Sigma + c\Phi & e_2e_3\Sigma + e_1s\Phi \\ e_3e_1\Sigma + e_2s\Phi & e_3e_2\Sigma - e_1s\Phi & e_3^2\Sigma + c\Phi \end{bmatrix} \tag{3.71}$$

where $\Sigma = 1 - c\Phi$. Again the short hand notation $c\Phi = \cos\Phi$ and $s\Phi = \sin\Phi$ was used here. The direction cosine matrix $[C]$ depends on four scalar quantities e_1, e_2, e_3 and Φ. However, only three degrees of freedom are present since the vector components e_i must abide by the unit constraint $\sum_i^3 e_i^2 = 1$.

By inspection of Eq. (3.71), the inverse transformation from the direction cosine matrix $[C]$ to the principal rotation elements is found to be

$$\cos\Phi = \frac{1}{2}\,(C_{11} + C_{22} + C_{33} - 1) \tag{3.72}$$

$$\hat{\boldsymbol{e}} = \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} = \frac{1}{2\sin\Phi}\begin{pmatrix} C_{23} - C_{32} \\ C_{31} - C_{13} \\ C_{12} - C_{21} \end{pmatrix} \tag{3.73}$$

Note that Eq. (3.72) will yield a principal rotation angle within the range $0 \leq \Phi \leq 180$ degrees. The direction of $\hat{\boldsymbol{e}}$ in Eq. (3.73) will be such that the principal rotation parameterizing $[C]$ will be through a positive angle Φ about $\hat{\boldsymbol{e}}$. To find the second possible principal rotation angle Φ' one subtracts 360 degrees from Φ.

$$\Phi' = \Phi - 2\pi \tag{3.74}$$

The angle Φ' is equally valid as Φ and yields the same principal rotation axis $\hat{\boldsymbol{e}}$. The only difference being that a *longer* rotation (for $|\Phi| \leq \pi$) is being performed in the opposite direction. As with the sequential Euler angle rotations, the instantaneous principal rotation parameters $\{e_1(t), e_2(t), e_3(t), \Phi(t)\}$ can be considered coordinates associated with the instantaneous direction cosine matrix $[C(t)]$, and obviously does not restrict the body to actually execute the principal rotation.

Example 3.3: Let the B frame attitude relative the $\mathcal{N}$ frame be given by the

(3-2-1) Euler angles (10,25,-15) degrees. Find the corresponding principal rotation axis and angles.

Using Eq. (3.33) the direction cosine matrix $[BN]$ is

$$[BN] = \begin{bmatrix} 0.892539 & 0.157379 & -0.422618 \\ -0.275451 & 0.932257 & -0.234570 \\ 0.357073 & 0.325773 & 0.875426 \end{bmatrix}$$

The first principal rotation angle Φ is found through Eq. (3.72).

$$\Phi = \cos^{-1}\left(\frac{1}{2}\left(0.892539 + 0.932257 + 0.875426 - 1\right)\right) = 31.7762^\circ$$

The corresponding principal rotation axis is given though Eq. (3.73).

$$\hat{\boldsymbol{e}} = \frac{1}{2\sin\left(31.7762^\circ\right)}\begin{pmatrix} -0.23457 - 0.325773 \\ 0.357073 - (-0.422618) \\ 0.157379 - (-0.275451) \end{pmatrix} = \begin{pmatrix} -0.532035 \\ 0.740302 \\ 0.410964 \end{pmatrix}$$

The second principal rotation angle Φ' calculated using Eq. (3.74).

$$\Phi' = 31.7762^\circ - 360^\circ = -328.2238^\circ$$

Either principal rotation element sets $(\hat{\boldsymbol{e}}, \Phi)$ or $(\hat{\boldsymbol{e}}, \Phi')$ describes the identical attitude as the original (3-2-1) Euler angles.

Many important attitude parameters that are derived from Euler's principal rotation axis $\hat{\boldsymbol{e}}$ and angle Φ can be written in the general form

$$\boldsymbol{p} = f(\Phi)\hat{\boldsymbol{e}} \tag{3.75}$$

where $f(\Phi)$ could be any scalar function of Φ. All these attitude coordinate vectors have the same direction and differ only by their magnitude $|\boldsymbol{p}\,| = f(\Phi)$.

The principal rotation vector $\boldsymbol{\gamma}$ is simply defined as

$$\boldsymbol{\gamma} = \Phi\hat{\boldsymbol{e}} \tag{3.76}$$

Therefore the magnitude of $\boldsymbol{\gamma}$ is $f(\Phi) = \Phi$. This attitude vector has a very interesting relationship to the direction cosine matrix that can be verified to also hold for higher dimensional orthogonal projections as shown in Ref. 17. To gain more insight, consider the special case of a pure single-axis rotation about a fixed $\hat{\boldsymbol{e}}$ with the rotation angle being Φ. The angular velocity vector for this case is

$$\boldsymbol{\omega} = \dot{\Phi}\hat{\boldsymbol{e}} \tag{3.77}$$

or in matrix form:

$$[\tilde{\boldsymbol{\omega}}] = \dot{\Phi}[\tilde{\boldsymbol{e}}] \tag{3.78}$$

Substituting Eq. (3.78) into Eq. (3.27) leads to the following development:

$$\begin{aligned} \frac{d[C]}{dt} &= -\frac{d\Phi}{dt}[\tilde{\boldsymbol{e}}][C] \\ \frac{d[C]}{d\Phi} &= -[\tilde{\boldsymbol{e}}][C] \\ [C] &= e^{-\Phi[\tilde{\boldsymbol{e}}]} \end{aligned} \tag{3.79}$$

The last step holds true for $[\tilde{\boldsymbol{e}}]$ being a constant matrix for a rotation about a fixed axis. Due to Euler's principal rotation theorem, however, any arbitrary rotation can be instantaneously described by the equivalent single-axis rotation. Euler's theorem means that Eq. (3.79) holds at any instant for an arbitrary time varying direction cosine matrix $[C]$. Note for time-varying $[C]$, however, that $\hat{\boldsymbol{e}}$ and Φ must be considered time-varying. Using Eq. (3.76) the rotation matrix $[C]$ is related to $\boldsymbol{\gamma}$ through

$$[C] = e^{-[\tilde{\gamma}]} = \sum_{n=0}^{\infty} \frac{1}{n!} \left(-[\tilde{\boldsymbol{\gamma}}]\right)^n \tag{3.80}$$

It turns out that this mapping also holds for higher dimensional proper orthogonal matrices $[C]$. For the case of three-dimensional rotations, the infinite power series in Eq. (3.80) can more conveniently be written as a finite, closed form solution.[5, 17]

$$[C] = e^{-\Phi[\tilde{\boldsymbol{e}}]} = [I_{3\times 3}]\cos\Phi - \sin\Phi[\tilde{\boldsymbol{e}}] + (1-\cos\Phi)\hat{\boldsymbol{e}}\hat{\boldsymbol{e}}^T \tag{3.81}$$

To find the inverse transformation from $[C]$ to $\boldsymbol{\gamma}$, the inverse matrix logarithm is taken.

$$[\tilde{\boldsymbol{\gamma}}] = -\ln[C] = \sum_{n=0}^{\infty} \frac{1}{n}(1-[C])^n \tag{3.82}$$

This inverse mapping is defined everywhere except for $\Phi = 0$ and $\Phi = \pm 180$ degree rotations. For these rotations, the non-uniqueness of the $\boldsymbol{\gamma}$ vector that leads to mathematical difficulties. Otherwise a vector $\boldsymbol{\gamma}$ is reliably returned corresponding to a principal rotation of less than or equal to 180 degrees.

Example 3.4: In Example 3.3 it was shown that the direction cosine matrix

$$[BN] = \begin{bmatrix} 0.892539 & 0.157379 & -0.422618 \\ -0.275451 & 0.932257 & -0.23457 \\ 0.357073 & 0.325773 & 0.875426 \end{bmatrix}$$

represents the equivalent orientation as the principal rotation vector

$$\boldsymbol{\gamma} = 0.55460 rad \begin{pmatrix} -0.532035 \\ 0.740302 \\ 0.410964 \end{pmatrix} = \begin{pmatrix} -0.295067 \\ 0.410571 \\ 0.227921 \end{pmatrix}$$

To verify the mapping in Eq. (3.80) let's write $[\tilde{\gamma}]$ using the definition of tilde matrix operator in Eq. (3.23).

$$[\tilde{\gamma}] = \begin{bmatrix} 0 & -0.227921 & 0.410571 \\ 0.227921 & 0 & 0.295067 \\ -0.410571 & -0.295067 & 0 \end{bmatrix}$$

Using software packages such as Mathematica or MATLAB, the matrix exponential mapping in Eq. (3.80) can be solved numerically for the corresponding direction cosine matrix $[BN]$.

$$[BN] = e^{-[\tilde{\gamma}]} = \begin{bmatrix} 0.892539 & 0.157379 & -0.422618 \\ -0.275451 & 0.932257 & -0.234570 \\ 0.357073 & 0.325773 & 0.875426 \end{bmatrix} \surd$$

Let $(\Phi_1, \hat{\boldsymbol{e}}_1)$ be the principal rotation elements that relate the $\mathcal{B}$ frame relative to the $\mathcal{N}$ frame, while $(\Phi_2, \hat{\boldsymbol{e}}_2)$ orients the $\mathcal{F}$ frame relative to the $\mathcal{B}$ frame. The $\mathcal{F}$ frame is related directly to the $\mathcal{N}$ frame by the elements $(\Phi, \hat{\boldsymbol{e}})$ through the relationship

$$[FN(\Phi, \hat{\boldsymbol{e}})] = [FB(\Phi_2, \hat{\boldsymbol{e}}_2)][BN(\Phi_1, \hat{\boldsymbol{e}}_1)] \tag{3.83}$$

Instead of solving for the overall principal rotation elements through the corresponding direction cosine matrix, it is possible to express $(\Phi, \hat{\boldsymbol{e}})$ directly in terms of $(\Phi_1, \hat{\boldsymbol{e}}_1)$ and $(\Phi_2, \hat{\boldsymbol{e}}_2)$ through[5]

$$\Phi = 2\cos^{-1}\left(\cos\frac{\Phi_1}{2}\cos\frac{\Phi_2}{2} - \sin\frac{\Phi_1}{2}\sin\frac{\Phi_2}{2}\hat{\boldsymbol{e}}_1 \cdot \hat{\boldsymbol{e}}_2\right) \tag{3.84}$$

$$\hat{\boldsymbol{e}} = \frac{\cos\frac{\Phi_2}{2}\sin\frac{\Phi_1}{2}\hat{\boldsymbol{e}}_1 + \cos\frac{\Phi_1}{2}\sin\frac{\Phi_2}{2}\hat{\boldsymbol{e}}_2 + \sin\frac{\Phi_1}{2}\sin\frac{\Phi_2}{2}\hat{\boldsymbol{e}}_1 \times \hat{\boldsymbol{e}}_2}{\sin\frac{\Phi}{2}} \tag{3.85}$$

This composite rotation property is easily derived from the Euler parameter composite rotation property shown in the next section. Given the two principal rotation element sets $(\Phi_1, \hat{\boldsymbol{e}}_1)$ and $(\Phi, \hat{\boldsymbol{e}})$, the relative orientation set $(\Phi_2, \hat{\boldsymbol{e}}_2)$ is expressed similarly through

$$\Phi_2 = 2\cos^{-1}\left(\cos\frac{\Phi}{2}\cos\frac{\Phi_1}{2} + \sin\frac{\Phi}{2}\sin\frac{\Phi_1}{2}\hat{\boldsymbol{e}} \cdot \hat{\boldsymbol{e}}_1\right) \tag{3.86}$$

$$\hat{\boldsymbol{e}}_2 = \frac{\cos\frac{\Phi_1}{2}\sin\frac{\Phi}{2}\hat{\boldsymbol{e}} - \cos\frac{\Phi}{2}\sin\frac{\Phi_1}{2}\hat{\boldsymbol{e}}_1 + \sin\frac{\Phi}{2}\sin\frac{\Phi_1}{2}\hat{\boldsymbol{e}} \times \hat{\boldsymbol{e}}_1}{\sin\frac{\Phi_2}{2}} \tag{3.87}$$

The kinematic differential equation of the principal rotation vector $\boldsymbol{\gamma}$ is given by[5, 18–20]

$$\dot{\boldsymbol{\gamma}} = \left[[I_{3\times3}] + \frac{1}{2}[\tilde{\gamma}] + \frac{1}{\Phi^2}\left(1 - \frac{\Phi}{2}\cot\left(\frac{\Phi}{2}\right)\right)[\tilde{\gamma}]^2\right]\boldsymbol{\omega} \tag{3.88}$$

where $\Phi = \| \boldsymbol{\gamma} \|$. The inverse transformation of Eq. (3.88) is

$$\boldsymbol{\omega} = \left[[I_{3\times 3}] - \left(\frac{1-\cos\Phi}{\Phi^2} \right) [\tilde{\gamma}] + \left(\frac{\Phi - \sin\Phi}{\Phi^3} \right) [\tilde{\gamma}]^2 \right] \dot{\boldsymbol{\gamma}} \tag{3.89}$$

As expected, the kinematic differential equation in Eq. (3.88) contains a 0/0 type mathematical singularity for zero rotations where $\Phi = 0$ degrees. Therefore, the principal rotation vector is not well suited for use in small motion feedback control type applications where the reference state is the zero rotation. Further, the mathematical expression in Eq. (3.88) is rather complex, containing polynomial fractions of degrees up to three in addition to trigonometric functions. This makes $\boldsymbol{\gamma}$ less attractive to describe large arbitrary rotations as compared to some other, closely related, attitude parameters that will be presented in the next few sections.

Example 3.5: Given the prescribed body angular velocity vector $\boldsymbol{\omega} = \omega(t)\hat{\boldsymbol{e}}$ for a single axis rotation, Eq. (3.88) yields the following kinematic differential equation for the principal rotation vector $\boldsymbol{\gamma} = \Phi\hat{\boldsymbol{e}}$.

$$\dot{\boldsymbol{\gamma}} = \left[[I_{3\times 3}] - \frac{\Phi}{2}[\tilde{\gamma}] + \frac{1}{\Phi^2}\left(1 - \frac{\Phi}{2}\cot\left(\frac{\Phi}{2}\right)\right) \Phi^2 [\tilde{\gamma}]^2 \right] \omega(t)\hat{\boldsymbol{e}}$$

Noting that $[\tilde{\gamma}]\hat{\boldsymbol{e}} = \Phi[\tilde{\boldsymbol{e}}]\hat{\boldsymbol{e}} = 0$, this is simplified to

$$\dot{\boldsymbol{\gamma}} = \omega(t)\hat{\boldsymbol{e}}$$

Therefore the general expression in Eq. (3.88) simplifies to the single axis result in Eq. (3.77).

The principal rotation elements $\hat{\boldsymbol{e}}$ and Φ have had a fundamental influence on the derivation of many sets of attitude coordinates. All of the following attitude parameters will be directly derived from these principal rotation elements.

3.4 Euler Parameters

Another popular set of attitude coordinates are the four Euler parameters (quaternions). They provide a redundant, nonsingular attitude description and are well suited to describe arbitrary, large rotations. The Euler parameter vector $\boldsymbol{\beta}$ is defined in terms of the principal rotation elements as

$$\beta_0 = \cos(\Phi/2) \tag{3.90a}$$
$$\beta_1 = e_1 \sin(\Phi/2) \tag{3.90b}$$
$$\beta_2 = e_2 \sin(\Phi/2) \tag{3.90c}$$
$$\beta_3 = e_3 \sin(\Phi/2) \tag{3.90d}$$

It is evident since $e_1^2+e_2^2+e_3^2=1$, that the β_i's satisfy the holonomic constraint

$$\beta_0^2+\beta_1^2+\beta_2^2+\beta_3^2=1 \tag{3.91}$$

Note that this constraint geometrically describes a four-dimensional unit sphere. Any rotation described through the Euler parameters has a trajectory on the surface of this constraint sphere. Given a certain attitude, there are actually two sets of Euler parameters that will describe the same orientation. This is due to the non-uniqueness of the principal rotation elements themselves. Switching between the sets $(\hat{\boldsymbol{e}},\Phi)$ and $(-\hat{\boldsymbol{e}},-\Phi)$ will yield the same Euler parameter vector $\boldsymbol{\beta}$. However, if the second principal rotation angle Φ' is used, another Euler parameter vector $\boldsymbol{\beta}'$ is found. Using Eq. (3.74) one can show that

$$\beta_0' = \cos\left(\frac{\Phi'}{2}\right) = \cos\left(\frac{\Phi}{2}-\pi\right) = -\cos\left(\frac{\Phi}{2}\right) = -\beta_0$$
$$\beta_i' = e_i\sin\left(\frac{\Phi'}{2}\right) = e_i\sin\left(\frac{\Phi'}{2}-\pi\right) = -e_i\sin\left(\frac{\Phi}{2}\right) = -\beta_i$$

Therefore the vector $\boldsymbol{\beta}'=-\boldsymbol{\beta}$ describes the same orientation as the vector $\boldsymbol{\beta}$. This results in the following interesting observation. Since any point on the unit constraint sphere surface represent a specific orientation, the anti-pole to that point represents the exact same orientation. The difference between the two attitude descriptions is that one specifies the orientation through the shortest single axis rotation, the other through the longest. From Eq.(3.90a) it is clear that in order to choose the Euler parameter vector corresponding to the shortest rotation (i.e. $|\Phi|\leq 180$ degrees), the coordinate β_0 must be chosen to be non-negative.

Using the trigonometric identities

$$\sin\Phi = 2\sin(\Phi/2)\cos(\Phi/2)$$
$$\cos\Phi = 2\cos^2(\Phi/2)-1$$

in Eq. (3.71), the direction cosine matrix can be written in terms of the Euler parameters as

$$[C]=\begin{bmatrix}\beta_0^2+\beta_1^2-\beta_2^2-\beta_3^2 & 2(\beta_1\beta_2+\beta_0\beta_3) & 2(\beta_1\beta_3-\beta_0\beta_2)\\ 2(\beta_1\beta_2-\beta_0\beta_3) & \beta_0^2-\beta_1^2+\beta_2^2-\beta_3^2 & 2(\beta_2\beta_3+\beta_0\beta_1)\\ 2(\beta_1\beta_3+\beta_0\beta_2) & 2(\beta_2\beta_3-\beta_0\beta_1) & \beta_0^2-\beta_1^2-\beta_2^2+\beta_3^2\end{bmatrix} \tag{3.92}$$

The fact that $\boldsymbol{\beta}$ and $-\boldsymbol{\beta}$ produces the same direction cosine matrix $[C]$ can be easily verified in Eq. (3.92). All Euler parameters appear in quadratic product pairs, thus changing the signs of all β_i components has no effect on the resulting $[C]$ matrix. It is evident that the most general angular motion of a reference frame generates two arcs on the four dimensional unit sphere (the geodesic arcs generated by $\boldsymbol{\beta}(t)$ and $-\boldsymbol{\beta}(t)$). This elegant description is universally nonsingular and is unique to within the sign $\pm\boldsymbol{\beta}(t)$. The inverse transformations from

$[C]$ to the Euler parameters can be found through inspection of Eq. (3.92) to be

$$\beta_0 = \pm\frac{1}{2}\sqrt{C_{11}+C_{22}+C_{33}+1} \tag{3.93a}$$

$$\beta_1 = \frac{C_{23}-C_{32}}{4\beta_0} \tag{3.93b}$$

$$\beta_2 = \frac{C_{31}-C_{13}}{4\beta_0} \tag{3.93c}$$

$$\beta_3 = \frac{C_{12}-C_{21}}{4\beta_0} \tag{3.93d}$$

Note that the non-uniqueness of the Euler parameters is evident again in this inverse transformation. By keeping the + sign in Eq. (3.93a) one restricts the corresponding principal rotation angle Φ to be less than or equal to 180 degrees. From a practical point of few this non-uniqueness does not pose any difficulties. Initially one simply picks an initial condition on one Euler parameter trajectory and then remains with it either through solving an associated kinematic differential developed below, or using elementary continuity logic.

Clearly Eq. (3.93) has a 0/0 type mathematical singularity whenever $\beta_0 \to 0$. This corresponds to the $\boldsymbol{\beta}$ vector describing any 180 degree principal rotation. A computationally superior algorithm has been developed by Stanley in Ref. 21. First the the four β_i^2 terms are computed.

$$\beta_0^2 = \frac{1}{4}\left(1 + Trace[C]\right) \tag{3.94a}$$

$$\beta_1^2 = \frac{1}{4}\left(1 + 2C_{11} - Trace[C]\right) \tag{3.94b}$$

$$\beta_2^2 = \frac{1}{4}\left(1 + 2C_{22} - Trace[C]\right) \tag{3.94c}$$

$$\beta_3^2 = \frac{1}{4}\left(1 + 2C_{33} - Trace[C]\right) \tag{3.94d}$$

Then Stanley takes the square root of the largest β_i^2 found in Eq. (3.94) where the sign of β_i is arbitrarily chosen to be positive. The other β_j's are found by dividing the appropriate three of the following six in Eq. (3.95) by the chosen largest β_i coordinate.

$$\beta_0\beta_1 = (C_{23}-C_{32})/4 \tag{3.95a}$$

$$\beta_0\beta_2 = (C_{31}-C_{13})/4 \tag{3.95b}$$

$$\beta_0\beta_3 = (C_{12}-C_{21})/4 \tag{3.95c}$$

$$\beta_2\beta_3 = (C_{23}+C_{32})/4 \tag{3.95d}$$

$$\beta_3\beta_1 = (C_{31}+C_{13})/4 \tag{3.95e}$$

$$\beta_1\beta_2 = (C_{12}+C_{21})/4 \tag{3.95f}$$

To find the alternate set of Euler parameter, the sign of the chosen β_i would simply be set negative.

Example 3.6: Let's use Stanley's method to find the Euler parameters of the direction cosine matrix $[C]$.

$$[C] = \begin{bmatrix} 0.892539 & 0.157379 & -0.422618 \\ -0.275451 & 0.932257 & -0.234570 \\ 0.357073 & 0.325773 & 0.875426 \end{bmatrix}$$

Using the expressions in Eq. (3.94) the absolute values of the four Euler parameter are found.

$$\beta_0^2 = 0.925055 \qquad \beta_1^2 = 0.021214$$
$$\beta_2^2 = 0.041073 \qquad \beta_3^2 = 0.012657$$

The β_0 term is selected as the largest element and used in Eqs. (3.95a) through (3.95c) to find the Euler parameter vector.

$$\boldsymbol{\beta} = (0.961798, -0.14565, 0.202665, 0.112505)^T$$

The alternate Euler parameter vector would be found be simply reversing the sign of each element in $\boldsymbol{\beta}$.

A very important *composite rotation* property of the Euler parameters is the manner in which they allow two sequential rotations to be combined into one overall composite rotation. Let the Euler parameter vector $\boldsymbol{\beta}'$ describe the first, $\boldsymbol{\beta}''$ the second and $\boldsymbol{\beta}$ the composite rotation. From Eq. (3.20) it is clear that

$$[FN(\boldsymbol{\beta})] = [FB(\boldsymbol{\beta}'')][BN(\boldsymbol{\beta}')] \tag{3.96}$$

Using Eq. (3.92) in Eq. (3.96) and equating corresponding elements leads to following elegant transformation that bi-linearly combines $\boldsymbol{\beta}'$ and $\boldsymbol{\beta}''$ into $\boldsymbol{\beta}$.

$$\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{bmatrix} \beta_0'' & -\beta_1'' & -\beta_2'' & -\beta_3'' \\ \beta_1'' & \beta_0'' & \beta_3'' & -\beta_2'' \\ \beta_2'' & -\beta_3'' & \beta_0'' & \beta_1'' \\ \beta_3'' & \beta_2'' & -\beta_1'' & \beta_0'' \end{bmatrix} \begin{pmatrix} \beta_0' \\ \beta_1' \\ \beta_2' \\ \beta_3' \end{pmatrix} \tag{3.97}$$

By *transmutation* of Eq.(3.97) an alternate expression $\boldsymbol{\beta} = [G(\boldsymbol{\beta}')]\boldsymbol{\beta}''$ is found

$$\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{bmatrix} \beta_0' & -\beta_1' & -\beta_2' & -\beta_3' \\ \beta_1' & \beta_0' & -\beta_3' & \beta_2' \\ \beta_2' & \beta_3' & \beta_0' & -\beta_1' \\ \beta_3' & -\beta_2' & \beta_1' & \beta_0' \end{bmatrix} \begin{pmatrix} \beta_0'' \\ \beta_1'' \\ \beta_2'' \\ \beta_3'' \end{pmatrix} \tag{3.98}$$

where the components of the matrix $[G(\boldsymbol{\beta}')]$ are given in Eq. (3.98). Note the useful identity

$$[G(\boldsymbol{\beta})]^T\boldsymbol{\beta} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{3.99}$$

By inspection, it is evident that the 4x4 matrices in Eqs. (3.97) and (3.98) are orthogonal. These transformations provide a simple, nonsingular and bilinear method to combine two successive rotations described through Euler parameters. For other attitude parameters such as the Euler angles, this same composite transformation would yield a very complicated, transcendental expression.

Example 3.7: Using Stanley's method, the direction cosine matrices $[BN]$ and $[FB]$ defined in Example 3.1 can be parameterized through the Euler parameter vectors $\boldsymbol{\beta}'$ and $\boldsymbol{\beta}''$ respectively as

$$[BN] \Rightarrow \quad \boldsymbol{\beta}' = \left(0, \tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}, 0\right)^T$$

$$[FB] \Rightarrow \quad \boldsymbol{\beta}'' = \left(\tfrac{1}{2}\sqrt{\tfrac{\sqrt{3}}{2}+1}, -\tfrac{1}{2}\sqrt{\tfrac{\sqrt{3}}{2}+1}, \tfrac{-\sqrt{2}}{4\sqrt{2+\sqrt{3}}}, \tfrac{\sqrt{2}}{4\sqrt{2+\sqrt{3}}}\right)^T$$

Note that the vector $\boldsymbol{\beta}'$ describes the attitude of the $\mathcal{B}$ frame relative to the $\mathcal{N}$ frame, while the vector $\boldsymbol{\beta}''$ describes the $\mathcal{F}$ frame attitude relative to the $\mathcal{B}$ frame. Eq. (3.97) can be used to combine the two successive attitude vectors into one vector $\boldsymbol{\beta}$ which directly describes the $\mathcal{F}$ frame orientation relative to the $\mathcal{N}$ frame.

$$\boldsymbol{\beta} = \frac{1}{2\sqrt{2}}\left(\sqrt{3}, \sqrt{3}, 1, 1\right)^T$$

To verify that $\boldsymbol{\beta}$ does indeed parameterize the direction cosine matrix $[FN]$ given in Example 3.1, it can be back substituted into Eq. (3.92) to yield

$$[FN] = \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ 0 & 0 & -1 \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \end{bmatrix} \quad \surd$$

The kinematic differential equation for the Euler parameters can be derived by differentiating the β_i's in Eq. (3.93). The following development will establish the kinematic equation for $\dot{\beta}_0$ only, the remaining $\dot{\beta}_i$ equations can be developed in an analogous manner. After taking the derivative of Eq. (3.93a), $\dot{\beta}_0$ is expressed as

$$\dot{\beta}_0 = \frac{\dot{C}_{11} + \dot{C}_{22} + \dot{C}_{33}}{8\beta_0} \tag{3.100}$$

After using the expressions for $\dot{C}_{ii}$ given in Eq. (3.27), the term $\dot{\beta}_0$ is rewritten as

$$\dot{\beta}_0 = \frac{1}{2}\left(-\frac{C_{23} - C_{32}}{4\beta_0}\omega_1 - \frac{C_{31} - C_{13}}{4\beta_0}\omega_2 - \frac{C_{12} - C_{21}}{4\beta_0}\omega_3\right) \tag{3.101}$$

Using Eqs. (3.93b) through (3.93d), the $\dot{\beta}_0$ differential equation is simplified to

$$\dot{\beta}_0 = \frac{1}{2}\left(-\beta_1\omega_1 - \beta_2\omega_2 - \beta_3\omega_3\right) \tag{3.102}$$

After performing a similar derivation for the $\dot{\beta}_1$, $\dot{\beta}_2$ and $\dot{\beta}_3$ terms, the four coupled kinematic differential equations for the Euler parameters are found to be the exceptionally elegant matrix form

$$\begin{pmatrix} \dot{\beta}_0 \\ \dot{\beta}_1 \\ \dot{\beta}_2 \\ \dot{\beta}_3 \end{pmatrix} = \frac{1}{2} \begin{bmatrix} 0 & -\omega_1 & -\omega_2 & -\omega_3 \\ \omega_1 & 0 & \omega_3 & -\omega_2 \\ \omega_2 & -\omega_3 & 0 & \omega_1 \\ \omega_3 & \omega_2 & -\omega_1 & 0 \end{bmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} \tag{3.103}$$

or by transmutation of Eq. (3.103), the kinematic differential equation has the elegant form

$$\begin{pmatrix} \dot{\beta}_0 \\ \dot{\beta}_1 \\ \dot{\beta}_2 \\ \dot{\beta}_3 \end{pmatrix} = \frac{1}{2} \begin{bmatrix} \beta_0 & -\beta_1 & -\beta_2 & -\beta_3 \\ \beta_1 & \beta_0 & -\beta_3 & \beta_2 \\ \beta_2 & \beta_3 & \beta_0 & -\beta_1 \\ \beta_3 & -\beta_2 & \beta_1 & \beta_0 \end{bmatrix} \begin{pmatrix} 0 \\ \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \tag{3.104}$$

Note that the transformation matrix relating $\dot{\boldsymbol{\beta}}$ and $\boldsymbol{\omega}$ is orthogonal and singularity free. The inverse transformation from $\boldsymbol{\omega}$ to $d(\boldsymbol{\beta})/dt$ is always defined. Further, the Euler parameter kinematic differential equation of Eq. (3.103) is rigorously linear if $\omega_i(t)$ are known functions of time only. If $\omega_i(t)$ are themselves coordinates, then Eqs. (3.103) and (3.104) are more generally considered bi-linear. This makes the Euler parameters very attractive attitude coordinates for attitude estimation problems where the kinematic differential equation is linearized. All three parameter sets of attitude coordinates always have kinematic differential equations which are nonlinear and contain 0/0 type mathematical singularities. In attitude estimation problems their linearization is only locally valid. Whereas the linear (or bi-linear) property of the Euler parameter kinematic differential equation is globally valid. The Euler parameter kinematic differential equation in Eq. (3.104) can be written compactly as

$$\dot{\boldsymbol{\beta}} = \frac{1}{2}[B(\boldsymbol{\beta})]\boldsymbol{\omega} \tag{3.105}$$

where the *4x3* matrix $[B(\boldsymbol{\beta})]$ is defined as

$$[B(\boldsymbol{\beta})] = \begin{bmatrix} -\beta_1 & -\beta_2 & -\beta_3 \\ \beta_0 & -\beta_3 & \beta_2 \\ \beta_3 & \beta_0 & -\beta_1 \\ -\beta_2 & \beta_1 & \beta_0 \end{bmatrix} \tag{3.106}$$

By carrying out the matrix algebra, the following useful identities can easily be verified.

$$[B(\boldsymbol{\beta})]^T\boldsymbol{\beta} = \mathbf{0} \tag{3.107}$$

$$[B(\boldsymbol{\beta})]^T\boldsymbol{\beta}' = -[B(\boldsymbol{\beta}')]^T\boldsymbol{\beta} \tag{3.108}$$

It is easily verified that the normalization condition $\boldsymbol{\beta}^T\boldsymbol{\beta} = 1$ is a rigorous analytical integral of Eqs. (3.103), (3.104). However, in practice the norm of $\boldsymbol{\beta}$ may slightly differ form 1 when numerically integrating Eq. (3.103). It is therefore necessary to take care to reimpose this condition differentially after each numerical integration step, if the solution is to remain valid over long time intervals. However, in contrast to the re-normalization of $[C(t)]$ to satisfy $[C]^T[C] = [I_{3\times 3}]$ when solving Eq. (3.27), only one scalar condition needs to be considered when integrating $\boldsymbol{\beta}(t)$.

In control applications, often the four Euler parameters are broken up into two groups. The parameter β_0 is single out since it contains no information regarding the corresponding principal rotation axis of the orientation being represented. In effect, if is a scalar measure of the three dimensional rigid body attitude measure whose value is +1 or -1 if the attitude is zero. The remaining three Euler parameters are grouped together into a three-dimensional vector as

$$\boldsymbol{\epsilon} \equiv (\beta_1, \beta_2, \beta_3)^T \tag{3.109}$$

If the attitude goes to zero, then so will this vector. From Euler parameter differential equation in Eq. (3.104), it is evident that the differential equations for $\dot{\beta}_0$ and $\dot{\boldsymbol{\epsilon}}$ are of the form

$$\dot{\beta}_0 = -\frac{1}{2}\boldsymbol{\epsilon}^T\boldsymbol{\omega} = -\frac{1}{2}\boldsymbol{\omega}^T\boldsymbol{\epsilon} \tag{3.110}$$

$$\dot{\boldsymbol{\epsilon}} = \frac{1}{2}[T]\boldsymbol{\omega} \tag{3.111}$$

The 3×3 matrix $[T]$ is defined as

$$[T(\beta_0, \boldsymbol{\epsilon})] = \beta_0[I_{3\times 3}] + [\tilde{\boldsymbol{\epsilon}}] \tag{3.112}$$

3.5 Classical Rodrigues Parameters

The origin of the classical Rodrigues parameter vector $\boldsymbol{q}$ (or Gibbs vector) dates back over a hundred years to the French mathematician O. M. Rodrigues. This rigid body attitude coordinate set reduces the redundant Euler parameters to a minimal three parameter set through the transformation

$$q_i = \frac{\beta_i}{\beta_0} \qquad i = 1, 2, 3 \tag{3.113}$$

The inverse transformation from classical Rodrigues parameters to Euler parameters is given by

$$\beta_0 = \frac{1}{\sqrt{1 + \boldsymbol{q}^T\boldsymbol{q}}} \tag{3.114a}$$

$$\beta_i = \frac{q_i}{\sqrt{1 + \boldsymbol{q}^T\boldsymbol{q}}} \qquad i = 1, 2, 3 \tag{3.114b}$$

Using the definitions in Eq. (3.90) the vector $\boldsymbol{q}$ is expressed directly in terms of the principal rotation elements as the elegant transformation

$$\boldsymbol{q} = \tan\frac{\Phi}{2}\hat{\boldsymbol{e}} \tag{3.115}$$

From Eqs. (3.113) and (3.115) it is evident that the classical Rodrigues parameters go singular whenever $\Phi \to \pm 180$ degrees. Very large rotations can be described with these parameters without ever approaching a geometric singularity. For rotations with $|\Phi| \leq 90^{\circ}$, it is evident that $\boldsymbol{q}(t)$ locates points near the origin bounded by the unit sphere. Compare this $\pm 180^{\circ}$ nonsingular range to the Euler angles where any orientation is never more than 90 degrees away from a singularity.

The small angle behavior of the classical Rodrigues parameters is also more linear than compared to the small angle behavior of any Euler angle set. Linearizing Eq. (3.115) it is evident that

$$\boldsymbol{q} \approx \frac{\Phi}{2}\hat{\boldsymbol{e}} \tag{3.116}$$

This means that classical Rodrigues parameters will linearize roughly to an "angle over 2" type quantity, whereas the Euler angles linearize as an angle type quantity well removed from singular points.

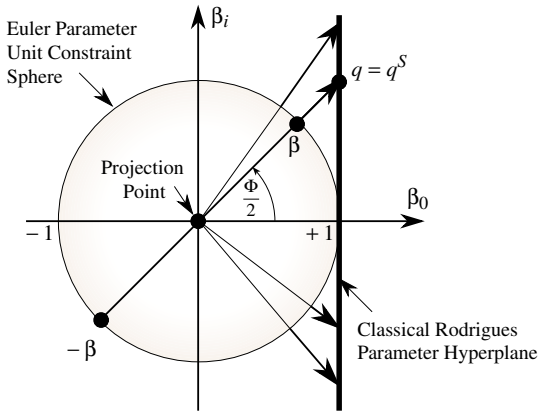


Figure 3.10: Stereographic Projection of Euler Parameters to Classical Rodrigues Parameters

As discussed in Ref. 22, the classical Rodrigues parameters can be viewed as a special set of stereographic orientation parameters. Stereographic projections are used to map a higher-dimensioned spherical surface onto a lower-dimensioned hyperplane. In this case, the surface of the four-dimensional Euler parameter unit constraint sphere in Eq. (3.91) is mapped (projected) onto a three-dimensional hyperplane though Eq. (3.113). Figure 3.10 illustrates how

such a projection would yield the classical Rodrigues parameters. The projection point is chosen to be the origin $\boldsymbol{\beta} = 0$ and the hyperplane upon which all Euler parameter coordinates are projected is the tangent surface at $\beta_0 = 1$. Note that on the constraint sphere surface $\beta_0 = 1$ corresponds to a $\Phi = 0$ degrees, $\beta_0 = 0$ corresponds to $\Phi = \pm 180$ degrees and β_0 = -1 represents $\Phi = \pm 360$ degrees. The transformation in Eq. (3.113) maps any Euler parameter set on the unit constraint sphere surface onto a corresponding point located on the classical Rodrigues parameter hyperplane.

All stereographic orientation parameters can be viewed as a projection of the constraint sphere onto some hyperplane. Since the Euler parameters themselves are not unique, the corresponding stereographic orientation parameters are also generally not unique. The set corresponding to the projection of the Euler parameter set $-\boldsymbol{\beta}$ is referred to as the *shadow* set and is differentiated from the original set by a superscript S.[22] However, it turns out that the shadow set of the classical Rodrigues parameters are indeed identical to the original classical Rodrigues parameters as is easily verified by reversing the β_i signs in Eq. (3.113) or by inspection of Figure 3.10.

$$q_i^S = \frac{-\beta_i}{-\beta_0} = q_i \tag{3.117}$$

The direction cosine matrix in terms of the classical Rodrigues parameters can be found by using their definition in Eq. (3.113) in the direction cosine matrix formulation in Eq. (3.92). The resulting parameterization is in matrix form[4, 22]

$$[C] = \frac{1}{1+\boldsymbol{q}^T\boldsymbol{q}} \begin{bmatrix} 1+q_1^2-q_2^2-q_3^2 & 2\left(q_1q_2+q_3\right) & 2\left(q_1q_3-q_2\right) \\ 2\left(q_2q_1-q_3\right) & 1-q_1^2+q_2^2-q_3^2 & 2\left(q_2q_3+q_1\right) \\ 2\left(q_3q_1+q_2\right) & 2\left(q_3q_2-q_1\right) & 1-q_1^2-q_2^2+q_3^2 \end{bmatrix} \tag{3.118}$$

and in vector form[5, 22]

$$[C] = \frac{1}{1+\boldsymbol{q}^T\boldsymbol{q}} \left(\left(1-\boldsymbol{q}^T\boldsymbol{q}\right)[I_{3\times 3}] + 2\boldsymbol{q}\boldsymbol{q}^T - 2[\tilde{\boldsymbol{q}}]\right) \tag{3.119}$$

The simplest way to extract the classical Rodrigues parameters from a given direction cosine matrix is to determine the Euler parameters first and then use Eq. (3.113) to find the corresponding Rodrigues parameters. Note the following useful identity.

$$[C(\boldsymbol{q})]^T = [C(-\boldsymbol{q})] \tag{3.120}$$

Since $\boldsymbol{q}$ defines the relative orientation of a second frame to a first frame, the relative orientation of the second frame relative to the first corresponds simply to reversing the sign of $\boldsymbol{q}$ as in

$$\{\hat{\boldsymbol{n}}\} = [C(\boldsymbol{q})]^T\{\hat{\boldsymbol{b}}\} = [C(-\boldsymbol{q})]\{\hat{\boldsymbol{b}}\} \tag{3.121}$$

This elegant property doesn't exist with Euler angles.

Similar to the direction cosine matrices and Euler parameters, the classical Rodrigues parameter vectors have a composite rotation property. Given two attitude vectors $\boldsymbol{q}'$ and $\boldsymbol{q}''$, let the overall composite attitude vector $\boldsymbol{q}$ be defined through the quadratically nonlinear condition

$$[FN(\boldsymbol{q})] = [FB(\boldsymbol{q}'')][BN(\boldsymbol{q}')] \tag{3.122}$$

However, solving for an overall transformation from $\boldsymbol{q}'$ and $\boldsymbol{q}''$ to $\boldsymbol{q}$ using Eq. (3.122) is very cumbersome. Using the successive rotation property of the Euler parameters and the definition of the classical Rodrigues parameters in Eq. (3.113), the composite attitude vector $\boldsymbol{q}$ is expressed directly in terms of $\boldsymbol{q}'$ and $\boldsymbol{q}''$ through[5, 23]

$$\boldsymbol{q} = \frac{\boldsymbol{q}'' + \boldsymbol{q}' - \boldsymbol{q}'' \times \boldsymbol{q}'}{1 - \boldsymbol{q}'' \cdot \boldsymbol{q}'} \tag{3.123}$$

Assume that the attitude vectors $\boldsymbol{q}$ and $\boldsymbol{q}'$ are given and the relative attitude vector $\boldsymbol{q}''$ is to be found. With direction cosine matrices and Euler parameters the two attitude descriptions were related through an orthogonal matrix which made finding the relative attitude description trivial. This is no longer the case with the classical Rodrigues parameter composite rotation property. However, we can use Eq. (3.122) to solve for $[FB(\boldsymbol{q}'')]$ first using the orthogonality of the direction cosine matrices.

$$[FB(\boldsymbol{q}'')] = [FN(\boldsymbol{q})][BN(\boldsymbol{q}')]^T \tag{3.124}$$

Using the identity in Eq. (3.120), this is rewritten as

$$[FB(\boldsymbol{q}'')] = [FN(\boldsymbol{q})][BN(-\boldsymbol{q}')] \tag{3.125}$$

which then leads to the desired direct transformation from $\boldsymbol{q}$ and $\boldsymbol{q}'$ to the relative orientation vector $\boldsymbol{q}''$.

$$\boldsymbol{q}'' = \frac{\boldsymbol{q} - \boldsymbol{q}' + \boldsymbol{q} \times \boldsymbol{q}'}{1 + \boldsymbol{q} \cdot \boldsymbol{q}'} \tag{3.126}$$

A similar transformation could be found to express $\boldsymbol{q}'$ in terms of $\boldsymbol{q}$ and $\boldsymbol{q}''$.

The kinematic differential equation of the classical Rodrigues parameters is found by taking the derivative of Eq. (3.113) and then substituting the corresponding expressions for $\dot{\beta}_i$ given in Eq. (3.104). The resulting matrix formulation is[4]

$$\dot{\boldsymbol{q}} = \frac{1}{2} \begin{bmatrix} 1 + q_1^2 & q_1 q_2 - q_3 & q_1 q_3 + q_2 \\ q_2 q_1 + q_3 & 1 + q_2^2 & q_2 q_3 - q_1 \\ q_3 q_1 - q_2 & q_3 q_2 + q_1 & 1 + q_3^2 \end{bmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \tag{3.127}$$

and the compact vector matrix form is

$$\dot{\boldsymbol{q}} = \frac{1}{2} \left[[I_{3\times 3}] + [\tilde{\boldsymbol{q}}] + \boldsymbol{q}\boldsymbol{q}^T \right] \boldsymbol{\omega} \tag{3.128}$$

Note that the above kinematic differential equation contains no trigonometric functions and only has a quadratic nonlinearity. It is defined for any rotation except for $\Phi = \pm 180$ degrees. As $\boldsymbol{q}(t)$ approaches $\Phi = \pm 180^o$, both $\boldsymbol{q}(t)$ and $\dot{\boldsymbol{q}}(t)$ diverge to infinity. The inverse transformation of Eq. (3.128) is given by[5]

$$\boldsymbol{\omega} = \frac{2}{1 + \boldsymbol{q}^T \boldsymbol{q}} \left([I_{3\times 3}] - [\tilde{\boldsymbol{q}}] \right) \dot{\boldsymbol{q}} \tag{3.129}$$

As is evident, for $(\boldsymbol{q}, \dot{\boldsymbol{q}}) \rightarrow \infty$, the transformation of Eq. (3.129) exhibits an ∞/∞ type singular behavior near $|\Phi| \rightarrow \pm 180$ degrees.

There exists a very elegant, analytically exact transformation between the orthogonal direction cosine matrix $[C]$ and the classical Rodrigues parameter vector $\boldsymbol{q}$ called the Cayley Transform.[4, 5, 10, 17, 24] What is remarkable is that this transformation holds for proper orthogonal matrices of dimensions higher than three. A proper orthogonal matrix is an orthogonal matrix with a determinant of +1. Thus it is possible to parameterize any proper orthogonal $[C]$ matrix by a minimal set of higher-dimensional classical Rodrigues parameters.

The Cayley Transform parameterizes a proper orthogonal matrix $[C]$ as a function of a skew-symmetric matrix $[Q]$:

$$[C] = ([I] - [Q]) ([I] + [Q])^{-1} = ([I] + [Q])^{-1} ([I] - [Q]) \tag{3.130}$$

The matrix product order is irrelevant in this transformation. Another surprising property of this transformation is that the inverse transformation from the skew-symmetric matrix Q back to the $[C]$ matrix *has exactly the same form as the forward transformation* in Eq. (3.130):

$$[Q] = ([I] - [C]) ([I] + [C])^{-1} = ([I] + [C])^{-1} ([I] - [C]) \tag{3.131}$$

For the case where $[C]$ is a 3x3 rotation matrix, the transformation in Eq. (3.130) yields the standard three-dimensional Rodrigues parameters. This can be verified by setting $[Q] = [\tilde{\boldsymbol{q}}]$ in Eq. (3.130), carrying out the 3×3 special case algebra implicit in Eq. (3.130) and comparing the result to Eq. (3.118). The kinematic differential equation of the $[C]$ is given in Eq. (3.27). This expression also holds for matrix dimensions higher than three.[4, 10] Since the Cayley Transform parameterizes a proper orthogonal matrix in terms of an "orientation coordinate" type quantity $[Q]$, the matrix $[\tilde{\boldsymbol{\omega}}]$ represents an analogous "angular velocity" cross product matrix which can be defined as[4, 10, 17]

$$[\tilde{\boldsymbol{\omega}}] = 2 ([I] + [Q])^{-1} [\dot{Q}] ([I] - [Q])^{-1} \tag{3.132}$$

The kinematic differential equation of the higher dimensional Rodrigues parameters is obtained by differentiation of Eq. (3.131), and substituting Eqs. (3.27) and (3.130) as

$$[\dot{Q}] = \frac{1}{2} ([I] + [Q]) [\tilde{\boldsymbol{\omega}}] ([I] - [Q]) \tag{3.133}$$

It can readily be verified that the 3×3 special case of Eq. (3.133) is equivalent to Eq. (3.127); so again the general $n \times n$ case contains the classical 3×3 results.

Example 3.8: Given the orthogonal 4x4 matrix $[C]$,

$$[C] = \begin{bmatrix} 0.505111 & -0.503201 & -0.215658 & 0.667191 \\ 0.563106 & -0.034033 & -0.538395 & -0.626006 \\ 0.560111 & 0.748062 & 0.272979 & 0.228387 \\ -0.337714 & 0.431315 & -0.767532 & 0.332884 \end{bmatrix}$$

it is easy to verify that $[C]$ can be parameterized in terms of higher dimensional classical Rodrigues parameters. Using MATLAB to solve Eq. (3.131), the skew-symmetric 4x4 matrix $[Q]$ is found to be

$$[Q] = \begin{bmatrix} 0 & 0.5 & 0.2 & -0.3 \\ -0.5 & 0 & 0.7 & 0.6 \\ -0.2 & -0.7 & 0 & -0.4 \\ 0.3 & -0.6 & 0.4 & 0 \end{bmatrix}$$

where the six upper diagonal elements of $[Q]$ are the higher dimensional classical Rodrigues elements.

3.6 Modified Rodrigues Parameters

The Modified Rodrigues Parameters (MRPs) are an elegant recent addition to the family of attitude parameters.[5, 22, 25–27] The MRP vector $\boldsymbol{\sigma}$ is defined in terms of the Euler parameters as the transformation

$$\sigma_i = \frac{\beta_i}{1+\beta_0} \qquad i = 1, 2, 3 \tag{3.134}$$

The inverse transformation is given by

$$\beta_0 = \frac{1-\sigma^2}{1+\sigma^2} \quad \beta_i = \frac{2\sigma_i}{1+\sigma^2} \quad i = 1, 2, 3 \tag{3.135}$$

where the notation $\sigma^{2n} = \left(\boldsymbol{\sigma}^T\boldsymbol{\sigma}\right)^n$ is introduced. Substituting Eq. (3.90) into Eq. (3.134) the MRP can be expressed in terms of the principal rotation elements as

$$\boldsymbol{\sigma} = \tan\frac{\Phi}{4}\hat{\boldsymbol{e}} \tag{3.136}$$

Studying Eq. (3.136), it is evident that the MRP have a geometric singularity at $\Phi = \pm 360$ degrees. Any rotation can be described except a complete revolution back to the original orientation. This gives $\boldsymbol{\sigma}$ twice the rotational range of the classical Rodrigues parameters. Also note that for small rotations the MRPs linearize as $\boldsymbol{\sigma} \approx (\Phi/4)\,\hat{\boldsymbol{e}}$.

Observing Eq. (3.134) it is evident that these equations are well-behaved except near the singularity at $\beta_0 = -1$, where $\Phi \to \pm 360^{\circ}$. Also, the inverse

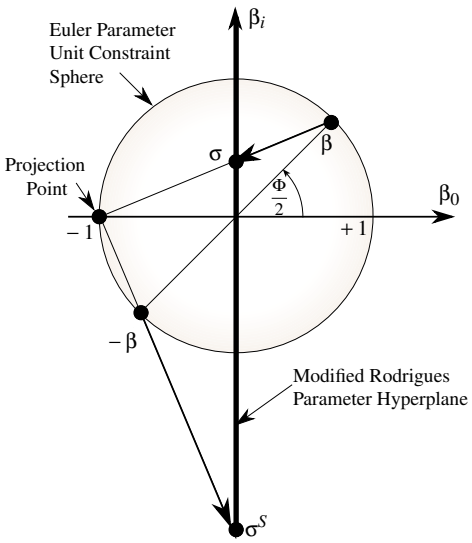


Figure 3.11: Stereographic Projection of Euler Parameters to Modified Rodrigues Parameters

transformation of Eq. (3.135) is well-behaved everywhere except at $|\boldsymbol{\sigma}| \to \infty$; we see from Eq. (3.136) that this again occurs at $\Phi \to 360^o$.

The MRP vector $\boldsymbol{\sigma}$ can be transformed directly into the classical Rodrigues parameter vector $\boldsymbol{q}$ through

$$\boldsymbol{q} = \frac{2\boldsymbol{\sigma}}{1 - \sigma^2} \tag{3.137}$$

with the inverse transformation being

$$\boldsymbol{\sigma} = \frac{\boldsymbol{q}}{1 + \sqrt{1 + \boldsymbol{q}^T \boldsymbol{q}}} \tag{3.138}$$

Naturally, these transformations are singular at $\Phi = \pm 180$ degrees since the classical Rodrigues parameters are singular at this orientation.

As are the classical Rodrigues parameters, the MRPs are also a particular set of stereographic orientation parameters. Equation (3.134) describes a stereographic projection of the Euler parameter unit sphere onto the MRP hyperplane normal to the β_0 axis at $\beta_0 = 0$, where the projection point is at $\boldsymbol{\beta} = (-1, 0, 0, 0)$. This is illustrated in Figure 3.11. As a ± 360 degree principal rotation is approached (i.e. $\beta_0 \to -1$), the projection of the corresponding point on the constraint sphere goes to infinity. This illustrates the singular behavior of the MRPs as they describe a complete revolution.

However, contrary to the classical Rodrigues parameters, the projection of the alternate Euler parameter vector $-\boldsymbol{\beta}$ results in a distinct set of shadow (or "image") MRPs as can be seen in Figure 3.11. Each MRP vector is an equally

valid attitude description satisfying the same kinematic differential equation. Therefore one can arbitrarily switch between the two vectors through the mapping[22, 26]

$$\sigma_i^S = \frac{-\beta_i}{1-\beta_0} = \frac{-\sigma_i}{\sigma^2} \quad i = 1,2,3 \tag{3.139}$$

where the choice as to which vector is the original and which the shadow vector is arbitrary. We usually let $\boldsymbol{\sigma}$ denote the mapping point interior to the unit sphere and $\boldsymbol{\sigma}^S$ the point point exterior to the unit sphere. As with the non-uniqueness of the principal rotation vector $\boldsymbol{\gamma}$ and the Euler parameter vector $\boldsymbol{\beta}$, one set of MRPs always corresponds to a principal rotation $\Phi \leq 180$ degrees and the other to $\Phi \geq 180$ degrees. From Eq. (3.136) it is clear that

$$\begin{aligned} |\boldsymbol{\sigma}| &\leq 1 \quad if \quad \Phi \leq 180^{\text{o}} \\ |\boldsymbol{\sigma}| &\geq 1 \quad if \quad \Phi \geq 180^{\text{o}} \\ |\boldsymbol{\sigma}| &= 1 \quad if \quad \Phi = 180^{\text{o}} \end{aligned} \tag{3.140}$$

The behavior is seen in Figure 3.11. The unit sphere $|\boldsymbol{\sigma}| = 1$, corresponding to all principal rotations of 180^{o} from the origin, is of particular importance. As one set of MRPs exits the unit sphere, the other (shadow) set enters. The mapping in Eq. (3.139) can be written in terms of the principal rotation elements using the definitions of β_i in Eq. (3.90) as

$$\boldsymbol{\sigma}^S = \tan\left(\frac{\Phi - 2\pi}{4}\right)\hat{\boldsymbol{e}} \tag{3.141}$$

Using Eq. (3.74) this can be written directly in terms of the alternate principal rotation angle Φ'.

$$\boldsymbol{\sigma}^S = \tan\left(\frac{\Phi'}{4}\right)\hat{\boldsymbol{e}} \tag{3.142}$$

Eq. (3.142) clearly shows that the shadow MRP vector is a direct result of the alternate principal rotation vector.

The shadow MRPs have a singular orientation at $\Phi = 0$ degrees as compared to the original MRPs, which are singular at $\Phi = \pm 360$ degrees. This allows one to avoid MRP singularities all together by switching between original and shadow MRP sets as one MRP vector approaches a singular orientation. On which surface $\boldsymbol{\sigma}^T\boldsymbol{\sigma} = c$ one switches is arbitrary. However, switching between the two MRPs whenever the vector $\boldsymbol{\sigma}$ penetrates the surface $\boldsymbol{\sigma}^T\boldsymbol{\sigma} = 1$ has many positive aspects. For one, the map between the two MRP vectors simplifies on this surface to $\boldsymbol{\sigma}^S = -\boldsymbol{\sigma}$. Further, the magnitude of $\boldsymbol{\sigma}$ will remain bounded above by 1. Having a bounded norm of an attitude description is useful since it reflects the fundamental fact that two orientations can only differ by a finite rotation. Also, the current MRP attitude description will always describe the shortest principal rotation because of Eq. (3.140). Therefore the combined set

of original and shadow MRPs with the switching surface $\boldsymbol{\sigma}^T\boldsymbol{\sigma} = 1$ provides for a nonsingular, bounded, minimal attitude description. It is ideally suited to describe large, arbitrary motions. The combined set is also useful in a feedback control type setting. For example, it linearizes well for small angles and has a bounded maximum norm of 1 which makes the selection of feedback gains easier.

The direction cosine matrix in terms of the MRP is found by substituting Eq. (3.135) into Eq. (3.92) and is given as[5, 22, 26, 27]

$$[C] = \frac{1}{(1+\sigma^2)^2}\begin{bmatrix} 4\left(\sigma_1^2-\sigma_2^2-\sigma_3^2\right)+(1-\sigma^2)^2 & 8\sigma_1\sigma_2+4\sigma_3(1-\sigma^2) & \cdots \\ 8\sigma_2\sigma_1-4\sigma_3(1-\sigma^2) & 4\left(-\sigma_1^2+\sigma_2^2-\sigma_3^2\right)+(1-\sigma^2)^2 & \\ 8\sigma_3\sigma_1+4\sigma_2(1-\sigma^2) & 8\sigma_3\sigma_2-4\sigma_1(1-\sigma^2) & \\ & 8\sigma_1\sigma_3-4\sigma_2(1-\sigma^2) & \\ \cdots & 8\sigma_2\sigma_3+4\sigma_1(1-\sigma^2) & \\ & 4\left(-\sigma_1^2-\sigma_2^2+\sigma_3^2\right)+(1-\sigma^2)^2 & \end{bmatrix} \tag{3.143}$$

In compact vector form $[C]$ is parameterized in terms of the MRP as[5, 22]

$$[C] = [I_{3\times 3}] + \frac{8[\tilde{\boldsymbol{\sigma}}]^2 - 4\left(1-\sigma^2\right)[\tilde{\boldsymbol{\sigma}}]}{\left(1+\sigma^2\right)^2} \tag{3.144}$$

As is the case with the classical Rodrigues parameters, the simplest method to extract the MRP from a given direction cosine matrix is the first extract the Euler parameters and then use Eq. (3.134) to find the MRP vector $\boldsymbol{\sigma}$. If $\beta_0 \geq 0$ is chosen when extracting the Euler parameters, then $|\boldsymbol{\sigma}| \leq 1$. If β_0 is chosen to be negative, then the alternate MRP vector corresponding to a larger principal rotation angle is found.

The MRPs enjoy the same relative rotation identity as did the classical Rodrigues parameters.

$$[C(\boldsymbol{\sigma})]^T = [C(-\boldsymbol{\sigma})] \tag{3.145}$$

Given two MRP vectors $\boldsymbol{\sigma}'$ and $\boldsymbol{\sigma}''$, let the overall MRP vector $\boldsymbol{\sigma}$ be defined through

$$[FN(\boldsymbol{\sigma})] = [FB(\boldsymbol{\sigma}'')][BN(\boldsymbol{\sigma}')] \tag{3.146}$$

Starting with the Euler parameter successive rotation property and using the MRP definitions in Eq. (3.134), the MRP successive rotation property is expressed as[5]

$$\boldsymbol{\sigma} = \frac{(1-|\boldsymbol{\sigma}'|^2)\boldsymbol{\sigma}'' + (1-|\boldsymbol{\sigma}''|^2)\boldsymbol{\sigma}' - 2\boldsymbol{\sigma}''\times\boldsymbol{\sigma}'}{1+|\boldsymbol{\sigma}'|^2|\boldsymbol{\sigma}''|^2 - 2\boldsymbol{\sigma}'\cdot\boldsymbol{\sigma}''} \tag{3.147}$$

Using Eq. (3.145), we are able to express the relative attitude vector $\boldsymbol{\sigma}''$ in terms of $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}'$ as

$$\boldsymbol{\sigma}'' = \frac{(1-|\boldsymbol{\sigma}'|^2)\boldsymbol{\sigma} - (1-|\boldsymbol{\sigma}|^2)\boldsymbol{\sigma}' + 2\boldsymbol{\sigma}\times\boldsymbol{\sigma}'}{1+|\boldsymbol{\sigma}'|^2|\boldsymbol{\sigma}|^2 + 2\boldsymbol{\sigma}'\cdot\boldsymbol{\sigma}} \tag{3.148}$$

While these expressions are more complicated than their Euler parameter or classical Rodrigues parameter counterparts, they do provide a numerically efficient method to compute the composition of two MRP vectors or find the relative MRP attitude vector.

Example 3.9: Given the Euler parameter vector $\boldsymbol{\beta}$

$$\boldsymbol{\beta} = (0.961798, -0.14565, 0.202665, 0.112505)^T$$

the MRP vector $\boldsymbol{\sigma}$ is found using Eq. (3.134)

$$\begin{aligned}
\sigma_1 &= \frac{-0.14565}{1 + 0.961798} = -0.0742431 \\
\sigma_2 &= \frac{0.202665}{1 + 0.961798} = 0.103306 \\
\sigma_3 &= \frac{0.112505}{1 + 0.961798} = 0.0573479
\end{aligned}$$

The alternate shadow MRP vector $\boldsymbol{\sigma}^S$ can be found using $-\boldsymbol{\beta}$ instead of $\boldsymbol{\beta}$ in Eq. (3.134).

$$\begin{aligned}
\sigma_1^S &= \frac{0.14565}{1 - 0.961798} = 3.81263 \\
\sigma_2^S &= \frac{-0.202665}{1 - 0.961798} = -5.30509 \\
\sigma_3^S &= \frac{-0.112505}{1 - 0.961798} = -2.945
\end{aligned}$$

Note that if the direct mapping in Eq. (3.139) is used the same vector $\boldsymbol{\sigma}^S$ is obtained. Since the vector $|\boldsymbol{\sigma}| = 0.139546 \leq 1$, it represents the shorter principal rotation angle of $\Phi = 7.94$ degrees. The vector $|\boldsymbol{\sigma}^S| = 7.16611 \geq 1$ represents the longer principal rotation angle $\Phi' = \Phi - 360° = -328.224°$.

The kinematic differential equation of the MRPs is found in a similar manner as the one for the classical Rodrigues parameters. The resulting matrix formulation is[22, 27]

$$\dot{\boldsymbol{\sigma}} = \frac{1}{4}\begin{bmatrix} 1 - \sigma^2 + 2\sigma_1^2 & 2(\sigma_1\sigma_2 - \sigma_3) & 2(\sigma_1\sigma_3 + \sigma_2) \\ 2(\sigma_2\sigma_1 + \sigma_3) & 1 - \sigma^2 + 2\sigma_2^2 & 2(\sigma_2\sigma_3 - \sigma_1) \\ 2(\sigma_3\sigma_1 - \sigma_2) & 2(\sigma_3\sigma_2 + \sigma_1) & 1 - \sigma^2 + 2\sigma_3^2 \end{bmatrix}\begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \tag{3.149}$$

The MRP kinematic differential equation in vector form is[5, 22]

$$\dot{\boldsymbol{\sigma}} = \frac{1}{4}\left[\left(1 - \sigma^2\right)[I_{3\times 3}] + 2[\tilde{\boldsymbol{\sigma}}] + 2\boldsymbol{\sigma}\boldsymbol{\sigma}^T\right]\boldsymbol{\omega} = \frac{1}{4}[B(\boldsymbol{\sigma})]\boldsymbol{\omega} \tag{3.150}$$

Note that the MRPs retain a kinematic differential equation very similar to the classical Rodrigues parameters with only quadratic nonlinearity present. This equation holds for either set of MRPs. However, the resulting vector $\dot{\boldsymbol{\sigma}}$ will

depend on which set of MRPs is being used. Just as a mapping exists between $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}^S$, a direct mapping between $\dot{\boldsymbol{\sigma}}$ and $\dot{\boldsymbol{\sigma}}^S$ is given by[28]

$$\dot{\boldsymbol{\sigma}}^S = -\frac{\dot{\boldsymbol{\sigma}}}{\sigma^2} + \frac{1}{2}\left(\frac{1+\sigma^2}{\sigma^4}\right)\boldsymbol{\sigma}\boldsymbol{\sigma}^T\boldsymbol{\omega} \tag{3.151}$$

Let the matrix $[B]$ transform $\boldsymbol{\omega}$ in Eqs. (3.149) and (3.150) into $\dot{\boldsymbol{\sigma}}$. Turns out that this $[B]$ matrix is almost orthogonal except for a generally non-unit scaling factor. The inverse of $[B]$ can be written as

$$[B]^{-1} = \frac{1}{\left(1+\sigma^2\right)^2}[B]^T \tag{3.152}$$

To prove Eq. (3.152) let's study the expression $[B]^T[B]$. Using Eq. (3.150) this is written as

$$[B]^T[B] = \left(\left(1-\sigma^2\right)[I_{3\times 3}] - 2[\tilde{\boldsymbol{\sigma}}] + 2\boldsymbol{\sigma}\boldsymbol{\sigma}^T\right)\left(\left(1-\sigma^2\right)[I_{3\times 3}] + 2[\tilde{\boldsymbol{\sigma}}] + 2\boldsymbol{\sigma}\boldsymbol{\sigma}^T\right)$$

After carrying out all the matrix multiplications the $[B]^T[B]$ expression is reduced to

$$[B]^T[B] = \left(1-\sigma^2\right)^2[I_{3\times 3}] - 4[\tilde{\boldsymbol{\sigma}}]^2 + 4\boldsymbol{\sigma}\boldsymbol{\sigma}^T$$

which can be further simplified using the identity $[\tilde{\boldsymbol{\sigma}}]^2 = \boldsymbol{\sigma}\boldsymbol{\sigma}^T - \sigma^2[I_{3\times 3}]$ to

$$[B]^T[B] = \left(1+\sigma^2\right)^2[I_{3\times 3}]$$

At this point it is trivial to verify that Eq. (3.152) must hold. The inverse transformation of Eqs. (3.149) and (3.150) then is in matrix notation

$$\boldsymbol{\omega} = \frac{4}{\left(1+\sigma^2\right)^2}[B]^T\dot{\boldsymbol{\sigma}} \tag{3.153}$$

and in vector form[5]

$$\boldsymbol{\omega} = \frac{4}{\left(1+\sigma^2\right)^2}\left[\left(1-\sigma^2\right)[I_{3\times 3}] - 2[\tilde{\boldsymbol{\sigma}}] + 2\boldsymbol{\sigma}\boldsymbol{\sigma}^T\right]\dot{\boldsymbol{\sigma}} \tag{3.154}$$

Like the classical Rodrigues parameters, the MRPs can also be used to minimally parameterize higher-dimensional proper orthogonal matrix $[C]$. Let the $[S]$ be a skew-symmetric matrix. The extended Cayley transform of $[C]$ in terms of $[S]$ is[17, 29]

$$[C] = \left([I_{3\times 3}] - [S]\right)^2\left(1+[S]\right)^{-2} = \left(1+[S]\right)^{-2}\left([I_{3\times 3}] - [S]\right)^2 \tag{3.155}$$

where the order of the matrix products is again irrelevant. For the case where $[C]$ is a 3x3 matrix, then $[S]$ is the same as $[\tilde{\boldsymbol{\sigma}}]$. Therefore Eq. (3.155) transforms a higher dimensional proper orthogonal $[C]$ into higher dimensional MRPs.

Unfortunately no direct inverse transformation exists like Eq. (3.131) for the higher order Cayley transforms.[17] The transformation is achieved indirectly through the matrix $[W]$, where it is defined as the matrix square root of $[C]$.

$$[C] = [W][W] \tag{3.156}$$

Since $[C]$ is orthogonal, it can be spectrally decomposed as

$$[C] = [V][D][V]^* \tag{3.157}$$

where $[V]$ is the orthogonal eigenvector matrix and $[D]$ is the diagonal eigenvalue matrix with entries of unit magnitude. The "$*$" operator stands for the adjoint operator which performs the complex conjugate transpose of a matrix. The matrix $[W]$ can be computed as

$$[W] = [V] \begin{bmatrix} \ddots & & 0 \\ & \sqrt{[D]_{ii}} & \\ 0 & & \ddots \end{bmatrix} [V]^T \tag{3.158}$$

The eigenvalues of $[C]$ are typically complex conjugate pairs. If the dimension of $[C]$ is odd, then the extra eigenvalue is real. For proper orthogonal matrices it is $+1$ and its square root is also chosen to be $+1$. The resulting $[W]$ matrix will then itself also be an proper orthogonal matrix. As Ref. 17 shows, the geometric interpretation of $[W]$ is that it represents the same "higher-dimensional" orientation as $[C]$ except that the corresponding principal rotation angles are halved.

The standard Cayley transforms in Eqs. (3.130) and (3.131) can be applied to map $[W]$ into $[S]$ and back.

$$[W] = ([I] - [S])([I] + [S])^{-1} = ([I] + [S])^{-1}([I] - [S]) \tag{3.159}$$

$$[S] = ([I] - [W])([I] + [W])^{-1} = ([I] + [W])^{-1}([I] - [W]) \tag{3.160}$$

Therefore, to obtain a higher-dimensional MRP representation of $[C]$, the matrix $[W]$ must be found first and then substituted into Eq. (3.160). Note that substituting Eq. (3.159) into Eq. (3.156) a direct forward transformation from $[S]$ to $[C]$ is found.

$$[C] = ([I] - [S])^2([I] + [S])^{-2} = ([I] + [S])^{-2}([I] - [S])^2 \tag{3.161}$$

The kinematic differential equations for $[S]$ are not written directly in terms of $[C]$ as they were for the classical Cayley transform. Instead the $[W]$ matrix is used. Being an orthogonal matrix, its kinematic differential equation is of the same form as Eq. (3.27)

$$[\dot{W}] = -[\tilde{\boldsymbol{\Omega}}][W] \tag{3.162}$$

where $[\tilde{\boldsymbol{\Omega}}]$ is the corresponding angular velocity matrix. It is related to the $[\tilde{\boldsymbol{\omega}}]$ matrix in Eq. (3.27) through

$$[\tilde{\boldsymbol{\omega}}] = [\tilde{\boldsymbol{\Omega}}] + [W][\tilde{\boldsymbol{\Omega}}][W]^T \tag{3.163}$$

Analogously to Eq. (3.133), the kinematic differential equation of the $[S]$ matrix is given by

$$[\dot{S}] = \frac{1}{2}\left([I] + [S]\right)[\tilde{\boldsymbol{\Omega}}]\left([I] - [S]\right) \tag{3.164}$$

Example 3.10: Consider the same orthogonal 4x4 matrix $[C]$ as is defined in Example 3.8. Using MATLAB, its matrix square root $[W]$ is found to be

$$[W] = \begin{bmatrix} 0.86416 & -0.35312 & -0.14580 & 0.32754 \\ 0.37209 & 0.69343 & -0.44177 & -0.43076 \\ 0.25488 & 0.50816 & 0.79065 & 0.22734 \\ -0.22320 & 0.36911 & -0.39807 & 0.80962 \end{bmatrix}$$

Using Eq. (3.160) the higher dimensional, skew-symmetric MRP matrix $[S]$ representing $[C]$ is found.

$$[S] = \begin{bmatrix} 0 & 0.20952 & 0.10114 & -0.14383 \\ -0.20952 & 0 & 0.28309 & 0.24040 \\ -0.10114 & -0.28309 & 0 & -0.17471 \\ 0.14383 & -0.24040 & 0.17471 & 0 \end{bmatrix}$$

By back substitution of this $[S]$ into Eq. (3.155) it can be verified that it does indeed parameterize $[C]$.

3.7 Other Attitude Parameters

There exists a multitude of other attitude parameters sets in addition to those discussed so far. This section will briefly outline a selected few.

3.7.1 Stereographic Orientation Parameters

The Stereographic Orientation Parameters (SOPs) are introduced in Ref. 22. They are formed by projecting the Euler parameter constraint surface, a four-dimensional unit hypersphere, onto a three-dimensional hyperplane. The projection point can be anywhere on or within the constraint hypersphere, while the mapping hyperplane is chosen to be a unit distance away from the projection point.

There are two types of SOPs, the symmetric and asymmetric sets. The symmetric sets have a mapping hyperplane that is perpendicular to the β_0 axis. Since $\beta_0 = \cos\Phi/2$ only contains information about the principal rotation angle, the resulting sets will all have a geometric singularity at a specific principal

rotation angle Φ only, regardless of the corresponding principal rotation axis $\hat{\boldsymbol{e}}$. The classical and modified Rodrigues parameters are examples of symmetric SOPs.

Asymmetric SOPs have a mapping hyperplane which is not perpendicular to the β_0 axis. The condition for a geometric singularity will now depend on both the principal rotation axis $\hat{\boldsymbol{e}}$ and the angle Φ. As an example, consider the asymmetric SOP vector $\boldsymbol{\eta}$. It is formed by having a projection point at $\beta_1 = -1$ and having a mapping hyperplane at $\beta_1 = 0$. In terms of the Euler parameters it is defined as

$$\eta_1 = \frac{\beta_0}{1+\beta_1} \quad \eta_2 = \frac{\beta_2}{1+\beta_1} \quad \eta_3 = \frac{\beta_3}{1+\beta_1} \tag{3.165}$$

with the inverse transformation being

$$\beta_0 = \frac{2\eta_1}{1+\eta^2} \quad \beta_1 = \frac{1-\eta^2}{1+\eta^2} \quad \beta_2 = \frac{2\eta_2}{1+\eta^2} \quad \beta_3 = \frac{2\eta_3}{1+\eta^2} \tag{3.166}$$

where $\eta^2 = \boldsymbol{\eta}^T\boldsymbol{\eta}$. From Eq. (3.165) it is evident that $\boldsymbol{\eta}$ has a geometric singularity whenever $\beta_1 \to -1$. This means that $\boldsymbol{\eta}$ goes singular whenever it represents a pure single-axis rotation about the first body axis by the principal angles Φ_1 = -180 degrees or Φ_2 = +540 degrees. This type of asymmetric principal angle rotation range is typical for all asymmetric SOPs. However, since the $\boldsymbol{\eta}$ vector has a distinct shadow counter part, any geometric singularities can be avoided by switching between the two sets through the mapping

$$\boldsymbol{\eta}^S = -\frac{\boldsymbol{\eta}}{\eta^2} \tag{3.167}$$

The direction cosine matrix is written in terms of the $\boldsymbol{\eta}$ vector components as

$$[C] = \tfrac{1}{(1+\eta^2)^2}\begin{bmatrix} 4\left(\eta_1^2-\eta_2^2-\eta_3^2\right)+(1-\eta^2)^2 & 8\eta_1\eta_3+4\eta_2(1-\eta^2) & \cdots \\ -8\eta_1\eta_3+4\eta_2(1-\eta^2) & 4\left(\eta_1^2+\eta_2^2-\eta_3^2\right)-(1-\eta^2)^2 & \cdots \\ 8\eta_1\eta_2+4\eta_3(1-\eta^2) & 8\eta_2\eta_3-4\eta_1(1-\eta^2) & \cdots \\ & -8\eta_1\eta_2+4\eta_3(1-\eta^2) & \\ & 8\eta_2\eta_3+4\eta_1(1-\eta^2) & \\ & 4\left(\eta_1^2-\eta_2^2+\eta_3^2\right)-(1-\eta^2)^2 & \end{bmatrix} \tag{3.168}$$

The kinematic differential equation of the $\boldsymbol{\eta}$ vector is

$$\dot{\boldsymbol{\eta}} = \frac{1}{4}\begin{bmatrix} -1-2\eta_1^2+\eta^2 & 2\left(\eta_1\eta_3-\eta_2\right) & -2\left(\eta_1\eta_2+\eta_3\right) \\ 2\left(\eta_3-\eta_1\eta_2\right) & 2\left(\eta_2\eta_3+\eta_1\right) & -1-2\eta_2^2+\eta^2 \\ -2\left(\eta_1\eta_3+\eta_2\right) & 1+2\eta_3^2-\eta^2 & 2\left(\eta_1-\eta_2\eta_3\right) \end{bmatrix}\boldsymbol{\omega} \tag{3.169}$$

Having a projection point on the constraint surface provides for the largest possible range of singularity free rotations. This is evident when comparing the classical and the modified Rodrigues parameters. The classical Rodrigues parameters have a projection point within the constraint hypersphere at $\beta_0 = 0$. Their principal rotation range is half of that of the MRPs whose projection is on the constraint surface at $\beta_0 = -1$.

3.7.2 Higher Order Rodrigues Parameters

The Higher Order Rodrigues Parameters (HORP) are introduced in Ref. 29. The classical Cayley transform in Eq. (3.130) is expanded such that it parameterized nxn orthogonal matrices through a skew-symmetric, higher order Rodrigues parameter matrix X.

$$[C] = \left([I_{3\times 3}] - X\right)^m \left([I_{3\times 3}] + X\right)^{-m} \tag{3.170}$$

The corresponding attitude vector $\boldsymbol{x}$ is given by

$$\boldsymbol{x} = \tan\left(\frac{\Phi}{2m}\right)\hat{\boldsymbol{e}} \tag{3.171}$$

For $m = 1$ the vector $\boldsymbol{x}$ is the classical Rodrigues vector and for $m = 2$ it is the MRP vector. Note that the domain of validity of the $\boldsymbol{x}$ vector is $|\Phi| < m\pi$. The HORP sets are generally also not unique as is the case with the MRPs. Corresponding "shadow" sets can be used here too to avoid any geometric singularities. Note that for a given m there are typically m sets of possible HORPs.

A particular set of HORP is the $\boldsymbol{\tau}$ vector where $m = 4$. In terms of the Euler parameters, the first two HORP vectors $\boldsymbol{\tau}$ are defined through

$$\tau_i = \frac{\beta_i}{1 + \beta_0 \pm \sqrt{2\left(1 + \beta_0\right)}} \quad i = 1, 2, 3 \tag{3.172}$$

with the inverse transformation being

$$\beta_0 = 2\left(\frac{1-\tau^2}{1+\tau^2}\right)^2 - 1 \quad \beta_i = \frac{4\tau_i\left(1-\tau^2\right)}{\left(1+\tau^2\right)^2} \quad i = 1, 2, 3 \tag{3.173}$$

where $\tau^{2n} = \left(\boldsymbol{\tau}^T\boldsymbol{\tau}\right)^n$. Each vector τ defined in Eq. (3.172) can be mapped to the corresponding shadow vector τ^S through

$$\boldsymbol{\tau}^S = -\boldsymbol{\tau}\left(\frac{1-\tau^2}{2\tau^2 + \left(1+\tau^2\right)\tau}\right) \tag{3.174}$$

where $\tau = \sqrt{\tau^2}$. Combined Eqs. (3.172) and (3.174) yield the four possible HORP vectors for $m = 4$. In terms of the principal rotation elements, the four sets can be expressed as

$$\boldsymbol{\tau} = \tan\left(\frac{\Phi - 2k\pi}{8}\right) \quad k = 0, 1, 2, 3 \tag{3.175}$$

Therefore it will always be possible to switch from one $\boldsymbol{\tau}$ vector to another in order to avoid geometric singularities.

The kinematic differential equations of the $\boldsymbol{\tau}$ vector are

$$\dot{\boldsymbol{\tau}} = \frac{1}{8\left(1-\tau^2\right)}\left[2\left(3-\tau^2\right)\boldsymbol{\tau}\boldsymbol{\tau}^T + 4\left(1-\tau^2\right)[\tilde{\boldsymbol{\tau}}] + (1-6\tau^2+\tau^4)[I_{3\times 3}]\right]\omega \tag{3.176}$$

Note that the kinematic differential equations of the HORP lose the simple second order polynomial form that is present for the classical and modified Rodrigues parameters. Also, while the $\boldsymbol{\tau}$ vector itself is defined for rotations up to $\Phi = m\pi$, the kinematic differential equations encounter mathematical singularities of the type 0/0 whenever $\tau^2 \to 0$. This corresponds to $\Phi \to \pm 360$ degrees. By using the mapping in Eq. (3.174) to transform a τ vector to an alternate set whenever $|\boldsymbol{\tau}| \geq \tan(\Phi/8)$ any geometrical and mathematical singularities are avoided all together.

3.7.3 The (w, z) Coordinates

The (w, z) attitude coordinates were introduced by Tsiotras and Longuski in Ref. 30. They are a minimal coordinate set and lend themselves well to be used in control problems of under actuated axially-symmetric spacecraft.[31] The complex coordinate w describes the heading of the one of the body axes, typically the spin axis. The coordinate z is the relative rotation angle about this axis defined by w. Let the heading of the chosen body axis be given by the vector $\hat{\boldsymbol{b}}_i = (a, b, c)^T$. Since the vector $\hat{\boldsymbol{b}}_i$ is a unit vector, the three components a, b and c are not independent. They must satisfy the constraint sphere equation

$$a^2 + b^2 + c^2 = 1 \tag{3.177}$$

By performing a stereographic projection of the constraint sphere from the projection point $(0, 0, -1$ onto the complex (w_1, w_2) plane, the three redundant axis heading coordinates (a, b, c) are reduced to the complex variable w.

$$w = w_1 + iw_2 = \frac{b - ia}{1 + c} \tag{3.178}$$

The inverse transformation from w to (a, b, c) is given by

$$a = \frac{i\,(w - \bar{w})}{1 + |w|^2} \quad b = \frac{w + \bar{w}}{1 + |w|^2} \quad c = \frac{1 - |w|^2}{1 + |w|^2} \tag{3.179}$$

Let's assume that the spin axis is the third body axis, then the direction cosine matrix in terms of (w, z) is given by

$$[C] = \frac{1}{1 + |w|^2} \begin{bmatrix} Re\left[\left(1 + w^2\right) e^{iz}\right] & Im\left[\left(1 + w^2\right) e^{iz}\right] & -2Im(w) \\ Im\left[\left(1 - \bar{w}^2\right) e^{-iz}\right] & Re\left[\left(1 - \bar{w}^2\right) e^{-iz}\right] & 2Re(w) \\ 2Im(we^{iz}) & -2Re(we^{iz}) & 1 - |w|^2 \end{bmatrix} \tag{3.180}$$

The kinematic differential equations of the (w, z) coordinates are given by

$$\dot{w}_1 = \omega_3 w_2 + \omega_2 w_1 w_2 + \frac{\omega_1}{2}\left(1 + w_1^2 - w_2^2\right) \tag{3.181a}$$

$$\dot{w}_2 = -\omega_3 w_1 + \omega_1 w_1 w_2 + \frac{\omega_2}{2}\left(1 + w_2^2 - w_1^2\right) \tag{3.181b}$$

$$\dot{z} = \omega_3 - \omega_1 w_2 \omega_2 w_1 \tag{3.181c}$$

3.7.4 Cayley-Klein Parameters

The Cayley-Klein parameters are a set of four complex parameters which are closely related to the Euler parameter vector $\boldsymbol{\beta}$. They form a once-redundant, non-singular set of attitude parameters. Let $i = \sqrt{-1}$, then they are defined in terms of $\boldsymbol{\beta}$ as[16]

$$\begin{array}{ll} \alpha = \beta_0 + i\beta_3 & \beta = -\beta_2 + i\beta_1 \\ \gamma = \beta_2 + i\beta_1 & \delta = \beta_0 - i\beta_3 \end{array} \tag{3.182}$$

The inverse transformation from the Euler parameters to the Cayley-Klein parameters is

$$\begin{array}{ll} \beta_0 = (\alpha + \delta)/2 & \beta_1 = -i(\beta + \gamma)/2 \\ \beta_2 = -(\beta - \gamma)/2 & \beta_3 = -i(\alpha - \delta)/2 \end{array} \tag{3.183}$$

The direction cosine matrix is parameterized by the Cayley-Klein parameters as

$$[C] = \begin{bmatrix} \left(\alpha^2 - \beta^2 - \gamma^2 + \delta^2\right)/2 & i\left(-\alpha^2 + \beta^2 - \gamma^2 + \delta^2\right)/2 & (\beta\delta - \alpha\gamma) \\ i\left(\alpha^2 + \beta^2 - \gamma^2 - \delta^2\right)/2 & \left(\alpha^2 + \beta^2 + \gamma^2 + \delta^2\right)/2 & -i(\alpha\gamma + \beta\delta) \\ (\gamma\delta - \alpha\beta) & i(\alpha\beta + \gamma\delta) & (\alpha\delta + \beta\gamma) \end{bmatrix} \tag{3.184}$$

3.8 Homogeneous Transformations

All previous sections in this chapter deal with methods to describe the relative orientation of one coordinate frame to another. In particular, the direction cosine matrix is a convenient tool to map a vector with components taken in one reference frame to a vector with components taken in another. However, one underlying assumption here is that both reference frames have the same origin. In other words, any translational differences between the two frames in questions is not taken into account when the vector components are mapped from one frame to another.

Figure 3.12 shows an illustration where two coordinates frames differ both in orientation and in their origins. Let us define the following two reference fame $\mathcal{N}$ and $\mathcal{B}$.

$$\begin{aligned} \mathcal{N} &: \{\mathcal{O}_\mathcal{N}, \hat{\boldsymbol{n}}_1, \hat{\boldsymbol{n}}_2, \hat{\boldsymbol{n}}_3\} \\ \mathcal{B} &: \{\mathcal{O}_\mathcal{B}, \hat{\boldsymbol{b}}_1, \hat{\boldsymbol{b}}_2, \hat{\boldsymbol{b}}_3\} \end{aligned}$$

Let the position vector from the $\mathcal{N}$ frame origin to the $\mathcal{B}$ frame origin be given by $\boldsymbol{r}_{\mathcal{B}/\mathcal{N}}$. The position vector of point P is expressed in $\mathcal{B}$ frame components as ${}^{\mathcal{B}}\boldsymbol{r}_p$. These vector components are mapped into $\mathcal{N}$ frame components by pre-multiplying by the direction cosine matrix $[NB]$. While this provides the correct $\mathcal{N}$ frame components of the vector $\boldsymbol{r}_P$, it does not provide the correct

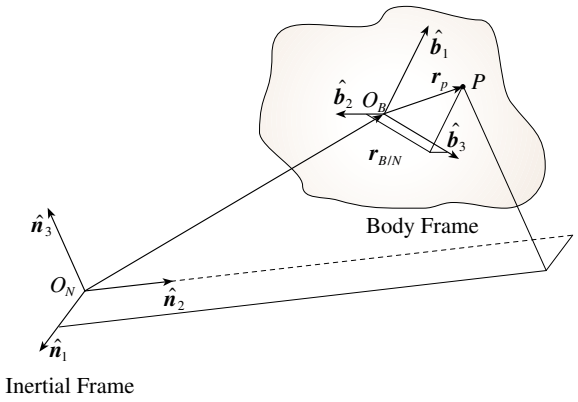


Figure 3.12: Illustration of two Coordinate Frames with Different Origins and Orientations

position vector of point P as seen by the $\mathcal{N}$ frame since the two frames have different origins. To obtain these vector components, we compute

$$^{\mathcal{N}}\boldsymbol{r}_p = {}^{\mathcal{N}}\boldsymbol{r}_{\mathcal{B}/\mathcal{N}} + [NB]\,{}^{\mathcal{B}}\boldsymbol{r}_p \tag{3.185}$$

By defining the 4×4 *homogeneous transformation*[32]

$$[\mathcal{N}\mathcal{B}] = \begin{bmatrix} NB & {}^{\mathcal{N}}\boldsymbol{r}_{\mathcal{B}/\mathcal{N}} \\ 0_{1\times 3} & 1 \end{bmatrix} \tag{3.186}$$

it is possible to transform the position vector taken in $\mathcal{B}$ frame components directly into the corresponding position vector in $\mathcal{N}$ frame components. In robotics literature, this transformation is typically referred to as ${}^{\mathcal{N}}_{\mathcal{B}}T$. To accomplish this, we define the 4×1 position vector

$${}^{\mathcal{B}}\boldsymbol{p} = \begin{bmatrix} {}^{\mathcal{B}}\boldsymbol{r}_p \\ 1 \end{bmatrix} \tag{3.187}$$

Observing Eq. (3.185), it is clear that

$${}^{\mathcal{N}}\boldsymbol{p} = [\mathcal{N}\mathcal{B}]\,{}^{B}\boldsymbol{p} \tag{3.188}$$

This formula is very convenient when computing the position coordinate of a chain of bodies such as are typically found in robotics applications. However, care must be taken when considering the order of the translational and rotational differences between the two frames. The homogenous transformation, as shown in Eq. (3.186), performs the translation first and the rotation second. This order is important. Assume a rotational joint has a telescoping member attached to it. To compute the homogeneous transformation from the joint to the telescoping member tip, a rotation must be performed first and a translation second.

Note that this homogenous transformation matrix abides by the same successive transformation property as the direction cosine matrix does. Consider the two vectors

$$^{\mathcal{A}}\boldsymbol{p} = [\mathcal{AB}]^{\mathcal{B}}\boldsymbol{p} \tag{3.189}$$

$$^{\mathcal{N}}\boldsymbol{p} = [\mathcal{NA}]^{\mathcal{A}}\boldsymbol{p} \tag{3.190}$$

Substituting Eq. (3.189) into (3.190), we find that

$$^{\mathcal{N}}\boldsymbol{p} = [\mathcal{NA}][\mathcal{AB}]^{\mathcal{B}}\boldsymbol{p} = [\mathcal{NB}]^{\mathcal{B}}\boldsymbol{p} \tag{3.191}$$

Thus, two successive transformations are combined through

$$[\mathcal{NB}] = [\mathcal{NA}][\mathcal{AB}] \tag{3.192}$$

However, the inverse matrix formula for the homogeneous transformation is not quite as elegant as the matrix inverse of the orthogonal direction cosine matrix. The following partitioned matrix inverse is convenient to compute the inverse of the $[\mathcal{NB}]$. Let $[M]$ be defined as[4]

$$[M] = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \tag{3.193}$$

Then the inverse is given by

$$[M]^{-1} = \begin{bmatrix} A^{-1} + A^{-1}B\Delta^{-1}CA^{-1} & -A^{-1}B\Delta^{-1} \\ -\Delta^{-1}CA^{-1} & \Delta^{-1} \end{bmatrix} \tag{3.194}$$

with the Schur complement being defined as

$$\Delta = D - CA^{-1}B \tag{3.195}$$

Substituting

$$[A] = [NB] \qquad [B] = [^{\mathcal{N}}\boldsymbol{r}_{\mathcal{B}/\mathcal{N}}]$$
$$[C] = [0_{1\times 3}] \qquad [D] = [1]$$

the Schur complement is given by

$$\Delta = [1] \tag{3.196}$$

and the inverse of the homogeneous transformation is the remarkable simply formula:

$$[\mathcal{NB}] = \begin{bmatrix} [NB]^T & -[NB]^T\, ^{\mathcal{N}}\boldsymbol{r}_{\mathcal{B}/\mathcal{N}} \\ 0_{1\times 3} & 1 \end{bmatrix} \tag{3.197}$$

Here the fact was used that $[NB]$ is orthogonal and that $[NB]^{-1} = [NB]^T$.

Problems

3.1 Given three reference frames $\mathcal{N}$, $\mathcal{B}$ and $\mathcal{F}$, let the unit base vectors of the reference frames $\mathcal{B}$ and $\mathcal{F}$ be

$$\hat{\boldsymbol{b}}_1 = \frac{1}{3}\begin{pmatrix}1\\2\\-2\end{pmatrix} \quad \hat{\boldsymbol{b}}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\1\end{pmatrix} \quad \hat{\boldsymbol{b}}_3 = \frac{1}{3\sqrt{2}}\begin{pmatrix}4\\-1\\1\end{pmatrix}$$

and

$$\hat{\boldsymbol{f}}_1 = \frac{1}{4}\begin{pmatrix}3\\-2\\\sqrt{3}\end{pmatrix} \quad \hat{\boldsymbol{f}}_2 = \frac{1}{2}\begin{pmatrix}-1\\0\\\sqrt{3}\end{pmatrix} \quad \hat{\boldsymbol{f}}_3 = \frac{-1}{4}\begin{pmatrix}\sqrt{3}\\2\sqrt{3}\\1\end{pmatrix}$$

where the base vector components are written in the $\mathcal{N}$ frame. Find the direction cosine matrices $[BF]$ that describes the orientation of the $\mathcal{B}$ frame relative to the $\mathcal{F}$ frame, along with the direction cosine matrices $[BN]$ and $[FN]$ that map vectors in the $\mathcal{N}$ frame into respective $\mathcal{B}$ or $\mathcal{F}$ frame vectors.

3.2 Let the vector $\boldsymbol{v}$ be written in $\mathcal{B}$ frame components as

$${}^{\mathcal{B}}\boldsymbol{v} = 1\hat{\boldsymbol{b}}_1 + 2\hat{\boldsymbol{b}}_2 - 3\hat{\boldsymbol{b}}_3$$

The orientation of the $\mathcal{B}$ frame relative to the $\mathcal{N}$ frame is given through the direction cosine matrix

$$[BN] = \begin{bmatrix}-0.87097 & 0.45161 & 0.19355\\-0.19355 & -0.67742 & 0.70968\\0.45161 & 0.58065 & 0.67742\end{bmatrix}$$

a) Find the direction cosine matrix $[NB]$ that maps vectors with components in the $\mathcal{B}$ frame into a vector with $\mathcal{N}$ frame components.

b) Find the $\mathcal{N}$ frame components of the vector $\boldsymbol{v}$.

3.3 Using the direction cosine matrix $[BN]$ in Problem 3.2, find its *real* eigenvalue and corresponding eigenvector.

3.4 ♣ The angular velocity vectors of a spacecraft $\mathcal{B}$ and a reference frame motion $\mathcal{R}$ relative to the inertial frame $\mathcal{N}$ are given by $\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}$ and $\boldsymbol{\omega}_{\mathcal{R}/\mathcal{N}}$. The vector $\boldsymbol{\omega}_{\mathcal{R}/\mathcal{N}}$ is given in $\mathcal{R}$ frame components, while $\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}$ is given in $\mathcal{B}$ frame components. The error angular velocity vector of the spacecraft relative to the reference motion is then given by $\delta\boldsymbol{\omega} = \boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}} - \boldsymbol{\omega}_{\mathcal{R}/\mathcal{N}}$. Find the relative error angular acceleration vector $\delta\dot{\boldsymbol{\omega}}$ with components expressed in the $\mathcal{B}$ frame.

a) Find $\delta\dot{\boldsymbol{\omega}}$ by only assigning vector frames at the last step.

b) Find $\delta\dot{\boldsymbol{\omega}}$ by first expressing $\delta\boldsymbol{\omega}$ in $\mathcal{B}$ frame components as ${}^{\mathcal{B}}\delta\boldsymbol{\omega} = {}^{\mathcal{B}}\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}} - [BR]{}^{\mathcal{R}}\boldsymbol{\omega}_{\mathcal{R}/\mathcal{N}}$ and then performing an inertial derivative.

3.5 The reference frames $\mathcal{N} : \{\hat{\boldsymbol{n}}_1, \hat{\boldsymbol{n}}_2, \hat{\boldsymbol{n}}_3\}$ and $\mathcal{B} : \{\hat{\boldsymbol{b}}_L, \hat{\boldsymbol{b}}_\theta, \hat{\boldsymbol{b}}_r\}$ are shown in Figure 3.13.

a) Find the direction cosine matrix $[BN]$ in terms of the angle ϕ.

b) Given the vector ${}^{\mathcal{B}}\boldsymbol{v} = 1\hat{\boldsymbol{b}}_r + 1\hat{\boldsymbol{b}}_\theta + 2\hat{\boldsymbol{b}}_L$, find the vector ${}^{\mathcal{N}}\boldsymbol{v}$.

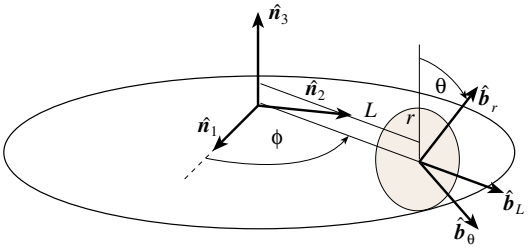


Figure 3.13: Disk Rolling on Circular Ring

3.6 Starting with Eqs. (3.21) and (3.22), verify Eq. (3.24).

3.7 Parameterize the direction cosine matrix $[C]$ in terms of (2-3-2) Euler angles. Also, find appropriate inverse transformations from $[C]$ back to the (2-3-2) Euler angles.

3.8 Find the kinematic differential equations of the (2-3-2) Euler angles. What is the geometric condition for which these equations will encounter a mathematical singularity.

3.9 Given the (3-2-1) Euler angles $\psi = 10°$, $\theta = -15°$ and $\phi = 20°$ and their rates $\dot{\psi} = 2°/s$, $\dot{\theta} = 1°/s$ and $\dot{\phi} = 0°/s$, find the vectors ${}^{\mathcal{B}}\boldsymbol{\omega}$ and ${}^{\mathcal{N}}\boldsymbol{\omega}$.

3.10 The orientation of an object is given in terms of the (3-1-3) Euler angles $(-30°, 40°, 20°)$,

a) find the corresponding principal rotation axis $\hat{\boldsymbol{e}}$

b) find the two principal rotation angles Φ and Φ'

3.11 A spacecraft performs a $45°$ single axis rotation about $\hat{\boldsymbol{e}} = \frac{1}{\sqrt{3}}(1,1,1)^T$. Find the corresponding (3-2-1) yaw, pitch and roll angles that relate the final attitude to the original attitude.

3.12 Verify that the exponential matrix mapping $[C] = e^{-\Phi[\tilde{\boldsymbol{e}}]}$ does have the finite form given in Eq. (3.81).

3.13 Verify that Eq. (3.89) is indeed the inverse mapping of the differential kinematic equation of $\dot{\boldsymbol{\gamma}}$ given in Eq. (3.88).

3.14 ♣ Starting from the direction cosine matrix $[C]$ in Eq. (3.71) written in terms of the principal rotation elements, derive the parameterization of $[C]$ in terms of the Euler parameters.

3.15 ♣ Derive the composite rotation property of the Euler parameter vector given in Eqs. (3.97) and (3.98).

3.16 Derive the kinematic differential equations for the second, third and fourth Euler parameter.

3.17 Verify the transformation in Eq. (3.114) which maps a classical Rodrigues parameter vector into an Euler parameter vector.

3.18 Show the details of transforming the classical Rodrigues parameter definition in terms of the Euler parameters $q_i = \beta_i/\beta_0$ into the expression $q_i = \tan\frac{\Phi}{2}\hat{e}_i$ which is in terms of the principal rotation elements.

3.19 ♣ Show that the classical Rodrigues parameters are indeed a stereographic projection of the Euler parameter constraint surface (a four-dimensional unit hypersphere) onto the three-dimensional hyperplane tangent to $\beta_0 = 1$ with the projection point being $\boldsymbol{\beta} = (0, 0, 0, 0)^T$.

3.20 Given the classical Rodrigues parameter vector $\boldsymbol{q} = (0.5, -0.2, 0.8)^T$. Use the Cayley transform in Eq. (3.130) to find the corresponding direction cosine matrix $[C]$. Also, verify that this $[C]$ is the same as is obtained through the mapping in Eq. (3.118) or (3.119).

3.21 Verify the transformation in Eq. (3.135) which maps a MRP vector into an Euler parameter vector.

3.22 Show the details of transforming the MRP definition in terms of the Euler parameters $\sigma_i = \beta_i/(1 + \beta_0)$ into the expression $\sigma_i = \tan\frac{\Phi}{4}\hat{e}_i$ which is in terms of the principal rotation elements.

3.23 ♣ Show that the MRPs are a stereographic projection of the Euler parameter constraint surface (a four-dimensional unit hypersphere) onto the three-dimensional hyperplane tangent to $\beta_0 = 0$ with the projection point being $\boldsymbol{\beta} = (-1, 0, 0, 0)^T$.

3.24 Derive the MRP parameterization of the direction cosine matrix $[C]$ given in Eq. (3.143).

3.25 Let the initial attitude vector be given through the MRP vector $\boldsymbol{\sigma}(t_0) = (0, 0, 0)^T$. The body angular velocity vector $\boldsymbol{\omega}(t)$ is given as $(1, 0.5, -0.7)^T$ rad/second. Integrate the resulting rotation for 5 seconds and use the mapping between "original" and "shadow" MRPs in Eq. (3.139) to enforce $|\boldsymbol{\sigma}| \leq 1$.

3.26 ♣ Derive the mapping between $\dot{\boldsymbol{\sigma}}$ and its shadow counter part $\dot{\boldsymbol{\sigma}}^S$ in Eq. (3.151) starting with the kinematic differential equation of the MRP in Eq. (3.150) and Eq. (3.139).

3.27 Given the MRP vector $\boldsymbol{\sigma} = (-0.25, -0.4, 0.3)^T$. Use the Cayley transform in Eq. (3.161) to find the corresponding direction cosine matrix $[C]$. Also, verify that this $[C]$ is the same as is obtained through the mapping in Eq. (3.143) or (3.144).

Bibliography

[1] Junkins, J. L. and Turner, J. D., *Optimal Spacecraft Rotational Maneuvers*, Elsevier Science Publishers, Amsterdam, Netherlands, 1986.

[2] Morton, H. S. and Junkins, J. L., *The Differential Equations of Rotational Motion*, 1986, In Preperation.

[3] Kaplan, W., *Advanced Calculus*, Addison-Wesley Publishing Company, Inc., New York, 4th ed., 1991.

[4] Junkins, J. L. and Kim, Y., *Introduction to Dynamics and Control of Flexible Structures*, AIAA Education Series, Washington D.C., 1993.

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

[6] Rugh, W. J., *Linear System Theory*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1993.

[7] Bowen, R. M. and Wang, C.-C., *Introduction to Vectors and Tensors*, Vol. 1, Plenum Press, New York, 1976.

[8] Goldstein, H., *Classical Mechanics*, Addison-Wesley, 1950.

[9] Likins, P. W., *Elements of Engineering Mechanics*, McGraw-Hill, New York, 1973.

[10] Bar-Itzhack, I. Y. and Markley, F. L., "Minimal Parameter Solution of the Orthogonal Matrix Differential Equation," *IEEE Transactions on Automatic Control*, Vol. 35, No. 3, March 1990, pp. 314–317.

[11] Nelson, R. C., *Flight Stability and Automatic Control*, McGraw-Hill, Inc., New York, 1989.

[12] Battin, R. H., *An Introduction to the Mathematics and Methods of Astrodynamics*, AIAA Education Series, New York, 1987.

[13] Junkins, J. L. and Shuster, M. D., "The Geometry of Euler Angles," *Journal of the Astronautical Sciences*, Vol. 41, No. 4, 1993, pp. 531–543.

[14] Euler, L., "Problema Algebraicum of Affectiones Psorsus Singulares Memorabile," .

[15] Rodriques, O., "Des Lois Geometriques qui Regissent Les Deplacements D'Un Systeme Solide Dans l'Espace, et de la Variation des Coordonnes Provenants de ces Deplacements Considers Independamment des Causes Qui Peuvent les Preduire," *LIOUV*, Vol. III, 1840, pp. 380–440.

[16] Whittaker, E. T., *Analytical Dynamics of Particles and Rigid Bodies*, Cambridge University Press, 1965 reprint, pp. 2–16.

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

[18] Nazaroff, G. J., "The Orientation Vector Differential Equation," *Journal of Guidance and Control*, Vol. 2, 1979, pp. 351–352.

[19] Jiang, Y. F. and Lin, Y. P., "On the Rotation Vector Differential Equation," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-27, 1991, pp. 181–183.

[20] Bharadwaj, S., Osipchuk, M., Mease, K. D., and Park, F. C., "Geometry and Optimality in Global Attitude Stabilization," *submitted to Journal of Guidance, Control and Dynamics*, July 1997.

[21] Stanley, W. S., "Quaternion from Rotation Matrix," *AIAA Journal of Guidance and Control*, Vol. I, No. 3, May 1978, pp. 223–224.

[22] 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.

[23] Federov, F., *The Lorentz Group*, Nauka, Moscow, 1979.

[24] Cayley, A., "On the Motion of Rotation of a Solid Body," *Cambridge Mathematics Journal*, Vol. 3, 1843, pp. 224–232.

[25] Wiener, T. F., *Theoretical Analysis of Gimballess Inertial Reference Equipment Using Delta-Modulated Instruments*, Ph.D. dissertation, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, March 1962.

[26] Marandi, S. R. and Modi, V. J., "A Preferred Coordinate System and the Associated Orientation Representation in Attitude Dynamics," *Acta Astronautica*, Vol. 15, No. 11, 1987, pp. 833–843.

[27] 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.

[28] Schaub, H., Robinett, R. D., and Junkins, J. L., "New Penalty Functions for Optimal Control Formulation for Spacecraft Attitude Control Problems," *Journal of Guidance, Control and Dynamics*, Vol. 20, No. 3, May–June 1997, pp. 428–434.

[29] Tsiotras, P., Junkins, J. L., and Schaub, H., "Higher Order Cayley Transforms with Applications to Attitude Representations," *Journal of Guidance, Control and Dynamics*, Vol. 20, No. 3, May–June 1997, pp. 528–534.

[30] Tsiotras, P. and Longuski, J. M., "A New Parameterization of the Attititude Kinematics," *Journal of the Astronautical Sciences*, Vol. 43, No. 3, 1996, pp. 342–262.

[31] Tsiotras, P., "On the Choice of Coordinates for Control Problems on SO(3)," *30th Annual Conference on Information Sciences and Systems*, Princeton University, March 20–22 1996, pp. 1238–1243.

[32] Craig, J. J., *Introduction to Robotics Mechanics and Control*, Addison-Wesley Publishing Company, 1989.

Chapter Four

Eulerian Mechanics

The dynamics of a continuous body, as presented in the chapter *Newtonian Mechanics*, is specialized in this chapter for the case of rigid body dynamics. This means that all continuous bodies studied will have a constant shape. This is the most common case for many applications. Systems such as satellites, aircraft or robots are all typically modeled as sets of rigid bodies. The rotational dynamics of a rigid body are often referred to as *Eulerian Mechanics*, since Euler's equation $\dot{\boldsymbol{H}} = \boldsymbol{L}$ and Euler's rotational equation of motion generally govern this field.

Unlike the chapter *Newtonian Mechanics*, this chapter will first investigate the rigid body angular momentum vector $\boldsymbol{H}$ and its derivative, along with the kinetic energy before developing the rotational equations of motion. Then the rigid body dynamics in a torque free environment will be studied in more detail. Further, the dynamics of a rigid body is studied when a set of variable speed control moment gyroscopes are present or the body is under the influence of gravity gradient torques.

4.1 Rigid Body Dynamics

4.1.1 Angular Momentum

The following development will parallel the development in Section 2.4 for the case where no body deformations were allowed. Let the moment be taken either about the center of mass or the inertial coordinate frame origin. In either case Euler's equation reduces to

$$\dot{\boldsymbol{H}} = \boldsymbol{L} \tag{4.1}$$

Let $\boldsymbol{R}$ be the inertial position vector of an infinitesimal mass element dm. Let's choose the moment to be defined about the coordinate frame origin O. It turns out this case will include the case of having the moment defined about the center

of mass $\boldsymbol{R}_c$. The angular momentum vector in Eq. (2.95) is reduced to

$$\boldsymbol{H}_O = \int_B \boldsymbol{R} \times \dot{\boldsymbol{R}} dm \tag{4.2}$$

Since $\boldsymbol{R} = \boldsymbol{R}_c + \boldsymbol{r}$ this is rewritten as

$$\boldsymbol{H}_O = \int_B (\boldsymbol{R}_c + \boldsymbol{r}) \times \left(\dot{\boldsymbol{R}}_c + \dot{\boldsymbol{r}}\right) dm \tag{4.3}$$

which is then expanded to

$$\boldsymbol{H}_O = \int_B \boldsymbol{R}_c \times \dot{\boldsymbol{R}}_c dm + \int_B \boldsymbol{r} dm \times \dot{\boldsymbol{R}}_c + \boldsymbol{R}_c \times \int_B \dot{\boldsymbol{r}} dm + \int_B \boldsymbol{r} \times \dot{\boldsymbol{r}} dm \tag{4.4}$$

Noting that the mass of the rigid body is constant and using the definition of the center of mass in Eq. (2.77), the angular momentum vector about the coordinate frame origin O is reduced to the expression

$$\boldsymbol{H}_O = \boldsymbol{R}_c \times M\dot{\boldsymbol{R}}_c + \int_B \boldsymbol{r} \times \dot{\boldsymbol{r}} dm \tag{4.5}$$

Eq. (4.5) is written for the general case where the rigid body $\mathcal{B}$ is rotating about its center of mass and the center of mass is moving independently at an inertial velocity $\dot{\boldsymbol{R}}_c$ as shown in Figure 4.1.

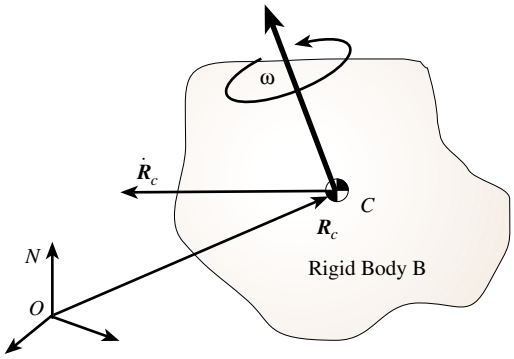


Figure 4.1: General Rigid Body Rotation

The first term of $\boldsymbol{H}_O$ is the angular momentum of the mass center about the origin and its behavior was studied when discussing the dynamics of a single particle. The second term is more interesting since it contains the angular momentum vector $\boldsymbol{H}_c$ of the rigid body $\mathcal{B}$ about its mass center $\boldsymbol{R}_c$. From here on we will be discussing mainly $\boldsymbol{H}_c$ and not the more general $\boldsymbol{H}_O$.

$$\boldsymbol{H}_c = \int_B \boldsymbol{r} \times \dot{\boldsymbol{r}} dm \tag{4.6}$$

At this point we will make use of some of the kinematics results from the previous chapter. By definition, the vector $\dot{\boldsymbol{r}}$ is defined to be an inertial derivative, therefore

$$\dot{\boldsymbol{r}} = \frac{{}^{\mathcal{N}}d}{dt}(\boldsymbol{r}) = \frac{{}^{\mathcal{B}}d}{dt}(\boldsymbol{r}) + \boldsymbol{\omega} \times \boldsymbol{r} \tag{4.7}$$

where the vector $\boldsymbol{\omega}$ is the instantaneous angular velocity vector of the rigid body $\mathcal{B}$ relative to the inertial frame $\mathcal{N}$. Since $\mathcal{B}$ is a rigid body the term ${}^{\mathcal{B}}d/dt\,(\boldsymbol{r})$ is zero. Thus $\dot{\boldsymbol{r}}$ reduces to

$$\dot{\boldsymbol{r}} = \boldsymbol{\omega} \times \boldsymbol{r} \tag{4.8}$$

The angular momentum vector about the center of mass is then defined as

$$\boldsymbol{H}_c = \int_B \boldsymbol{r} \times (\boldsymbol{\omega} \times \boldsymbol{r})\, dm = \left(\int_B -[\tilde{\boldsymbol{r}}][\tilde{\boldsymbol{r}}] dm \right) \boldsymbol{\omega} \tag{4.9}$$

Let the vectors $\hat{\boldsymbol{b}}_i$ be the $\mathcal{B}$ frame unit direction vectors, then the vector $\boldsymbol{r}$, $\boldsymbol{\omega}$ and $\boldsymbol{H}_c$ are written in $\mathcal{B}$ frame coordinates as

$$\boldsymbol{r} = r_1 \hat{\boldsymbol{b}}_1 + r_2 \hat{\boldsymbol{b}}_2 + r_3 \hat{\boldsymbol{b}}_3 \tag{4.10}$$

$$\boldsymbol{\omega} = \omega_1 \hat{\boldsymbol{b}}_1 + \omega_2 \hat{\boldsymbol{b}}_2 + \omega_3 \hat{\boldsymbol{b}}_3 \tag{4.11}$$

$$\boldsymbol{H}_c = H_{c_1} \hat{\boldsymbol{b}}_1 + H_{c_2} \hat{\boldsymbol{b}}_2 + H_{c_3} \hat{\boldsymbol{b}}_3 \tag{4.12}$$

After carrying out the triple cross product and collecting all terms, the angular momentum vector $\boldsymbol{H}_c$ is expressed as

$$\boldsymbol{H}_c = {}^{\mathcal{B}}\begin{pmatrix} H_{c_1} \\ H_{c_2} \\ H_{c_3} \end{pmatrix} = \int_B {}^{\mathcal{B}}\begin{bmatrix} r_2^2 + r_3^2 & -r_1 r_2 & -r_1 r_3 \\ -r_1 r_2 & r_1^2 + r_3^2 & -r_2 r_3 \\ -r_1 r_3 & -r_2 r_3 & r_1^2 + r_2^2 \end{bmatrix} {}^{\mathcal{B}}\begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} dm \tag{4.13}$$

The entries in the 3x3 matrix are the moments and products of inertia of the rigid body $\mathcal{B}$ about its center of mass. Note that since the $\boldsymbol{r}$ vector components were taken in the $\mathcal{B}$ frame, the corresponding matrix components are also taken in the $\mathcal{B}$ frame. If a different coordinate system were assigned to the rigid body, the corresponding inertia matrix would be different too. Let this *symmetric* inertia matrix be called $[I_c]$ where the subscript letter c indicates about which point the moments and products of inertia were taken. If this letter is omitted, then it is understood that this inertia matrix is defined about the center of mass.

$${}^{\mathcal{B}}[I_c] = \int_B -[\tilde{\boldsymbol{r}}][\tilde{\boldsymbol{r}}] dm = \int_B {}^{\mathcal{B}}\begin{bmatrix} r_2^2 + r_3^2 & -r_1 r_2 & -r_1 r_3 \\ -r_1 r_2 & r_1^2 + r_3^2 & -r_2 r_3 \\ -r_1 r_3 & -r_2 r_3 & r_1^2 + r_2^2 \end{bmatrix} dm \tag{4.14}$$

Since $\boldsymbol{\omega}$ does not vary over the volume it can be taken outside the integral. Unless noted otherwise, from here on it will be assumed that the vectors and inertia matrices are written in the $\mathcal{B}$ frame and the superscript letter $\mathcal{B}$ will be dropped. The angular momentum vector of a rigid body about its center of mass can then be written in its simplest form as[1]

$$\boldsymbol{H}_c = [I_c]\boldsymbol{\omega} \tag{4.15}$$

4.1.2 Inertia Matrix Properties

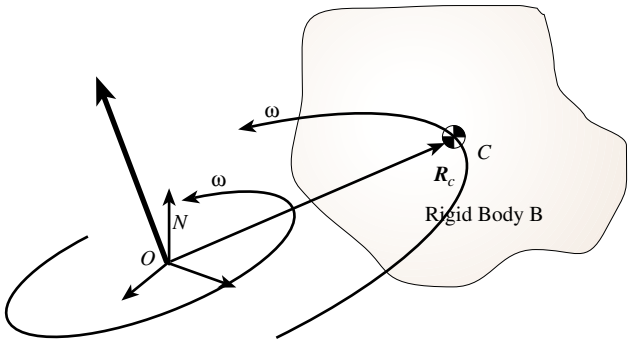


Figure 4.2: Rigid Body Rotation about Origin

Developing Eq. (4.15) it was assumed that the rigid body $\mathcal{B}$ was free to rotate in space. Now it is assumed that the rigid body is no longer rotating independently from the center of mass motion, but instead it is orbiting a fixed point O such that it always keeps the same side facing this point as shown in Figure 4.2. Examples of this type of rotation would be the moon orbiting Earth or a rigid body swinging back and forth at the end of a suspended rope. The center of mass position vector $\boldsymbol{R}_c$ with this type of rotation is fixed in the $\mathcal{B}$ frame and therefore has the following inertial derivative.

$$\dot{\boldsymbol{R}}_c = \frac{{}^{\mathcal{B}}d}{dt}(\boldsymbol{R}_c) + \boldsymbol{\omega} \times \boldsymbol{R}_c = \boldsymbol{\omega} \times \boldsymbol{R}_c \tag{4.16}$$

Substituting Eq. (4.16) into Eq. (4.5) and making use of Eq. (4.15), the angular momentum vector about the origin O is written as

$$\boldsymbol{H}_0 = M\boldsymbol{R}_c \times \boldsymbol{\omega} \times \boldsymbol{R}_c + [I_c]\boldsymbol{\omega} \tag{4.17}$$

After making use of the skew-symmetric tilde operator defined in Eq. (3.23), the vector $\boldsymbol{H}_O$ is written as

$$\boldsymbol{H}_0 = \Big([I_c] - M[\tilde{\boldsymbol{R}}_c][\tilde{\boldsymbol{R}}_c]\Big)\,\boldsymbol{\omega} \tag{4.18}$$

This leads to the famous *parallel axis theorem*. Given the moment of inertia matrix $[I_c]$ of a rigid body $\mathcal{B}$ about its center of mass and the position vector $\boldsymbol{R}_c$ of this center of mass relative some some fixed point O, then the inertia matrix of $\mathcal{B}$ about P is given through the transformation

$$[I_O] = [I_c] + M[\tilde{\boldsymbol{R}}_c][\tilde{\boldsymbol{R}}_c]^T \tag{4.19}$$

Note that the fixed point O does not have to be the origin, but it can be any inertially fixed location. Also, note that when expressing $[I_c]$ and $\boldsymbol{R}_c$ in component form, for the matrix subtraction in Eq. (4.19) to be meaningful, both $[I_c]$ and $\boldsymbol{R}_c$ must be expressed in the same coordinate frame. The resulting matrix $[I_O]$ will also have components taken in the same frame.

Example 4.1: Consider the oblate disk of mass m and radius r rolling on the level surface as shown in Figure 4.3. The disk is attached to a vertical shaft through a massless rod of length L. This horizontal rod is clamped to the center of the disk and pinned to the vertical shaft which is rotating at a constant rate $\dot{\phi}$. What is the normal force $\boldsymbol{N}$ that the surface is exerting onto the disk?

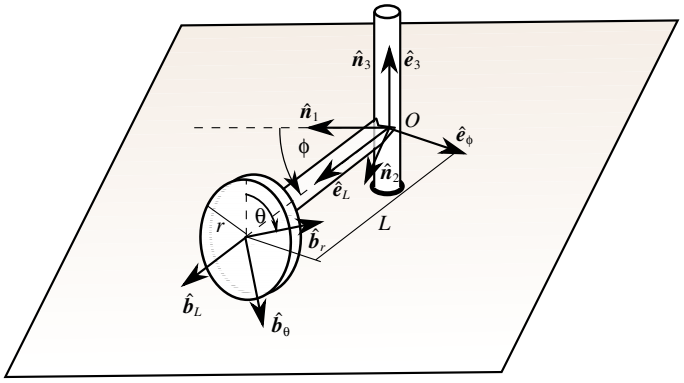


Figure 4.3: Oblate Disk Rolling on Level Surface

Let the coordinate frame $\mathcal{B} : \{\hat{\boldsymbol{b}}_L, \hat{\boldsymbol{b}}_\theta, \hat{\boldsymbol{b}}_r\}$ be attached to the rolling disk, $\mathcal{E} : \{\hat{\boldsymbol{e}}_L, \hat{\boldsymbol{e}}_\phi, \hat{\boldsymbol{e}}_3\}$ be attached to the rotating support rod and $\mathcal{N} : \{\hat{\boldsymbol{n}}_1, \hat{\boldsymbol{n}}_2, \hat{\boldsymbol{n}}_3\}$ be an inertial frame. Since the disk is rolling without slip, the angular rate $\dot{\theta}$ can be related to the shaft rotating rate $\dot{\phi}$ through

$$\dot{\theta} = \frac{L}{r}\dot{\phi}$$

The angular velocity vectors between the respective frames are

$$\boldsymbol{\omega}_{\mathcal{E}/\mathcal{N}} = \dot{\phi}\hat{\boldsymbol{n}}_3 = \dot{\phi}\hat{\boldsymbol{e}}_3 = {}^{\mathcal{E}}\begin{pmatrix} 0 \\ 0 \\ \dot{\phi} \end{pmatrix}$$

$$\boldsymbol{\omega}_{\mathcal{B}/\mathcal{E}} = -\dot{\theta}\hat{\boldsymbol{e}}_L = -\frac{L}{r}\dot{\phi}\hat{\boldsymbol{b}}_L = -\frac{L}{r}\,{}^{\mathcal{B}}\begin{pmatrix} \dot{\phi} \\ 0 \\ 0 \end{pmatrix}$$

Let I_s be the disk inertia about its spin axis $\hat{\boldsymbol{b}}_L$ and I_t the transverse inertias, then the disk inertia matrix about its center of mass is given in $\mathcal{B}$ frame components by the diagonal matrix

$$[I_c] = {}^{\mathcal{B}}\begin{bmatrix} I_s & 0 & 0 \\ 0 & I_t & 0 \\ 0 & 0 & I_t \end{bmatrix}$$

Because the disk is axi-symmetric about the $\hat{\boldsymbol{b}}_L = \hat{\boldsymbol{e}}_L$ axis, for this example ${}^{\mathcal{B}}[I_c] = {}^{\mathcal{E}}[I_c]$ must hold. Since the disk is rotating at an offset distance L

about the $\hat{\boldsymbol{n}}_3$ axis, to find the disk inertia matrix about the point O we must use the parallel axis theorem in Eq. (4.19). The position vector of the disk center of mass is

$$\boldsymbol{R}_c = L\hat{\boldsymbol{e}}_L = \overset{\mathcal{E}}{\begin{pmatrix} L \\ 0 \\ 0 \end{pmatrix}}$$

The disk inertia matrix $[I_O]$ about point O is then given in $\mathcal{E}$ frame components by

$$[I_O] = [I_c] + m[\tilde{\boldsymbol{R}}_c][\tilde{\boldsymbol{R}}_c]^T = \overset{\mathcal{E}}{\begin{bmatrix} I_s & 0 & 0 \\ 0 & I_t + mL^2 & 0 \\ 0 & 0 & I_t + mL^2 \end{bmatrix}}$$

The angular momentum vector $\boldsymbol{H}_O$ of the disk about the point O is the sum of the angular momentum due to the shaft rotation about the $\hat{\boldsymbol{n}}_3$ direction and the rolling about the $\hat{\boldsymbol{b}}_L$ direction.

$$\begin{aligned} \boldsymbol{H}_O &= [I_c]\boldsymbol{\omega}_{\mathcal{B}/\mathcal{E}} + [I_O]\boldsymbol{\omega}_{\mathcal{E}/\mathcal{N}} \\ &= -I_s\frac{L}{r}\dot{\phi}\hat{\boldsymbol{e}}_L + \left(I_t + mL^2\right)\dot{\phi}\hat{\boldsymbol{n}}_3 \end{aligned}$$

The inertial angular momentum vector rate $\dot{\boldsymbol{H}}_O$ is found using the transport theorem.

$$\begin{aligned} \dot{\boldsymbol{H}}_O &= -\frac{^{\mathcal{E}}d}{dt}\left(I_s\frac{L}{r}\dot{\phi}\hat{\boldsymbol{e}}_L\right) - \boldsymbol{\omega}_{\mathcal{E}/\mathcal{N}} \times I_s\frac{L}{r}\dot{\phi}\hat{\boldsymbol{e}}_L + \frac{^{\mathcal{N}}d}{dt}\left(\left(I_t + mL^2\right)\dot{\phi}\hat{\boldsymbol{n}}_3\right) \\ &= -\frac{I_s L}{r}\dot{\phi}^2\hat{\boldsymbol{e}}_\phi \end{aligned}$$

The normal force is defined as $\boldsymbol{N} = N\hat{\boldsymbol{e}}_3$ and the gravity force is given by $\boldsymbol{F}_g = -mg\hat{\boldsymbol{e}}_3$. The torque about point O due to these forces is

$$\boldsymbol{L}_O = \boldsymbol{R}_c \times (\boldsymbol{F}_g + \boldsymbol{N}) = L\left(mg - N\right)\hat{\boldsymbol{e}}_\phi$$

Note that by taking all moments about point O the reaction forces of the pin joint at point O don't appear. Using Euler's equation $\dot{\boldsymbol{H}}_O = \boldsymbol{L}_O$ the normal force component N can be solved for.

$$N = mg + \frac{I_s}{r}\dot{\phi}^2$$

The polar moment of inertia of a circular disk of mass m and radius r is

$$I_s = \frac{m}{2}r^2$$

which allows N to be written as

$$N = m\left(g + \frac{r}{2}\dot{\phi}^2\right)$$

Eq. (4.15) is valid for any choice of body fixed coordinate axes with their origin at the body center of mass. Note that the inertia matrix $[I]$ is calculated for a specific coordinate system. Let the reference frames $\mathcal{B}$ and $\mathcal{F}$ both be proper body fixed coordinate systems. All angular velocities are measured relative to an inertial reference frame $\mathcal{N}$. Let the direction cosine matrix $[FB]$ transform vectors written in the $\mathcal{B}$ frame into vectors expressed in the $\mathcal{F}$ frame. Therefore, using Eq. (3.17) we can write

$$^{\mathcal{F}}\boldsymbol{H}_c = [FB]\,^{\mathcal{B}}\boldsymbol{H}_c \tag{4.20}$$

$$^{\mathcal{F}}\boldsymbol{\omega} = [FB]\,^{\mathcal{B}}\boldsymbol{\omega} \tag{4.21}$$

Let us use the following notation. The matrix $^{\mathcal{F}}[I]$ is the inertia matrix written in the respective $\mathcal{F}$ frame and $^{\mathcal{B}}[I]$ is the inertia matrix in the $\mathcal{B}$ frame. Eq. (4.15) can then be written as

$$^{\mathcal{B}}\boldsymbol{H}_c = {}^{\mathcal{B}}[I]\,{}^{\mathcal{B}}\boldsymbol{\omega} \tag{4.22}$$

which is expanded using Eqs. (3.18), (4.20) and (4.21) to

$$^{\mathcal{F}}\boldsymbol{H}_c = [FB]\,{}^{\mathcal{B}}[I]\,[FB]^T\,{}^{\mathcal{F}}\boldsymbol{\omega} = {}^{\mathcal{F}}[I]\,{}^{\mathcal{F}}\boldsymbol{\omega} \tag{4.23}$$

Thus, an inertia matrix written in the $\mathcal{B}$ frame is rewritten into the $\mathcal{F}$ frame through the similarity transformation

$$^{\mathcal{F}}[I] = [FB]\,{}^{\mathcal{B}}[I]\,[FB]^T \tag{4.24}$$

Whereas Eq. (3.17) maps a vector written in one frame into a vector expressed in another frame, Eq. (4.24) performs the analogous operation for matrices. It allows matrices with components taken in one frame to be expressed with components taken in another frame through the use of the corresponding direction cosine matrix between the two frames.

Given this similarity transformation, the following question arises. Is there a judicious rotation matrix $[C]$ which will rotate the current coordinate frame $\mathcal{B}$ into a new frame $\mathcal{F}$ such that the inertia matrix in this $\mathcal{F}$ frame is diagonal? The answer to this is yes, this is always possible. Let's define $^{\mathcal{F}}[I]$ to be diagonal. Then Eq. (4.24) can be rewritten as

$$\begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix} {}^{\mathcal{B}}\begin{bmatrix} I_{11} & I_{12} & I_{13} \\ I_{12} & I_{22} & I_{23} \\ I_{13} & I_{23} & I_{33} \end{bmatrix} = {}^{\mathcal{F}}\begin{bmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{bmatrix} \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix} \tag{4.25}$$

After carrying out the algebra and equating the proper components, Eq. (4.25) can be reduced to

$${}^{\mathcal{B}}\begin{bmatrix} I_{11} & I_{12} & I_{13} \\ I_{12} & I_{22} & I_{23} \\ I_{13} & I_{23} & I_{33} \end{bmatrix} \begin{pmatrix} C_{i1} \\ C_{i2} \\ C_{i3} \end{pmatrix} = I_{ii\,\mathcal{F}} \begin{pmatrix} C_{i1} \\ C_{i2} \\ C_{i3} \end{pmatrix} \tag{4.26}$$

for i=1,2,3. Studying Eq. (4.26) it is evident that each row of the desired $[C]$ matrix is an eigenvector of the ${}^{\mathcal{B}}[I]$ inertia matrix. Assuming that $\boldsymbol{v}_i$ are the eigenvectors of ${}^{\mathcal{B}}[I]$ we have

$$[C] = [V]^T = \begin{bmatrix} \boldsymbol{v}_1^T \\ \boldsymbol{v}_2^T \\ \boldsymbol{v}_3^T \end{bmatrix} \tag{4.27}$$

The diagonal entries of the new ${}^{\mathcal{F}}[I]$ are the eigenvalues of the old ${}^{\mathcal{B}}[I]$ matrix. Note that the eigenvectors will always be orthogonal since $[C]$ is an orthogonal rotation matrix. The new set of body fixed coordinate axes whose inertia matrix is diagonal are called the *principal axes*. Many analytical problems only consider the simpler case of diagonal inertia matrices since they assume that the appropriate coordinate transformation has already been done. However, in practice it is often difficult to find the exact principal axes of a given body. Here a set of coordinate axes are typically chosen that are close, but not perfectly aligned with the principal axes. The resulting inertia matrix will have dominant diagonal and small off-diagonal terms.

Example 4.2: Find the rotation matrix $[C]$ that will transform the current coordinate frame to a new frame $\mathcal{F}$ which diagonalizes the inertia matrix

$$[I] = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 5 & 2 \\ 1 & 2 & 4 \end{bmatrix}$$

Using Matlab, the eigenvector matrix $[V]$ and eigenvalue vector $\boldsymbol{\Lambda}$ are found to be

$$[V] = \begin{bmatrix} 0.32799 & 0.59101 & 0.73698 \\ 0.73698 & 0.32799 & -0.59101 \\ 0.59101 & -0.73698 & 0.32799 \end{bmatrix} \quad \boldsymbol{\Lambda} = \begin{pmatrix} 7.04892 \\ 2.30798 \\ 2.64310 \end{pmatrix}$$

Note numerical software packages will not necessarily return eigenvectors of unit length. If they are not unit length, they would have to be normalized at this point. In our case the eigenvectors returned are already of unit length. Secondly, we must verify that the set of eigenvectors $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ form a right-handed set. By inspection it is clear that our first eigenvector $\boldsymbol{v}_1$ crossed with the second eigenvector $\boldsymbol{v}_2$ does not yield the third eigenvector $\boldsymbol{v}_3$, but rather $-\boldsymbol{v}_3$. To correct this we change the sign of $\boldsymbol{v}_3$ by simply reversing the sign of each element of the third column of $[V]$. The proper, orthogonal, right-handed $[V]$ matrix is then

$$[V] = \begin{bmatrix} 0.32799 & 0.59101 & -0.73698 \\ 0.73698 & 0.32799 & 0.59101 \\ 0.59101 & -0.73698 & -0.32799 \end{bmatrix}$$

Since each row of the desired $[C]$ matrix is an eigenvector of $[I]$, then

$$[C] = [V]^T = \begin{bmatrix} 0.32799 & 0.73698 & 0.59101 \\ 0.59101 & 0.32799 & -0.73698 \\ -0.73698 & 0.59101 & -0.32799 \end{bmatrix}$$

The new principal inertia matrix components are the eigenvalues of $[I]$.

$$^{\mathcal{F}}I_1 = 7.04892 \quad ^{\mathcal{F}}I_2 = 2.30798 \quad ^{\mathcal{F}}I_3 = 2.64310$$

4.1.3 Euler's Rotational Equations of Motion

Given the previous results, the equations of motion of a rigid body can be developed in a very straight forward fashion. Using the transport theorem, Euler's equation is expressed as

$$\dot{\boldsymbol{H}}_c = \frac{^{\mathcal{B}}d}{dt}(\boldsymbol{H}_c) + \boldsymbol{\omega} \times \boldsymbol{H}_c = \boldsymbol{L}_c \tag{4.28}$$

Using Eqs. (4.15) and the fact that $[I]$ is constant as seen from the $\mathcal{B}$ frame for a rigid body, the derivative of the angular momentum vector $\boldsymbol{H}_c$ as seen in the $\mathcal{B}$ frame is written as

$$\frac{^{\mathcal{B}}d}{dt}(\boldsymbol{H}_c) = \frac{^{\mathcal{B}}d}{dt}([I])\,\boldsymbol{\omega} + [I]\frac{^{\mathcal{B}}d}{dt}(\boldsymbol{\omega}) = [I]\dot{\boldsymbol{\omega}} \tag{4.29}$$

The last step in Eq. (4.29) is true since the derivative of the body angular velocity vector $\boldsymbol{\omega}$ is the same as seen in the $\mathcal{B}$ and the $\mathcal{N}$ frame.

$$\dot{\boldsymbol{\omega}} = \frac{^{\mathcal{N}}d}{dt}(\boldsymbol{\omega}) = \frac{^{\mathcal{B}}d}{dt}(\boldsymbol{\omega}) + \boldsymbol{\omega} \times \boldsymbol{\omega} = \frac{^{\mathcal{B}}d}{dt}(\boldsymbol{\omega}) \tag{4.30}$$

Substituting Eqs. (4.15), (4.29) into Eq. (4.28) yields

$$\boldsymbol{L}_c = [I]\dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times ([I]\boldsymbol{\omega}) \tag{4.31}$$

Using Eq. (3.23), the famous *Euler rotational equations of motion* are[1]

$$[I]\dot{\boldsymbol{\omega}} = -[\tilde{\omega}][I]\boldsymbol{\omega} + \boldsymbol{L}_c \tag{4.32}$$

By choosing a body fixed coordinate system which is aligned with the principal body axes, the inertia matrix $[I]$ will be diagonal and Eq. (4.32) reduces to[2]

$$I_{11}\dot{\omega}_1 = -(I_{33} - I_{22})\omega_2\omega_3 + L_1 \tag{4.33a}$$
$$I_{22}\dot{\omega}_2 = -(I_{11} - I_{33})\omega_3\omega_1 + L_2 \tag{4.33b}$$
$$I_{33}\dot{\omega}_3 = -(I_{22} - I_{11})\omega_1\omega_2 + L_3 \tag{4.33c}$$

For the special case where the body is axially symmetric and no external torques are present, the the rotational equations of motion in Eq. (4.33) are reduced to

$$I_T\dot{\omega}_1 = -(I_{33} - I_T)\omega_2\omega_3 \tag{4.34a}$$
$$I_T\dot{\omega}_2 = (I_{33} - I_T)\omega_3\omega_1 \tag{4.34b}$$
$$I_{33}\dot{\omega}_3 = 0 \tag{4.34c}$$

where the transverse inertia I_T is given by

$$I_T = I_{11} = I_{22} \tag{4.35}$$

From Eq. (4.34c) it is clear that body angular velocity component ω_3 along the axis of symmetry will remain constant. Using this fact while differentiating Eq. (4.34a) and substituting Eq. (4.34b), a second order differential equation for ω_1 is found:

$$\ddot{\omega}_1 + (\frac{I_{33}}{I_T} - 1)^2 \omega_3^2 \omega_1 = 0 \tag{4.36}$$

Similarly, we can find the second order differential equation of ω_2 to be

$$\ddot{\omega}_2 + (\frac{I_{33}}{I_T} - 1)^2 \omega_3^2 \omega_2 = 0 \tag{4.37}$$

Note that these differential equations have the standard form of undamped oscillators. Therefore the solution of $\omega_1(t)$ and $\omega_2(t)$ are given by

$$\omega_1(t) = A_1 \cos \omega_p t + B_1 \sin \omega_p t \tag{4.38a}$$
$$\omega_2(t) = A_2 \cos \omega_p t + B_2 \sin \omega_p t \tag{4.38b}$$

where ω_p is defined as

$$\omega_p = \left(\frac{I_{33}}{I_T} - 1 \right) \omega_3 \tag{4.39}$$

Let ω_{i_0} be the initial body angular velocity components , then the constants A_1 and A_2 must be

$$A_1 = \omega_{1_0} \qquad A_2 = \omega_{2_0} \tag{4.40}$$

Differentiating Eqs (4.38a) and (4.38b) and substituting them into Eqs. (4.34a) and (4.34b), the constants B_1 and B_2 are found to be

$$B_1 = -A_2 \qquad B_2 = A_1 \tag{4.41}$$

The closed form solution of the body angular velocity components are given for this axially symmetric, torque-free case through

$$\omega_1(t) = \omega_{1_0} \cos \omega_p t - \omega_{2_0} \sin \omega_p t \tag{4.42a}$$
$$\omega_2(t) = \omega_{2_0} \cos \omega_p t + \omega_{1_0} \sin \omega_p t \tag{4.42b}$$
$$\omega_3(t) = \omega_{3_0} \tag{4.42c}$$

4.1.4 Kinetic Energy

The kinetic energy of a continuous system is shown in Eq. (2.84) to be

$$T = \frac{1}{2} M \dot{\boldsymbol{R}}_c \cdot \dot{\boldsymbol{R}}_c + \frac{1}{2} \int_B \dot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}} dm = T_{trans} + T_{rot} \tag{4.43}$$

Since we are now only dealing with non-deformable rigid bodies, the kinetic energy component T_{rot} only describes the rotational energy of the rigid body $\mathcal{B}$.

$$T_{rot} = \frac{1}{2}\int_B \dot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}} dm \tag{4.44}$$

After substituting Eq. (4.8), the rotational kinetic energy of a rigid body is expressed as

$$T_{rot} = \frac{1}{2}\int_B (\boldsymbol{\omega} \times \boldsymbol{r}) \cdot (\boldsymbol{\omega} \times \boldsymbol{r})\, dm \tag{4.45}$$

After making use of the trigonometric identity

$$(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c} \equiv \boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}) \tag{4.46}$$

the rotational kinetic energy is rewritten as

$$T_{rot} = \frac{1}{2}\boldsymbol{\omega} \cdot \int_B \boldsymbol{r} \times (\boldsymbol{\omega} \times \boldsymbol{r})\, dm \tag{4.47}$$

Note that the integral is exactly equal to $\boldsymbol{H}_c$ in Eq. (4.9). Therefore, after using Eqs. (4.9) and (4.15) the rigid body rotational kinetic energy expression is simplified to the form

$$T_{rot} = \frac{1}{2}\boldsymbol{\omega} \cdot \boldsymbol{H}_c = \frac{1}{2}\boldsymbol{\omega}^T [I] \boldsymbol{\omega} \tag{4.48}$$

The total kinetic energy of a rigid body $\mathcal{B}$ is the sum of translational and rotational energy as shown in Eq. (4.43) and is given by

$$T = \frac{1}{2} M \dot{\boldsymbol{R}}_c \cdot \dot{\boldsymbol{R}}_c + \frac{1}{2}\boldsymbol{\omega}^T [I] \boldsymbol{\omega} \tag{4.49}$$

To find the work done onto a rigid body $\mathcal{B}$, let us find the derivative of Eq. (4.48).

$$\dot{T}_{rot} = \frac{1}{2}\dot{\boldsymbol{\omega}} \cdot \boldsymbol{H}_c + \frac{1}{2}\boldsymbol{\omega} \cdot \dot{\boldsymbol{H}}_c \tag{4.50}$$

Using Eq. (4.1) and (4.15) this is rewritten as

$$\dot{T}_{rot} = \frac{1}{2}\dot{\boldsymbol{\omega}}^T [I] \boldsymbol{\omega} + \frac{1}{2}\boldsymbol{\omega} \cdot \boldsymbol{L}_c \tag{4.51}$$

After substituting Eq. (4.32) and simplifying the resulting expression, the rotational kinetic energy rate for a rigid body is found to be

$$\dot{T}_{rot} = \boldsymbol{\omega} \cdot \boldsymbol{L}_c \tag{4.52}$$

Using Eq. (2.87), the total kinetic energy rate is then given by

$$\dot{T} = \boldsymbol{F} \cdot \dot{\boldsymbol{R}}_c + \boldsymbol{L}_c \cdot \boldsymbol{\omega} \tag{4.53}$$

If the force vector $\boldsymbol{F}$ is conservative and due to a potential function $V_c(\boldsymbol{R}_c)$ and the torque vector $\boldsymbol{L}_c$ is also conservative and due to a potential function V_T, then Eq. (4.53) can be written as

$$\frac{dT}{dt} + \frac{dV_c}{dt} + \frac{dV_T}{dt} = 0 \tag{4.54}$$

which states that the total system energy $E = T + V_c + V_T$ is conserved in this case.

To find the work W done onto the rigid body $\mathcal{B}$ between two time steps, Eq. (4.53) is integrated once to yield

$$W = T(t_2) - T(t_1) = \int_{t_1}^{t_2} \boldsymbol{F} \cdot \dot{\boldsymbol{R}}_c dt + \int_{t_1}^{t_2} \boldsymbol{L}_c \cdot \boldsymbol{\omega} dt \tag{4.55}$$

Example 4.3: Let us investigate the dynamical system shown in Figure 4.4 where one solid disk of radius r and mass m is rolling off another disk of radius R without slip. The coordinate frame $\mathcal{N} = \{\hat{\boldsymbol{n}}_1, \hat{\boldsymbol{n}}_2, \hat{\boldsymbol{n}}_3\}$ is an inertial frame with it's origin O attached to the center of the stationary disk of radius R. A second coordinate from $\mathcal{E} = \{\hat{\boldsymbol{e}}_r, \hat{\boldsymbol{e}}_\theta, \hat{\boldsymbol{e}}_3\}$ has the same origin O. Note that $\boldsymbol{e}_r$ tracks the heading of the disk center O' relative to O. The angle θ specifies the angular position of the disk center O', while the angle ϕ defines the orientation of the rolling disk relative to the inertial $\hat{\boldsymbol{n}}_2$ axis.

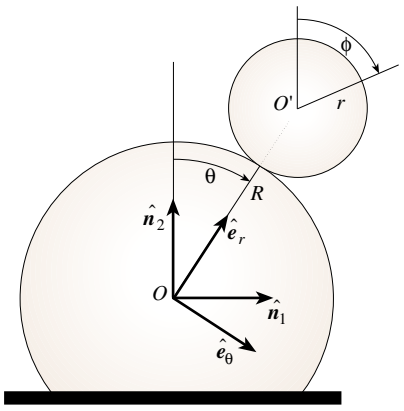


Figure 4.4: Disk Rolling Off Another Disk

Since the disk rolls without slip, the angles θ and ϕ must be related through

$$(R + r)\theta = r\phi$$

First, let's find an expression for the normal force that the lower disk exerts onto the rolling disk. The center of mass position vector of the rolling disk is given through

$$\boldsymbol{r}_c = (R + r)\hat{\boldsymbol{e}}_r$$

The angular velocity of the disk center of mass to the $\mathcal{N}$ frame is

$$\boldsymbol{\omega}_{\mathcal{E}/\mathcal{N}} = \dot{\theta}\hat{\boldsymbol{e}}_3$$

Upon differentiating $\hat{\boldsymbol{r}}_c$, the inertial velocity and acceleration of the disk center of mass are found to be

$$\begin{aligned}\dot{\boldsymbol{r}}_c &= (R+r)\dot{\theta}\hat{\boldsymbol{e}}_\theta \\ \ddot{\boldsymbol{r}}_c &= -(R+r)\dot{\theta}^2\hat{\boldsymbol{e}}_r + (R+r)\ddot{\theta}\hat{\boldsymbol{e}}_\theta\end{aligned}$$

Let N be the normal force component acting along the $\hat{\boldsymbol{e}}_r$ direction and F_f be the frictional force component acting along the $\hat{\boldsymbol{e}}_\theta$ direction. Considering the constant gravity field case, the total force vector acting on the rolling disk is given by

$$\boldsymbol{F} = (N - mg\cos\theta)\,\hat{\boldsymbol{e}}_r + (mg\sin\theta - F_f)\,\hat{\boldsymbol{e}}_\theta$$

The *Super Particle Theorem* for a continuous body states that

$$m\ddot{\boldsymbol{r}}_c = \boldsymbol{F}$$

which leads to

$$-m(R+r)\dot{\theta}^2\hat{\boldsymbol{e}}_r + m(R+r)\ddot{\theta}\hat{\boldsymbol{e}}_\theta = (N - mg\cos\theta)\,\hat{\boldsymbol{e}}_r + (mg\sin\theta - F_f)\,\hat{\boldsymbol{e}}_\theta$$

Equating $\hat{\boldsymbol{e}}_r$ and $\hat{\boldsymbol{e}}_\theta$ components expressions are found for the normal force component N and the friction component F_f.

$$\begin{aligned}N &= mg\cos\theta - m(R+r)\dot{\theta}^2 \\ F_f &= mg\sin\theta - m(R+r)\ddot{\theta}\end{aligned}$$

To write F_f purely in terms of θ and not $\ddot{\theta}$ we study the rotational motion of the disk about its center of mass. The torque $L_{O'}$ experienced by the disk is

$$L_{O'} = rF_f$$

For this simple planar disk Euler's rotational equations of motion simplify to

$$I_c\ddot{\phi} = L_{O'} = rF_f$$

where I_c is the polar mass moment of inertia of the disk given by

$$I_c = \frac{m}{2}r^2$$

Using the relationship $\ddot{\phi} = \frac{R+r}{r}\ddot{\theta}$ the angular acceleration $\ddot{\theta}$ is expressed as

$$\ddot{\theta} = \frac{2F_f}{m(R+r)}$$

The friction force component F_f can now be expressed as

$$F_f = \frac{1}{3}mg\sin\theta$$

Note that the friction component only depends on the angle θ, and not on the disk radius r. The only assumption made here is that the disk inertia satisfies the formula used for I_c.

To find at what angle θ the rolling disk will leave the lower disk, the normal force component N is set to zero. This leads to the first condition that θ must satisfy when the disk leaves the surface.

$$mg\cos\theta = m(R+r)\dot{\theta}^2$$

Let the scalar function $V(\theta)$ be the potential function of the rolling disk.

$$V(\theta) = mg(R+r)\cos\theta$$

The kinetic energy of the disk is given by

$$T = \frac{1}{2}m\dot{\boldsymbol{r}}_c \cdot \dot{\boldsymbol{r}}_c + \frac{1}{2}I_c\dot{\phi}^2 = \frac{3}{4}m\left(R+r\right)^2\dot{\theta}^2$$

Recall that ϕ is measured relative to an inertial axis. Since the dynamical system is conservative, the total energy is conserved. The initial energy E_0 is

$$E_0 = mg(R+r)$$

The total energy at θ is

$$E = mg(R+r)\cos\theta + \frac{3}{4}m\left(R+r\right)^2\dot{\theta}^2$$

Equating the two energy states leads to the expression

$$\dot{\theta}^2 = \frac{4}{3}g\frac{(1-\cos\theta)}{(R+r)}$$

Substituting this $\dot{\theta}^2$ into the first condition on θ (setting the normal force component N equal to zero) leads to

$$\theta = \cos^{-1}\left(\frac{4}{7}\right) = 55.15 \text{ degrees}$$

Thus the normal force component becomes zero at the same angle of θ regardless of the disk mass m and radius r. Again, the underlying assumption is that $I_c = \frac{m}{2}r^2$ is satisfied.

4.2 Torque-Free Rigid Body Rotation

4.2.1 Energy and Momentum Integrals

If no external torques are acting on a system, then Eqs. (2.38), (2.72) and (2.101) show that the total angular momentum vector $\boldsymbol{H}$ is constant. This truth does not depend on whether the system is a single particle, a collection of particles or a continuous body. If no external forces are present, then the

rigid body rotational kinetic energy will also be a constant as seen in Eq. (4.52). Let us write the the angular momentum vector $\boldsymbol{H}$ in terms of body frame $\mathcal{B}$ components as

$$\boldsymbol{H} = {}^{\mathcal{B}}\boldsymbol{H} = H_1\hat{\boldsymbol{b}}_1 + H_2\hat{\boldsymbol{b}}_2 + H_3\hat{\boldsymbol{b}}_3 \tag{4.56}$$

Note that $\dot{\boldsymbol{H}}$ is a derivative taken relative to the inertial reference frame $\mathcal{N}$. Since $\dot{\boldsymbol{H}} = 0$, the angular momentum vector will appear constant only when seen from the inertial $\mathcal{N}$ frame. Relative to the body fixed $\mathcal{B}$ reference frame, the vector $\boldsymbol{H}$ will generally not appear to be constant but rotating. Therefore the $\mathcal{B}$ frame H_i vector components will be time varying. However, the magnitude of $\boldsymbol{H}$ will be constant in all frames.

The present discussion will assume that the body fixed coordinate axes are all aligned with principal inertia axes, therefore the rigid body inertia matrix is diagonal. For notational compactness, let us use the short-hand notation $I_i = I_{ii}$. The angular momentum vector is then written as

$$\boldsymbol{H} = {}^{\mathcal{B}}\boldsymbol{H} = \begin{pmatrix} H_1 \\ H_2 \\ H_3 \end{pmatrix}^{\mathcal{B}} = \begin{pmatrix} I_1\omega_1 \\ I_2\omega_2 \\ I_3\omega_3 \end{pmatrix}^{\mathcal{B}} \tag{4.57}$$

Since the angular momentum magnitude is constant, all possible angular velocities must lie on the surface of the following momentum ellipsoid.

$$H^2 = \boldsymbol{H}^T\boldsymbol{H} = I_1^2\omega_1^2 + I_2^2\omega_2^2 + I_3^2\omega_3^2 \tag{4.58}$$

Because the kinetic energy is constant too, the angular velocities must also lie on the surface following energy ellipsoid.

$$T = \frac{1}{2}I_1\omega_1^2 + \frac{1}{2}I_2\omega_2^2 + \frac{1}{2}I_3\omega_3^2 \tag{4.59}$$

Thus, the dynamical torque-free rotation of a rigid body must be such that the corresponding body angular velocity vector $\boldsymbol{\omega}(t)$ satisfies both Eqs. (4.58) and (4.59). The geometric interpretation of this is that $\boldsymbol{\omega}(t)$ must lie on the intersection of the momentum and energy ellipsoid surfaces.

To more easily visualize the intersection of these two ellipsoids, they are written in terms of the $\mathcal{B}$ frame angular momentum vector components H_i instead of the body angular velocity vector components ω_i. Using H_i as independent coordinates, the momentum ellipsoid becomes the momentum sphere

$$H^2 = H_1^2 + H_2^2 + H_3^2 \tag{4.60}$$

and the energy ellipsoid is written as

$$1 = \frac{H_1^2}{2I_1T} + \frac{H_2^2}{2I_2T} + \frac{H_3^2}{2I_3T} \tag{4.61}$$

where $\sqrt{2I_iT}$ are the corresponding semi-axes. In order for the torque-free rotation to satisfy both Eqs. (4.60) and (4.61), the energy ellipsoid and the momentum sphere must intersect. The intersection forms a trajectory of feasible $\boldsymbol{\omega}(t)$ as illustrated in Figure 4.5. This geometrical interpretations is very useful to make qualitative studies on the nature and limiting properties of large rotations.

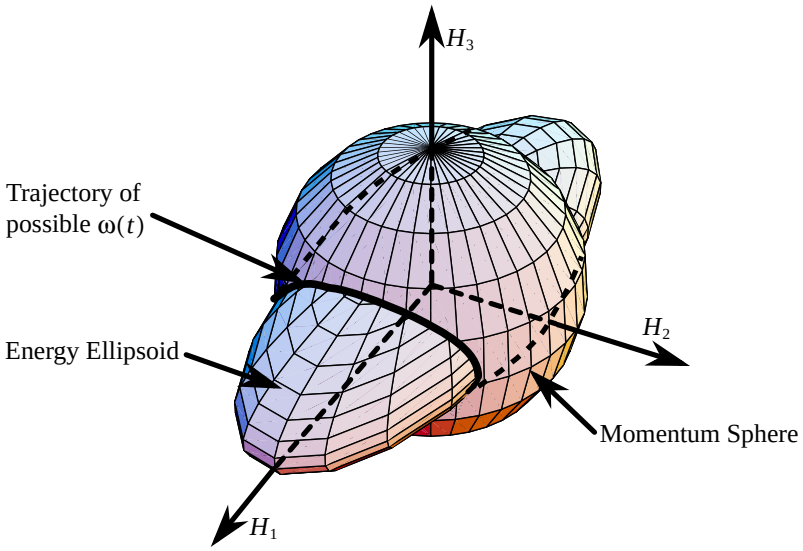


Figure 4.5: General Intersection of the Momentum Sphere and the Energy Ellipsoid

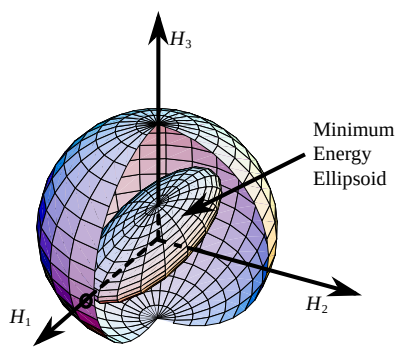
Clearly, for a given $|\boldsymbol{H}|$, only a certain range of kinetic energy is possible. For the current discussion, let us hold the angular momentum vector magnitude constant and sweep the kinetic energy through its two extrema. Also, assume that the inertia matrix entries I_i are ordered such that

$$I_1 \geq I_2 \geq I_3 \tag{4.62}$$

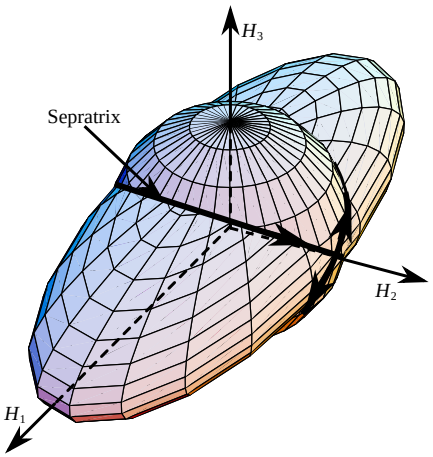
With this ordering of inertias, the largest kinetic energy ellipsoid semi-axis $\sqrt{2I_1T}$ occurs about the $\hat{\boldsymbol{b}}_1$ axis as shown in Figure 4.5, and the smallest semi-axis is about the $\hat{\boldsymbol{b}}_3$ axis. Eq. (4.61) shows that varying T will only uniformly scale the corresponding kinetic energy ellipsoid. The overall shape and aspect ratio of the ellipsoid will remain the same for each choice in T.

Three special energy cases are shown in Figure 4.6. Since the kinetic energy ellipsoid and the momentum sphere must intersect, the smallest possible T would be scaled the energy ellipsoid such that its largest semi-axis is equal to $H = |\boldsymbol{H}|$. The momentum sphere perfectly envelops the energy ellipsoid as shown in Figure 4.6(i). The only points of intersection are at

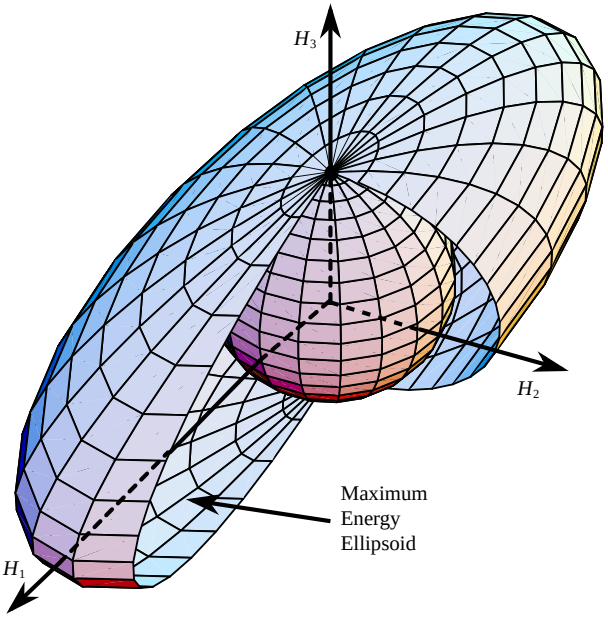
$$^{\mathcal{B}}\boldsymbol{H} = \pm H\hat{\boldsymbol{b}}_1 \tag{4.63}$$



(i) Minimum Energy Case



(ii) Intermediate Energy Case



(iii) Maximum Energy Case

Figure 4.6: Special Cases of Kinetic Energy Ellipsoid and Momentum Sphere Intersections

Therefore, for this *minimum kinetic energy* case, the rigid body $\mathcal{B}$ is spinning purely about its axis of maximum inertia $\hat{\boldsymbol{b}}_1$ and the corresponding kinetic energy is

$$T_{min} = \frac{H^2}{2I_1} \tag{4.64}$$

As the kinetic energy T is enlarged, the next special case arises when the intermediate energy ellipsoid semi-axis is equal to H as shown in Figure 4.6(ii). The intersection curve between the momentum sphere and the energy ellipsoid is called the sepratrix. The kinetic energy for any motion along the sepratrix is given by

$$T_{int} = \frac{H^2}{2I_2} \tag{4.65}$$

Note that any small departure from the pure spin case about the intermediate inertia axis $\hat{\boldsymbol{b}}_2$ will result in general "tumbling" motion. This result agrees well with the common experience that it is very difficult to throw an object into the air and have it spin purely about the intermediate inertia axis without starting to turn and twist about the other axes.

As the kinetic energy is enlarged to its largest possible value, the corresponding kinetic energy ellipsoid perfectly envelops the momentum sphere as shown in Figure 4.6(iii). This *maximum kinetic energy* case

$$T_{max} = \frac{H^2}{2I_3} \tag{4.66}$$

corresponds to a pure spin about the smallest axis of inertia $\hat{\boldsymbol{b}}_3$ since the only intersection point is at

$${}^{\mathcal{B}}\boldsymbol{H} = \pm H\hat{\boldsymbol{b}}_3 \tag{4.67}$$

For a general rigid body motion as shown in Figure 4.5, once the initial kinetic energy T and angular momentum vector $\boldsymbol{H}$ are established, the angular velocity vector $\boldsymbol{\omega}$ will theoretically trace out a particular intersection curve forever. The assumption here is that the body $\mathcal{B}$ is perfectly rigid and that no energy is lost (i.e. no internal dampening, heat loss, ...). However, this assumption is highly idealistic. No body is perfectly rigid and devoid of internal damping. Therefore real rigid bodies spinning in a torque free environment do actually lose energy, though typically at a slow rate.

Figure 4.7 shows a family of energy ellipsoid and momentum sphere intersections for varying levels of kinetic energy. Note that except for the sepratrix case, all feasible $\boldsymbol{\omega}(t)$ paths form closed trajectories. A typical example of a torque-free rigid body rotation would be a rigid satellite launched into an Earth orbit. Once the satellite is spun up about a particular axis and the thrusters are shut down, the satellite won't experience any external torques and the $\boldsymbol{H}$ vector will remain constant. We are ignoring here the affects of atmospheric and solar

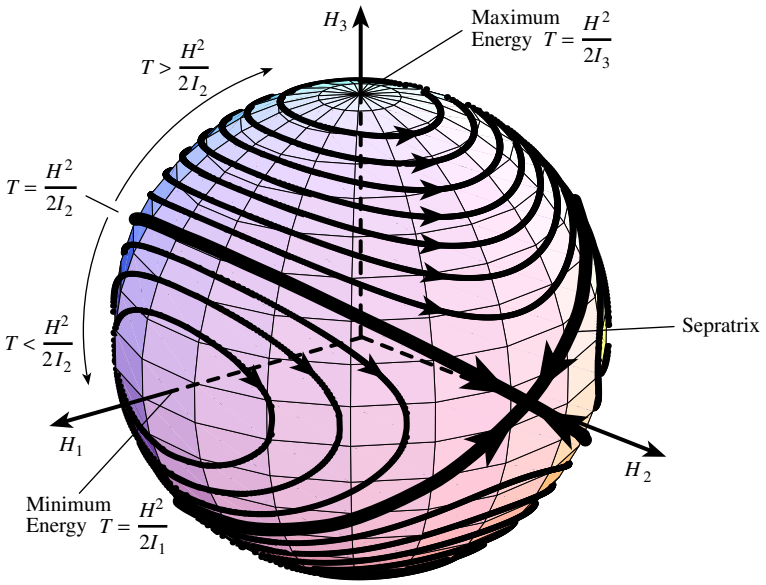


Figure 4.7: A Family of Energy Ellipsoid and Momentum Sphere Intersections

drag. Let's study what happens if a satellite is spun up about the axis of least inertia $\hat{\boldsymbol{b}}_3$. For a given angular momentum, this corresponds to the *maximum kinetic energy* case. Since any real rigid body will loose energy over time simply due to internal damping, this satellite's energy is expected to decrease over time. Figure 4.7 shows how the satellite will start to "wobble" about the $\hat{\boldsymbol{b}}_3$ axis as the energy ellipsoid is reduced. After some time the $\boldsymbol{\omega}(t)$ curves will cross the sepratrix and the satellite will start to "wobble" about the axis of maximum inertia $\hat{\boldsymbol{b}}_1$. Ultimately, as the energy approaches the minimum energy ellipsoid, the satellite will be spinning purely about the $\hat{\boldsymbol{b}}_1$ axis. Therefore, under the presence of a negative energy rate, only the spin about the axis of maximum inertia is a stable spin. The pure spin case about $\hat{\boldsymbol{b}}_3$ will become unstable over time.

This behavior is demonstrated in nature in that all planets are essentially spinning about their axis of maximum inertia. This fact was rediscovered during early space explorations when *Explorer 1* was launched into orbit spinning about its axis of least inertia. It took less than a fraction of an orbit before it started to tumble.

4.2.2 General Free Rigid Body Motion

In this section we would like to derive the general rotational equations of motion for a rigid body free of any external torques. The attitude coordinates are chosen to be the (3-2-1) Euler angles, also known as the yaw, pitch and roll angles

(ψ, θ, ϕ). However, the method used here to derive the equations of motion could be used for any set of attitude coordinates presented in Chapter 3. Assume the rigid body has a coordinate system $\mathcal{B}$ attached to it which is aligned with the principal inertia axes and let $\mathcal{N}$ be an inertial reference frame. For the free motion of a rigid body the angular momentum vector $\boldsymbol{H}$ will remain constant. Using a trick due to Jacobi, we can therefore always align our inertial space unit axes $\hat{\boldsymbol{n}}_i$ such that $\hat{\boldsymbol{n}}_3$ is aligned with $-\boldsymbol{H}$.

$$\boldsymbol{H} = {}^{\mathcal{N}}\boldsymbol{H} = H\hat{\boldsymbol{n}}_3 = \begin{pmatrix} 0 \\ 0 \\ -H \end{pmatrix}^{\mathcal{N}} \tag{4.68}$$

The direction cosine matrix $[BN]$ translates any vector written in $\mathcal{N}$ frame components into a vector in $\mathcal{B}$ frame components as shown in Eq. (3.17).

$${}^{\mathcal{B}}\boldsymbol{H} = [BN]\,{}^{\mathcal{N}}\boldsymbol{H} \tag{4.69}$$

The direction cosine matrix in terms of the (3-2-1) Euler angles is given in Eq. (3.33). After using Eq. (4.68) to carry out the matrix multiplication and equating the resulting $\mathcal{B}$ frame components to Eq. (4.56) we obtain

$$H_1 = H\sin\theta \qquad = I_1\omega_1 \tag{4.70a}$$
$$H_2 = -H\sin\phi\cos\theta \; = I_2\omega_2 \tag{4.70b}$$
$$H_3 = -H\cos\phi\cos\theta = I_3\omega_3 \tag{4.70c}$$

which can be solved for the body angular velocity vector $\boldsymbol{\omega}$ as

$$\begin{pmatrix} \frac{H}{I_1}\sin\theta \\ -\frac{H}{I_2}\sin\phi\cos\theta \\ -\frac{H}{I_3}\cos\phi\cos\theta \end{pmatrix} = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \tag{4.71}$$

To find an expression for the individual Euler angle rates $(\dot{\psi}, \dot{\theta}, \dot{\phi})$, we substitute Eq. (3.55) into Eq. (4.71) and obtain the following equations of motion for a torque-free rigid body.

$$\dot{\psi} = -H\left(\frac{\sin^2\phi}{I_2} + \frac{\cos^2\phi}{I_3}\right) \tag{4.72a}$$
$$\dot{\theta} = \frac{H}{2}\left(\frac{1}{I_3} - \frac{1}{I_2}\right)\sin 2\phi\cos\theta \tag{4.72b}$$
$$\dot{\phi} = H\left(\frac{1}{I_1} - \frac{\sin^2\phi}{I_2} - \frac{\cos^2\phi}{I_3}\right)\sin\theta \tag{4.72c}$$

Note that $\dot{\psi}$ in Eq. (4.72a) cannot be positive, while $\dot{\theta}$ and $\dot{\phi}$ can have either sign.

4.2.3 Axisymmetric Rigid Body Motion

The equations of motion in Eqs. (4.72) are valid for a general rigid body with the body fixed axes aligned with the principal inertia axes. Now we would like to study a particular case of these equations where the rigid body is is axisymmetric. Without loss of generality, assume that $I_2 = I_3$. Then the yaw, pitch and roll angle rates in Eqs. (4.72) simplify to

$$\dot{\psi} = -\frac{H}{I_2} \tag{4.73a}$$

$$\dot{\theta} = 0 \tag{4.73b}$$

$$\dot{\phi} = H\left(\frac{I_2 - I_1}{I_1 I_2}\right)\sin\theta \tag{4.73c}$$

Having chosen the inertial angular momentum vector ${}^{\mathcal{N}}\boldsymbol{H}$ to be in the positive $\hat{\boldsymbol{n}}_3$ direction, the precession rate $\dot{\psi}$ will be a positive constant for an axisymmetric rigid body. The relative spin rate $\dot{\phi}$ is also a constant like precession rate. However, the sign of $\dot{\phi}$ in Eq. (4.73c) depends on the relative size of I_1 and I_2 and on the pitch angle θ. On the other hand, the pitch rate $\dot{\theta}$ is zero for axisymmetric rigid body rotations, therefore the pitch angle θ will remain constant throughout the motion.

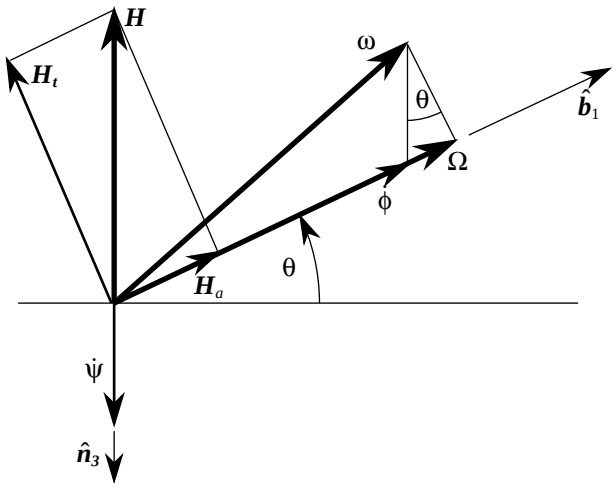


Figure 4.8: Angular Velocity and Momentum Vector Relationship for the Case $I_2 > I_1$

Let $\Omega = \omega_1$ be the body spin rate about its axis of symmetry $\hat{\boldsymbol{b}}_1$ as shown in Figure 4.8. Using Eq. (4.70a) it is expressed as

$$\Omega = \frac{H}{I_1}\sin\theta \tag{4.74}$$

For a positive pitch angle $0 \leq \theta \leq \pi/2$ the body spin rate Ω about the symmetry axis must be positive. Instead of being written in terms of the angular momentum magnitude H, the precession rate $\dot{\psi}$ and the relative spin rate $\dot{\phi}$ can now be expressed in terms of Ω as

$$\dot{\psi} = -\frac{I_1}{I_2}\frac{\Omega}{\sin\theta} \tag{4.75}$$

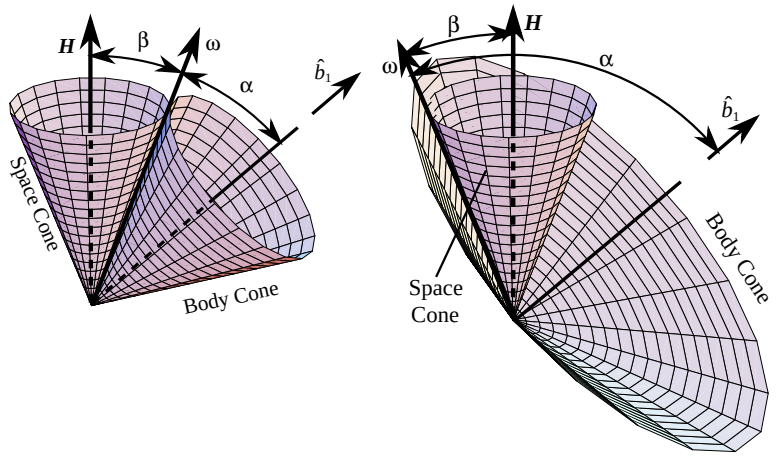
$$\dot{\phi} = \frac{I_2 - I_1}{I_2}\Omega \tag{4.76}$$

The angular momentum vector along the axis of symmetry $\hat{\boldsymbol{b}}_1$ is labeled in Figure 4.8 as $\boldsymbol{H}_a$. It is defined as

$$\boldsymbol{H}_a = I_1\Omega\hat{\boldsymbol{b}}_1 \tag{4.77}$$

As is easily seen in Figure 4.8, for positive pitch angles θ and $I_2 > I_1$, the axisymmetric rigid body $\mathcal{B}$ will have a positive spin rate about $\hat{\boldsymbol{b}}_1$.

Since the pitch angle θ is shown to remain constant during this torque-free rotation, the resulting motion can be visualized by two cones rolling on each other. Figure 4.9 shows the two cases where either $I_2 > I_1$ or $I_2 < I_1$.



(i) $I_2 > I_1$ **(ii)** $I_2 < I_1$

Figure 4.9: Conic Illustration of Direct and Retrograde Precession of a Freely Rotating Rigid Body

The *space cone* is fixed in space and its cone axis is always aligned with the angular momentum vector $\boldsymbol{H}$. The cone angle β is defined as the angle between the vectors $\boldsymbol{H}$ and $\boldsymbol{\omega}$. The *body cone* axis is aligned with the body axis $\hat{\boldsymbol{b}}_1$ and has the cone angle α which is the angle between $\boldsymbol{\omega}$ and $\hat{\boldsymbol{b}}_1$. If $I_2 > I_1$ then the body cone will roll on the outside of the space cone as shown in Figure 4.9(i) and the resulting motion is called a *direct precession*. If $I_2 < I_1$ then the space

cone will lie on the inside of the body cone and the resulting motion is called a retrograde precession.

4.3 Momentum Exchange Devices

Instead of using thrusters to perform precise spacecraft attitude maneuvers, typically control moment gyros (CMGs) or reaction wheels (RWs) are used. Either devices can change its internal angular momentum vector and thus, through Euler's equation, produce an effective torque on the spacecraft. A single-gimbal CMG contains a wheel spinning at a constant rate. To exert a torque onto the spacecraft this wheel is gimbaled or rotated about a fixed axis.[3–5] The rotation axis and rotation angle are referred to as the gimbal axis and gimbal angle respectively. A separate feedback control loop is used to spin up the rotor to the required spin rate and maintain it. The advantage of a CMG is that a relatively small gimbal torque input is required to produce a large effective torque output on the spacecraft. This makes CMGs a very popular devices for reorienting large space structures such as the space station. The drawback of the single-gimbal CMGs is that their control laws can be fairly complex and that such CMG systems encounter certain singular gimbal angle configurations. At these singular configurations the CMG cluster is unable to produce the required torque exactly, or any torque at all if the required torque is orthogonal to the plane of allowable torques. Several papers deal with this issue and present various solutions.[3–6] However, even with singularity robust steering laws or when various singularity avoidance strategies are applied, the actual torque produced by the CMG cluster is never equal to the required torque when maneuvering in the proximity of a singularity. The resulting motion may be stable, but these path deviations can be highly undesirable in some applications. Double-gimbal CMGs have fewer problems with singularities. However, they are also much more costly and complicated devices than the single-gimbal CMGs.

Reaction wheels, on the other hand, have a wheel spinning about a body fixed axis whose spin speed is variable. Torques are produced on the spacecraft by accelerating or decelerating the reaction wheels.[1, 7] RW systems don't have singular configurations and typically have simpler control laws than CMG clusters. Drawbacks to the reaction wheels include a relatively small effective torque being produced on the spacecraft and the possible reaction wheel saturation. To exert a given torque onto a spacecraft, reaction wheels typically requires more energy than CMGs.

Variable Speed Control Moment Gyroscopes (VSCMGs) combine positive features of both the single-gimbal CMGs and the RWs. The spinning disk can be rotated or gimbaled about a single body fixed axis, while the disk spin rate is also free to be controlled.[8, 9] This adds an extra degree of control to the classical single-gimbal CMG device. Note that adding this variable speed feature would not require the single-gimbal CMG to be completely reengineered. These devices already have a separate feedback loop that maintains a constant spin rate. What would need to be changed is that the torque motor controlling the RW spin rate

would need to be larger, and the constant speed law abandoned. With this extra control singular configurations, in the classical CMG sense, will not be present. This section will first develop the equations of motion of a spacecraft containing VSCMGs using Euler's equation. The resulting formulation contains the two classical cases of having either pure RWs or CMGs.

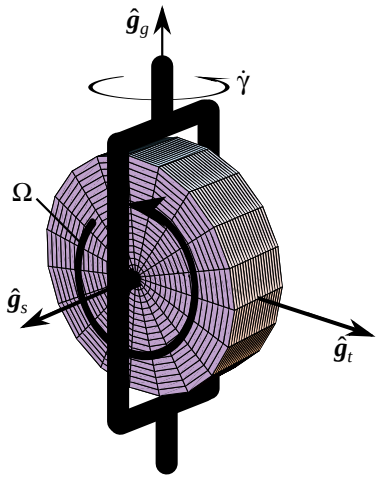


Figure 4.10: Illustration of a Variable Speed Control Moment Gyroscope

4.3.1 Spacecraft with Single VSCMG

To simplify the development and notation, the rotational equations of motion are first derived for the case where only one VSCMG is attached to a rigid spacecraft. Afterwards, the result is expanded to incorporate a system of N VSCMGs. Let $\mathcal{G}$ be the gimbal reference frame whose orientation is given by the triad of unit vectors $\{\hat{\boldsymbol{g}}_s, \hat{\boldsymbol{g}}_t, \hat{\boldsymbol{g}}_g\}$ as shown in Figure 4.10. The vector components of the unit vectors $\hat{\boldsymbol{g}}_i$ are assumed to be given in the spacecraft reference frame $\mathcal{B}$. Note that since the VSCMG gimbal axis $\hat{\boldsymbol{g}}_g$ is fixed relative to $\mathcal{B}$, only the orientation of the spin axis $\hat{\boldsymbol{g}}_s$ and the transverse axis $\hat{\boldsymbol{g}}_t$ will be time varying as seen from the $\mathcal{B}$ frame. Given an initial gimbal angle γ_0, the spin and transverse axis at a gimbal angle $\gamma(t)$ are given by

$$\hat{\boldsymbol{g}}_s(t) = \cos(\gamma(t) - \gamma_0)\,\hat{\boldsymbol{g}}_s(t_0) + \sin(\gamma(t) - \gamma_0)\,\hat{\boldsymbol{g}}_t(t_0) \tag{4.78}$$

$$\hat{\boldsymbol{g}}_t(t) = -\sin(\gamma(t) - \gamma_0)\,\hat{\boldsymbol{g}}_s(t_0) + \cos(\gamma(t) - \gamma_0)\,\hat{\boldsymbol{g}}_t(t_0) \tag{4.79}$$

The spin rate of the VSCMG about $\hat{\boldsymbol{g}}_s$ is denoted by Ω. The angular velocity vector of the $\mathcal{G}$ frame relative to the $\mathcal{B}$ frame is

$$\boldsymbol{\omega}_{\mathcal{G}/\mathcal{B}} = \dot{\gamma}\hat{\boldsymbol{g}}_g \tag{4.80}$$

The angular velocity vector of the reaction wheel frame $\mathcal{W}$ relative to the gimbal frame $\mathcal{G}$ is

$$\boldsymbol{\omega}_{\mathcal{W}/\mathcal{G}} = \Omega \hat{\boldsymbol{g}}_s \tag{4.81}$$

To indicate in which reference frame vector or matrix components are taken, a superscript letter is added before the vector or matrix name. Since the $\mathcal{G}$ frame unit axes are aligned with the principal gimbal frame axes, the gimbal frame inertia matrix $[I_G]$ expressed in the $\mathcal{G}$ frame is the constant diagonal matrix.

$$[I_G] = {}^{\mathcal{G}}[I_G] = {}^{\mathcal{G}}\begin{bmatrix} I_{G_s} & 0 & 0 \\ 0 & I_{G_t} & 0 \\ 0 & 0 & I_{G_g} \end{bmatrix} \tag{4.82}$$

where I_{G_s}, I_{G_t} and I_{G_g} are the gimbal frame inertias about the corresponding spin, transverse and gimbal axes. The reaction wheel inertia about the same axes are denoted by I_{W_s} and $I_{W_t} = I_{W_g}$.

$$[I_W] = {}^{\mathcal{W}}[I_W] = {}^{\mathcal{W}}\begin{bmatrix} I_{W_s} & 0 & 0 \\ 0 & I_{W_t} & 0 \\ 0 & 0 & I_{W_t} \end{bmatrix} \tag{4.83}$$

Note that since the disk is symmetric about the $\hat{\boldsymbol{g}}_s$ axis ${}^{\mathcal{W}}[I_W] = {}^{\mathcal{G}}[I_W]$. In practice I_{W_s} is typically much larger than any of the other gimbal frame or RW inertias. In this development the RW and gimbal frame inertias are not combined early on into one overall VSCMG inertia matrix; rather, they are retained as separate entities until later into the development. This will allow for a precise formulation of the actual physical motor torques that drive the RWs or the CMGs.

The $\mathcal{G}$ frame orientation is related to the $\mathcal{B}$ frame orientation through the direction cosine matrix $[BG]$ which is expressed in terms of the gimbal frame unit direction vectors as

$$[BG] = [\hat{\boldsymbol{g}}_s \, \hat{\boldsymbol{g}}_t \, \hat{\boldsymbol{g}}_g] \tag{4.84}$$

In Eq. (4.84) the $\hat{\boldsymbol{g}}_i$ vector components are taken in the $\mathcal{B}$ frame. The rotation matrix $[BG]$ maps a vector with components taken in the $\mathcal{G}$ frame into a vector with components in the $\mathcal{B}$ frame. The constant diagonal inertia matrices ${}^{\mathcal{G}}[I_G]$ and ${}^{\mathcal{G}}[I_W]$ are expressed with components taken in the $\mathcal{B}$ frame as the time varying matrices[2, 10]

$${}^{\mathcal{B}}[I_G] = [BG]\,{}^{\mathcal{G}}[I_G]\,[BG]^T = I_{G_s}\hat{\boldsymbol{g}}_s\hat{\boldsymbol{g}}_s^T + I_{G_t}\hat{\boldsymbol{g}}_t\hat{\boldsymbol{g}}_t^T + I_{G_g}\hat{\boldsymbol{g}}_g\hat{\boldsymbol{g}}_g^T \tag{4.85}$$

$${}^{\mathcal{B}}[I_W] = [BG]\,{}^{\mathcal{G}}[I_W]\,[BG]^T = I_{W_s}\hat{\boldsymbol{g}}_s\hat{\boldsymbol{g}}_s^T + I_{W_t}\hat{\boldsymbol{g}}_t\hat{\boldsymbol{g}}_t^T + I_{W_t}\hat{\boldsymbol{g}}_g\hat{\boldsymbol{g}}_g^T \tag{4.86}$$

The total angular momentum of the spacecraft and the VSCMG about the spacecraft center of mass is given by

$$\boldsymbol{H} = \boldsymbol{H}_B + \boldsymbol{H}_G + \boldsymbol{H}_W \tag{4.87}$$

where $\boldsymbol{H}_B$ is the angular momentum component of the spacecraft, $\boldsymbol{H}_G$ is the angular momentum of the gimbal frame and $\boldsymbol{H}_W$ is the angular momentum of the RW. Let $\mathcal{N}$ be an inertial reference frame and $\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}$ be the relative angular velocity vector, then $\boldsymbol{H}_B$ is written as

$$\boldsymbol{H}_B = [I_s]\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}} \tag{4.88}$$

The matrix $[I_s]$ contains the spacecraft inertia terms and the VSCMG inertia components due to the fact that the VSCMG center of mass is not located at the spacecraft center of mass. Note that ${}^{\mathcal{B}}[I_s]$ is a constant matrix as seen from the $\mathcal{B}$ frame. The gimbal frame angular momentum $\boldsymbol{H}_G$ is given by

$$\boldsymbol{H}_G = [I_G]\boldsymbol{\omega}_{\mathcal{G}/\mathcal{N}} \tag{4.89}$$

where $\boldsymbol{\omega}_{\mathcal{G}/\mathcal{N}} = \boldsymbol{\omega}_{\mathcal{G}/\mathcal{B}} + \boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}$. Using Eqs. (4.80), (4.82) and (4.85) this is rewritten as

$$\boldsymbol{H}_G = \left(I_{G_s}\hat{\boldsymbol{g}}_s\hat{\boldsymbol{g}}_s^T + I_{G_t}\hat{\boldsymbol{g}}_t\hat{\boldsymbol{g}}_t^T + I_{G_g}\hat{\boldsymbol{g}}_g\hat{\boldsymbol{g}}_g^T\right)\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}} + I_{G_g}\dot{\gamma}\hat{\boldsymbol{g}}_g \tag{4.90}$$

To simplify the following notation, let the variables ω_s, ω_t and ω_g be the projection of $\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}$ onto the $\mathcal{G}$ frame unit axes.

$$\omega_s = \hat{\boldsymbol{g}}_s^T\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}} \tag{4.91a}$$

$$\omega_t = \hat{\boldsymbol{g}}_t^T\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}} \tag{4.91b}$$

$$\omega_g = \hat{\boldsymbol{g}}_g^T\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}} \tag{4.91c}$$

The angular momentum $\boldsymbol{H}_G$ is then written as

$$\boldsymbol{H}_G = I_{G_s}\omega_s\hat{\boldsymbol{g}}_s + I_{G_t}\omega_t\hat{\boldsymbol{g}}_t + I_{G_g}\left(\omega_g + \dot{\gamma}\right)\hat{\boldsymbol{g}}_g \tag{4.92}$$

The RW angular momentum $\boldsymbol{H}_W$ is given by

$$\boldsymbol{H}_W = [I_W]\boldsymbol{\omega}_{\mathcal{W}/\mathcal{N}} \tag{4.93}$$

where $\boldsymbol{\omega}_{\mathcal{W}/\mathcal{N}} = \boldsymbol{\omega}_{\mathcal{W}/\mathcal{G}} + \boldsymbol{\omega}_{\mathcal{G}/\mathcal{B}} + \boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}$. Using analogous definitions as for $\boldsymbol{H}_G$, $\boldsymbol{H}_W$ is rewritten as

$$\boldsymbol{H}_W = I_{W_s}\left(\omega_s + \Omega\right)\hat{\boldsymbol{g}}_s + I_{W_t}\omega_t\hat{\boldsymbol{g}}_t + I_{W_t}\left(\omega_g + \dot{\gamma}\right)\hat{\boldsymbol{g}}_g \tag{4.94}$$

To simplify the notation from here on, let us use the short hand notation $\boldsymbol{\omega} = \boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}$. In some calculations it will be convenient to express $\boldsymbol{\omega}$ in the $\mathcal{G}$ frame as

$${}^{\mathcal{G}}\boldsymbol{\omega} = \omega_s\hat{\boldsymbol{g}}_s + \omega_t\hat{\boldsymbol{g}}_t + \omega_g\hat{\boldsymbol{g}}_g \tag{4.95}$$

The equations of motion of a system of rigid bodies follow from Euler's equation

$$\dot{\boldsymbol{H}} = \boldsymbol{L} \tag{4.96}$$

if all moments are taken about the center of mass. The vector $\boldsymbol{L}$ represents the sum of all the external torques experienced by the spacecraft. To find the inertial derivatives of $\boldsymbol{H}_G$ and $\boldsymbol{H}_W$, the inertial derivatives of the vectors $\{\hat{\boldsymbol{g}}_s, \hat{\boldsymbol{g}}_t, \hat{\boldsymbol{g}}_g\}$ are required. Using the transport theorem we find

$$\dot{\hat{\boldsymbol{g}}}_s = \frac{\mathcal{B}d}{dt}(\hat{\boldsymbol{g}}_s) + \boldsymbol{\omega} \times \hat{\boldsymbol{g}}_s = (\dot{\gamma} + \omega_g)\,\hat{\boldsymbol{g}}_t - \omega_t \hat{\boldsymbol{g}}_g \tag{4.97a}$$

$$\dot{\hat{\boldsymbol{g}}}_t = \frac{\mathcal{B}d}{dt}(\hat{\boldsymbol{g}}_t) + \boldsymbol{\omega} \times \hat{\boldsymbol{g}}_t = -(\dot{\gamma} + \omega_g)\,\hat{\boldsymbol{g}}_s + \omega_s \hat{\boldsymbol{g}}_g \tag{4.97b}$$

$$\dot{\hat{\boldsymbol{g}}}_g = \frac{\mathcal{B}d}{dt}(\hat{\boldsymbol{g}}_g) + \boldsymbol{\omega} \times \hat{\boldsymbol{g}}_g = \omega_t \hat{\boldsymbol{g}}_s - \omega_s \hat{\boldsymbol{g}}_t \tag{4.97c}$$

since the $\mathcal{B}$ frame derivatives are

$$\frac{\mathcal{B}d}{dt}(\hat{\boldsymbol{g}}_s) = \dot{\gamma}\hat{\boldsymbol{g}}_t \tag{4.98}$$

$$\frac{\mathcal{B}d}{dt}(\hat{\boldsymbol{g}}_t) = -\dot{\gamma}\hat{\boldsymbol{g}}_s \tag{4.99}$$

$$\frac{\mathcal{B}d}{dt}(\hat{\boldsymbol{g}}_g) = 0 \tag{4.100}$$

as can be verified through Eqs. (4.78) and (4.79). The inertial derivatives of the $\mathcal{G}$ frame body angular velocity components are

$$\dot{\omega}_s = \dot{\hat{\boldsymbol{g}}}_s^T \boldsymbol{\omega} + \hat{\boldsymbol{g}}_s^T \dot{\boldsymbol{\omega}} = \dot{\gamma}\omega_t + \hat{\boldsymbol{g}}_s^T \dot{\boldsymbol{\omega}} \tag{4.101a}$$

$$\dot{\omega}_t = \dot{\hat{\boldsymbol{g}}}_t^T \boldsymbol{\omega} + \hat{\boldsymbol{g}}_t^T \dot{\boldsymbol{\omega}} = -\dot{\gamma}\omega_s + \hat{\boldsymbol{g}}_t^T \dot{\boldsymbol{\omega}} \tag{4.101b}$$

$$\dot{\omega}_g = \dot{\hat{\boldsymbol{g}}}_g^T \boldsymbol{\omega} + \hat{\boldsymbol{g}}_g^T \dot{\boldsymbol{\omega}} = \hat{\boldsymbol{g}}_g^T \dot{\boldsymbol{\omega}} \tag{4.101c}$$

Using these definitions the inertial derivative of $\boldsymbol{H}_W$ is expressed as

$$\begin{aligned}\dot{\boldsymbol{H}}_W = {} & \hat{\boldsymbol{g}}_s I_{W_s}\left(\dot{\Omega} + \hat{\boldsymbol{g}}_s^T \dot{\boldsymbol{\omega}} + \dot{\gamma}\omega_t\right) \\ & + \hat{\boldsymbol{g}}_t\left(I_{W_s}\dot{\gamma}\omega_s + I_{W_t}\hat{\boldsymbol{g}}_t^T \dot{\boldsymbol{\omega}} + (I_{W_s} - I_{W_t})\,\omega_s\omega_g + I_{W_s}\Omega(\dot{\gamma} + \omega_g)\right) \\ & + \hat{\boldsymbol{g}}_g\left(I_{W_t}\hat{\boldsymbol{g}}_g^T(\dot{\boldsymbol{\omega}} + \ddot{\gamma}) + (I_{W_t} - I_{W_s})\,\omega_s\omega_t + I_{W_s}\Omega\omega_t\right)\end{aligned} \tag{4.102}$$

Let $\boldsymbol{L}_W$ be the torque the gimbal frame exerts on the RW. Isolating the dynamics of the RW, Euler's equation states that $\dot{\boldsymbol{H}}_W = \boldsymbol{L}_W$. The torque components in the $\hat{\boldsymbol{g}}_t$ and $\hat{\boldsymbol{g}}_g$ direction are produced by the gimbal frame itself. However, the torque component u_s about the $\hat{\boldsymbol{g}}_s$ axis is produced by the RW torque motor. Therefore, from Eq. (4.102) the spin control torque u_s is given by

$$u_s = I_{W_s}\left(\dot{\Omega} + \hat{\boldsymbol{g}}_s^T \dot{\boldsymbol{\omega}} + \dot{\gamma}\omega_t\right) \tag{4.103}$$

After differentiating Eq. (4.92) and using the definitions in Eqs. (4.97) and

(4.101), $\dot{\boldsymbol{H}}_G$ is expressed as

$$\begin{aligned}\dot{\boldsymbol{H}}_G = &\hat{\boldsymbol{g}}_s \left(\left(I_{G_s} - I_{G_t} + I_{G_g} \right) \dot{\gamma}\omega_t + I_{G_s} \hat{\boldsymbol{g}}_s^T \dot{\boldsymbol{\omega}} + \left(I_{G_g} - I_{G_t} \right) \omega_t \omega_g \right) \\ &+ \hat{\boldsymbol{g}}_t \left(\left(I_{G_s} - I_{G_t} - I_{G_g} \right) \dot{\gamma}\omega_s + I_{G_t} \hat{\boldsymbol{g}}_t^T \dot{\boldsymbol{\omega}} + \left(I_{G_s} - I_{G_t} \right) \omega_s \omega_g \right) \\ &+ \hat{\boldsymbol{g}}_g \left(I_{G_g} \left(\hat{\boldsymbol{g}}_g^T \dot{\boldsymbol{\omega}} + \ddot{\gamma} \right) + \left(I_{G_t} - I_{G_s} \right) \omega_s \omega_t \right)\end{aligned} \tag{4.104}$$

From here on it is convenient to combine the inertia matrices of the RW and the gimbal frame into one VSCMG inertia matrix $[J]$ as

$$[J] = [I_G] + [I_W] = \overset{\mathcal{G}}{\begin{bmatrix} J_s & 0 & 0 \\ 0 & J_t & 0 \\ 0 & 0 & J_g \end{bmatrix}} \tag{4.105}$$

Let $\boldsymbol{L}_G$ be the torque vector that the combined RW and CMG system exerts onto the spacecraft, then Euler's equation states that $\dot{\boldsymbol{H}}_G + \dot{\boldsymbol{H}}_W = \boldsymbol{L}_G$. The $\boldsymbol{L}_G$ torque component about the $\hat{\boldsymbol{g}}_g$ axis is produced by the gimbal torque motor. Adding Eqs. (4.102) and (4.104) and making use of the definition in Eq. (4.105), the gimbal torque u_g is then expressed as

$$u_g = J_g \left(\hat{\boldsymbol{g}}_g^T \dot{\boldsymbol{\omega}} + \ddot{\gamma} \right) - \left(J_s - J_t \right) \omega_s \omega_t - I_{W_s} \Omega \omega_t \tag{4.106}$$

The inertial derivative of $\boldsymbol{H}_B$ is simply

$$\dot{\boldsymbol{H}}_B = [I_s]\dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times [I_s]\boldsymbol{\omega} \tag{4.107}$$

To further simplify the equations of motions, the total spacecraft inertia matrix $[I]$ is defined as

$$[I] = [I_s] + [J] \tag{4.108}$$

Substituting Eqs. (4.102), (4.104) and (4.107) back into Eq. (4.96) and making use of the definition in Eq. (4.108), the equations of motion for a rigid spacecraft containing one VSCMG are

$$\begin{aligned}[I]\dot{\boldsymbol{\omega}} = -\boldsymbol{\omega} \times [I]\boldsymbol{\omega} &- \hat{\boldsymbol{g}}_s \left(J_s \dot{\gamma}\omega_t + I_{W_s} \dot{\Omega} - \left(J_t - J_g \right) \omega_t \dot{\gamma} \right) \\ &- \hat{\boldsymbol{g}}_t \left(\left(J_s \omega_s + I_{W_s} \Omega \right) \dot{\gamma} - \left(J_t + J_g \right) \omega_s \dot{\gamma} + I_{W_s} \Omega \omega_g \right) \\ &- \hat{\boldsymbol{g}}_g \left(J_g \ddot{\gamma} - I_{W_s} \Omega \omega_t \right) + \boldsymbol{L}\end{aligned} \tag{4.109}$$

where the identity

$$\boldsymbol{\omega} \times [J]\boldsymbol{\omega} = \left(J_g - J_t \right) \omega_t \omega_g \hat{\boldsymbol{g}}_s + \left(J_s - J_g \right) \omega_s \omega_g \hat{\boldsymbol{g}}_t + \left(J_t - J_s \right) \omega_s \omega_s \hat{\boldsymbol{g}}_g \tag{4.110}$$

is used to combine terms into $\boldsymbol{\omega} \times [I]\boldsymbol{\omega}$. From here on the common assumption will be made that $J_s \approx I_{W_s}$, i.e,. that the gimbal frame inertia I_{G_s} about the spin axis is negligible. The corresponding equations of motion are simplified to

$$\begin{aligned}[I]\dot{\boldsymbol{\omega}} = -\boldsymbol{\omega} \times [I]\boldsymbol{\omega} &- \hat{\boldsymbol{g}}_s \left(J_s \left(\dot{\Omega} + \dot{\gamma}\omega_t \right) - \left(J_t - J_g \right) \omega_t \dot{\gamma} \right) \\ &- \hat{\boldsymbol{g}}_t \left(J_s \left(\omega_s + \Omega \right) \dot{\gamma} - \left(J_t + J_g \right) \omega_s \dot{\gamma} + J_s \Omega \omega_g \right) \\ &- \hat{\boldsymbol{g}}_g \left(J_g \ddot{\gamma} - J_s \Omega \omega_t \right) + \boldsymbol{L}\end{aligned} \tag{4.111}$$

Note that the equations of motion in Eq. (4.111) incorporates both classical cases of having *either* a single-gimbal CMG or a RW attached.

Example 4.4: To reduce the general equations of motion in Eq. (4.111) to that of a spacecraft with a single CMG, the RW spin speed is forced to remain constant by setting $\dot{\Omega} = 0$. Quickly the standard single-gimbal CMG equations of motion are retrieved to be

$$\begin{aligned}[I]\dot{\boldsymbol{\omega}} = -\boldsymbol{\omega} \times [I]\boldsymbol{\omega} &- \hat{\boldsymbol{g}}_s \left(J_s \dot{\gamma} \omega_t - \left(J_t - J_g\right) \omega_t \dot{\gamma}\right) \\ &- \hat{\boldsymbol{g}}_t \left(J_s \left(\omega_s + \Omega\right) \dot{\gamma} - \left(J_t + J_g\right) \omega_s \dot{\gamma} + J_s \Omega \omega_g\right) \\ &- \hat{\boldsymbol{g}}_g \left(J_g \ddot{\gamma} - J_s \Omega \omega_t\right) + \boldsymbol{L}\end{aligned}$$

This is also the form that is commonly used when designing control laws since CMGs are controlled at a gimbal velocity $\dot{\gamma}$ level. There is no need to have the gimbal motor torque u_g explicitly present in this formulation.

To retrieve the RW equations of motion from Eq. (4.111), the gimbal rates and accelerations $\dot{\gamma}$ and $\ddot{\gamma}$ are set to zero. The resulting equations of motion of a spacecraft with a single RW attached are

$$[I]\dot{\boldsymbol{\omega}} = -\boldsymbol{\omega} \times [I]\boldsymbol{\omega} - \hat{\boldsymbol{g}}_s J_s \dot{\Omega} - J_s \Omega \left(\omega_g \hat{\boldsymbol{g}}_t - \omega_t \hat{\boldsymbol{g}}_g\right) + \boldsymbol{L}$$

which can be simplified using the cross product operator to be

$$[I]\dot{\boldsymbol{\omega}} = -\boldsymbol{\omega} \times [I]\boldsymbol{\omega} - \hat{\boldsymbol{g}}_s J_s \dot{\Omega} - \boldsymbol{\omega} \times J_s \Omega \hat{\boldsymbol{g}}_s + \boldsymbol{L}$$

However, many times it is convenient to have these equations of motion written in terms of the u_s instead of $\dot{\Omega}$ to which results in control laws that directly find the required RW motor torque. The motor torque given in Eq. (4.103) is simplified for this case to be

$$u_s = J_s \left(\dot{\Omega} + \hat{\boldsymbol{g}}_s^T \dot{\boldsymbol{\omega}}\right) \tag{4.112}$$

Extracting the J_s component of the inertia matrix $[I]$, the modified inertia matrix $[I_{RW}]$ is defined as

$$[I_{RW}] = [I_s] + J_t \hat{\boldsymbol{g}}_t \hat{\boldsymbol{g}}_t^T + J_g \hat{\boldsymbol{g}}_g \hat{\boldsymbol{g}}_g^T$$

which allows the equations of motion to be written in the standard form terms of u_s.[7]

$$[I_{RW}]\dot{\boldsymbol{\omega}} = -\boldsymbol{\omega} \times [I_{RW}]\boldsymbol{\omega} - \boldsymbol{\omega} \times J_s \hat{\boldsymbol{g}}_s \left(\omega_s \Omega\right) - u_s \hat{\boldsymbol{g}}_s + \boldsymbol{L} \tag{4.113}$$

4.3.2 Spacecraft with Multiple VSCMGs

To obtain the equations of motion of a rigid spacecraft with several VSCMGs attached, the effects of each $\dot{\boldsymbol{H}}_G$ and $\dot{\boldsymbol{H}}_W$ are added up. To simplify notation,

let us define the following useful matrices. The $3\text{x}N$ matrices $[G_s]$, $[G_t]$ and $[G_g]$ contain the unit direction vectors of each VSCMG gimbal frame.

$$[G_s] = [\hat{\boldsymbol{g}}_{s_1} \cdots \hat{\boldsymbol{g}}_{s_N}] \tag{4.114a}$$

$$[G_t] = [\hat{\boldsymbol{g}}_{t_1} \cdots \hat{\boldsymbol{g}}_{t_N}] \tag{4.114b}$$

$$[G_g] = [\hat{\boldsymbol{g}}_{g_1} \cdots \hat{\boldsymbol{g}}_{g_N}] \tag{4.114c}$$

The total spacecraft inertia matrix is expressed as

$$[I] = [I_s] + \sum_{i=1}^{N} [J_i] = [I_s] + \sum_{i=1}^{N} J_{s_i} \hat{\boldsymbol{g}}_{s_i} \hat{\boldsymbol{g}}_{s_i}^T + J_{t_i} \hat{\boldsymbol{g}}_{t_i} \hat{\boldsymbol{g}}_{t_i}^T + J_{g_i} \hat{\boldsymbol{g}}_{g_i} \hat{\boldsymbol{g}}_{g_i}^T \tag{4.115}$$

The torque-like quantities τ_{s_i}, τ_{t_i} and τ_{g_i} are defined as

$$\boldsymbol{\tau}_s = \begin{bmatrix} J_{s_1} \left(\dot{\Omega}_1 + \dot{\gamma}_1 \omega_{t_1} \right) - \left(J_{t_1} - J_{g_1} \right) \omega_{t_1} \dot{\gamma}_1 \\ \vdots \\ J_{s_N} \left(\dot{\Omega}_N + \dot{\gamma}_N \omega_{t_N} \right) - \left(J_{t_N} - J_{g_N} \right) \omega_{t_N} \dot{\gamma}_N \end{bmatrix} \tag{4.116a}$$

$$\boldsymbol{\tau}_s = \begin{bmatrix} J_{s_1} \left(\Omega_1 + \omega_{s_1} \right) \dot{\gamma}_1 - \left(J_{t_1} + J_{g_1} \right) \omega_{s_1} \dot{\gamma}_1 + J_{s_1} \Omega_1 \omega_{g_1} \\ \vdots \\ J_{s_N} \left(\Omega_N + \omega_{s_N} \right) \dot{\gamma}_N - \left(J_{t_N} + J_{g_N} \right) \omega_{s_N} \dot{\gamma}_N + J_{s_N} \Omega_N \omega_{g_N} \end{bmatrix} \tag{4.116b}$$

$$\boldsymbol{\tau}_g = \begin{bmatrix} J_{g_1} \ddot{\gamma}_1 - J_{s_1} \Omega_1 \omega_{t_1} \\ \vdots \\ J_{g_N} \ddot{\gamma}_N - J_{s_N} \Omega_N \omega_{t_N} \end{bmatrix} \tag{4.116c}$$

The rotational equations of motion for a rigid body containing N VSCMGs is then written compactly as[9]

$$[I]\dot{\boldsymbol{\omega}} = -\boldsymbol{\omega} \times [I]\boldsymbol{\omega} - [G_s]\boldsymbol{\tau}_s - [G_t]\boldsymbol{\tau}_t - [G_g]\boldsymbol{\tau}_g + \boldsymbol{L} \tag{4.117}$$

The rotational kinetic energy T of a rigid spacecraft with N VSCMGs is given by

$$T = \frac{1}{2}\boldsymbol{\omega}^T [I_s]\boldsymbol{\omega} + \frac{1}{2} \sum_{i=1}^{N} J_{s_i} \left(\Omega_i + \omega_{s_i} \right)^2 + J_{t_i} \omega_{t_i}^2 + J_{g_i} \left(\omega_{g_i} + \dot{\gamma}_i \right)^2 \tag{4.118}$$

The kinetic energy rate, also known as the work rate, is found after differentiating Eq. (4.118) and performing some lengthy algebra to be

$$\dot{T} = \boldsymbol{\omega}^T \boldsymbol{L} + \sum_{i=1}^{N} \dot{\gamma}_i u_{g_i} + \Omega_i u_{s_i} \tag{4.119}$$

This energy rate for this system of rigid bodies was apriori known from the Work-Energy-Rate principle[11] shown in Eq. (4.53) and is thus a validation of the presented equations of motion. Also, checking the kinetic energy time history is a convenient way to check the accuracy of the numerical simulations.

Example 4.5: The equations of motion of a spacecraft with several single-gimbal CMGs or RWs attached can easily be extracted from Eq. (4.117). This example discusses the generalization of Eq. (4.113) for the case of multiple RWs. The inertia matrix $[I_{RW}]$ is now defined as

$$[I_{RW}] = [I_s] + \sum_{i=1}^{N} \left(J_{t_i} \hat{\boldsymbol{g}}_{t_i} \hat{\boldsymbol{g}}_{t_i}^T + J_{g_i} \hat{\boldsymbol{g}}_{g_i} \hat{\boldsymbol{g}}_{g_i}^T \right)$$

Let the i-th components of the vector $\boldsymbol{u}_s$ be the RW motor torques given in Eq. (4.112) and let the vector $\boldsymbol{h}_s$ be defined as

$$\boldsymbol{h}_s = \begin{pmatrix} \vdots \\ J_{s_i} \left(\omega_{s_i} + \Omega_i \right) \\ \vdots \end{pmatrix}$$

The desired equations of motion are then written as

$$[I_{RW}]\dot{\boldsymbol{\omega}} = -\boldsymbol{\omega} \times [I_{RW}]\boldsymbol{\omega} - \boldsymbol{\omega} \times [G_s]\boldsymbol{h}_s - [G_s]\boldsymbol{u}_s + \boldsymbol{L} \tag{4.120}$$

Often three RWs are built into a spacecraft such that their spin axis $\hat{\boldsymbol{g}}_{s_i}$ align with the principal body axis. For this special case the matrix $[G_s]$ is reduced to the identity matrix. The equations of motion for this non-redundant RW setup are

$$[I_{RW}]\dot{\boldsymbol{\omega}} = -\boldsymbol{\omega} \times [I_{RW}]\boldsymbol{\omega} - \boldsymbol{\omega} \times \boldsymbol{h}_s - \boldsymbol{u}_s + \boldsymbol{L}$$

While the redundant setup with a general $[G_s]$ matrix can accommodate RW failures, the minimal RW setup cannot have one control wheel fail and still perform general three dimensional rotations.

4.4 Gravity Gradient Satellite

An object in Low Earth Orbit (LEO) does not experience the same gravitational pull on all parts of its body. As is described in Newton's Law of Universal Gravitation in Eq. (2.4), portions closer to Earth are attracted more strongly than portions further removed. While this force is relatively weak, it is enough to stabilize some satellites in a vertical orientation relative to the local horizon. The oldest and most famous gravity gradient stabilized satellite in Earth's orbit is the moon. This section will study the effect of the gravity gradient torque on a rigid object in an inverse square gravity field.

4.4.1 Gravity Gradient Torque

Assume an object $\mathcal{B}$ is in LEO and its center of mass has the inertial position vector $\boldsymbol{R}_c$ relative to Earth's center. Let the vector $\boldsymbol{L}_G$ be the external gravity

torque experienced by a rigid object measured about its center of mass. For a solid body this torque is defined through Eq. (2.99) to be

$$\boldsymbol{L}_G = \int_{\mathcal{B}} \boldsymbol{r} \times d\boldsymbol{F}_G \tag{4.121}$$

where the vector $\boldsymbol{r}$ is the position vector of an infinitesimal body element relative to the center of mass and $\boldsymbol{F}_G$ is the gravitational attraction experienced by this element. Using Newton's Gravitational Law in Eq. (2.4) this force is written as

$$d\boldsymbol{F}_G = -\frac{GM_e}{|\boldsymbol{R}|^3}\boldsymbol{R}dm \tag{4.122}$$

where M_e is Earth's mass, dm is the body element mass and $\boldsymbol{R}$ is its inertial position vector measured from Earth's center.

$$\boldsymbol{R} = \boldsymbol{R}_c + \boldsymbol{r} \tag{4.123}$$

Substituting Eq. (4.122) into the $\boldsymbol{L}_G$ expression yields

$$\boldsymbol{L}_G = -\int_{\mathcal{B}} \boldsymbol{r} \times \frac{GM_e}{|\boldsymbol{R}|^3}\left(\boldsymbol{R}_c + \boldsymbol{r}\right) dm \tag{4.124}$$

The $\boldsymbol{R}_c$ vector is constant within this integral and can be taken outside. After cancelling the $\boldsymbol{r} \times \boldsymbol{r}$ term and rearranging the expression slightly, the torque $\boldsymbol{L}_G$ is written as

$$\boldsymbol{L}_G = GM_e\boldsymbol{R}_c \times \int_{\mathcal{B}} \frac{\boldsymbol{r}}{|\boldsymbol{R}|^3} dm \tag{4.125}$$

To evaluate the integral, the integral denominator $|\boldsymbol{R}|^3$ must be simplified. Let R_c and r be the magnitude of the vector $\boldsymbol{R}_c$ and $\boldsymbol{r}$ respectively.

$$\begin{aligned} |\boldsymbol{R}|^{-3} = |\boldsymbol{R}_c + \boldsymbol{r}|^{-3} &= \left(R_c^2 + 2\boldsymbol{R}_c \cdot \boldsymbol{r} + r^2\right)^{-3/2} \\ &= \frac{1}{R_c^3}\left(1 + \frac{2\boldsymbol{R}_c \cdot \boldsymbol{r}}{R_c^2} + \left(\frac{r}{R_c}\right)^2\right)^{-3/2} \\ &\approx \frac{1}{R_c^3}\left(1 - \frac{3\boldsymbol{R}_c \cdot \boldsymbol{r}}{R_c^2} + \cdots\right) \end{aligned} \tag{4.126}$$

The approximation in the last step was performed using a binomial expansion and dropping the higher order terms. Substituting Eq. (4.126) into Eq. (4.125) yields

$$\boldsymbol{L}_G = \frac{GM_e}{R_c^3}\boldsymbol{R}_c \times \int_{\mathcal{B}} \boldsymbol{r}\left(1 - \frac{3\boldsymbol{R}_c \cdot \boldsymbol{r}}{R_c^2}\right) dm \tag{4.127}$$

Since by definition the vector $\boldsymbol{r}$ is measured relative to the center of mass, the term $\int_{\mathcal{B}} \boldsymbol{r} dm$ is zero and drops out. The gravity gradient torque $\boldsymbol{L}_G$ can then be written as

$$\boldsymbol{L}_G = \frac{3GM_e}{R_c^5}\boldsymbol{R}_c \times \int_{\mathcal{B}} -\boldsymbol{r}\left(\boldsymbol{r} \cdot \boldsymbol{R}_c\right) dm \tag{4.128}$$

Using the vector identity

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \boldsymbol{c})\, \boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\, \boldsymbol{c} \tag{4.129}$$

the integrant is rewritten in the form

$$\boldsymbol{L}_G = \frac{3GM_e}{R_c^5} \boldsymbol{R}_c \times \int_{\mathcal{B}} - (\boldsymbol{r} \times \boldsymbol{r} \times \boldsymbol{R}_c + (\boldsymbol{r} \cdot \boldsymbol{r})\, \boldsymbol{R}_c)\, dm \tag{4.130}$$

After using the definition of the tilde matrix in Eq. (3.23), the torque $\boldsymbol{L}_G$ expression is written as

$$\boldsymbol{L}_G = \frac{3GM_e}{R_c^5} \boldsymbol{R}_c \times \left(\int_{\mathcal{B}} -[\tilde{\boldsymbol{r}}][\tilde{\boldsymbol{r}}] dm \right) \boldsymbol{R}_c - \frac{3GM_e}{R_c^5} \left(\int_{\mathcal{B}} r^2 dm \right) \boldsymbol{R}_c \times \boldsymbol{R}_c \tag{4.131}$$

Note that the first integrant is equal to the inertia matrix definition in Eq. (4.14), while the second cross product term is zero. Therefore, the gravity gradient torque vector $\boldsymbol{L}_G$ acting on a rigid body in an inverse square gravity field is written in its most general form as[1, 10]

$$\boldsymbol{L}_G = \frac{3GM_e}{R_c^5} \boldsymbol{R}_c \times [I] \boldsymbol{R}_c \tag{4.132}$$

The only approximation made was the truncation of the binomial series. If the inertia matrix $[I]$ is assumed to be of diagonal form, i.e. the chosen coordinates axes are the principal body axes, then the gravity gradient torque expression can be further simplified. Let the center of mass vector $\boldsymbol{R}_c$ be given in body frame components as

$$\boldsymbol{R}_c = {}^{\mathcal{B}}\begin{pmatrix} R_{c_1} \\ R_{c_2} \\ R_{c_3} \end{pmatrix} \tag{4.133}$$

After carrying out the algebra in Eq. (4.132), the simplified gravity gradient torque vector is given in $\mathcal{B}$ frame components as

$$\begin{pmatrix} L_{G_1} \\ L_{G_2} \\ L_{G_3} \end{pmatrix} = \frac{3GM_e}{R_c^5} \begin{pmatrix} R_{c_2} R_{c_3} (I_{33} - I_{22}) \\ R_{c_1} R_{c_3} (I_{11} - I_{33}) \\ R_{c_1} R_{c_2} (I_{22} - I_{11}) \end{pmatrix} \tag{4.134}$$

Studying Eq. (4.134) it is clear that that several situations will lead to no gravity gradient torque being produced on a spacecraft. Symmetric spacecraft with $I_{11} = I_{22} = I_{33}$ have a zero torque $\boldsymbol{L}_G$ vector. Assume that the i-th principal body axis is a symmetry axis, then the spacecraft will not experience any gravity gradient torque about its i-th body axis. Lastly, if the center of mass vector $\boldsymbol{R}_c$ is parallel with any of the principal body axes, then two of the three R_{c_i} vector components will be zero, which results in the $\boldsymbol{L}_G$ vector itself being zero.

4.4.2 Rotational - Translational Motion Coupling

To study the coupling effect of the translational motion of a spacecraft $\mathcal{B}$ versus its rotational motion and vice versa, the total gravity force vector $\boldsymbol{F}_G$ acting on the rigid body needs to be investigated. This force vector determines the center of mass motion of the spacecraft. Using Eqs. (2.75) and (4.122) it is written as

$$\boldsymbol{F}_G = \int_{\mathcal{B}} d\boldsymbol{F}_G = -GM_e \int_{\mathcal{B}} \frac{\boldsymbol{R}}{|\boldsymbol{R}|^3} dm \tag{4.135}$$

After using the approximation for $|\boldsymbol{R}|^{-3}$ given in Eq. (4.126) and the definition of $\boldsymbol{R}$ and expanding the resulting product, the gravity force vector $\boldsymbol{F}_G$ is rewritten as

$$\boldsymbol{F}_G = -\frac{GM_e}{R_c^3}\left(\int_{\mathcal{B}} \boldsymbol{r} dm - \frac{3}{R_c^2}\int_{\mathcal{B}} (\boldsymbol{r}\cdot\boldsymbol{R}_c)\,\boldsymbol{r} dm + \boldsymbol{R}_c \int_{\mathcal{B}} dm \right. \\ \left. -\frac{3}{R_c^2}\int_{\mathcal{B}} (\boldsymbol{R}_c\cdot\boldsymbol{r})\,\boldsymbol{R}_c dm\right) \tag{4.136}$$

Note that the first term in the parenthesis is zero due to the definition of the spacecraft center of mass. Let m be the total spacecraft mass. After using the vector identity in Eq. (4.129) $\boldsymbol{F}_G$ is written as

$$\boldsymbol{F}_G = -\frac{GM_e}{R_c^3}\left(m\boldsymbol{R}_c - \frac{3}{R_c^2}\int_{\mathcal{B}} \left(\boldsymbol{r}\times\boldsymbol{r}\times\boldsymbol{R}_c + r^2\boldsymbol{R}_c\right) dm \right. \\ \left. -\frac{3}{R_c^2}\boldsymbol{R}_c\cdot\left(\int_{\mathcal{B}} \boldsymbol{r} dm\right)\boldsymbol{R}_c\right) \tag{4.137}$$

where the last term in the parenthesis is zero due to the definition of the center of mass. Using the definition of the inertia matrix in Eq. (4.14) the gravity force vector of a rigid body in an inverse square gravitational field is expressed

$$\boldsymbol{F}_G = -\frac{GM_e}{R_c^3}\left(m\boldsymbol{R}_c + \frac{3}{R_c}\left([I] - \int_{\mathcal{B}} r^2 dm\right)\frac{\boldsymbol{R}_c}{R_c}\right) \tag{4.138}$$

Observe that the first term in Eq. (4.138) is the standard gravitational force experienced by a body of mass m in Earth's inverse square gravitational field. The second term is due to the gravitational gradient effect. Since typical man-made spacecraft are of rather small size compared to their orbital radius R_c, the gravitational gradient term is always very small compared to the gravitational attraction of the center of mass. If the spacecraft $\mathcal{B}$ is rotating about its center of mass, then the unit vector $\boldsymbol{R}_c/Rc$ components mapped into the body frame $\mathcal{B}$ will vary with time. However, this rotation to translation coupling effect is negligible and can be ignore for a first order approximation. To study the effect of the spacecraft translation onto its rotation Eq. (4.132) is used. Clearly the orbit radius R_c directly effects the magnitude of the gravity gradient torque vector experienced. While $\boldsymbol{L}_G$ is typically a small quantity, its effect can be quite substantial on elliptic orbits where R_c can vary greatly with time.

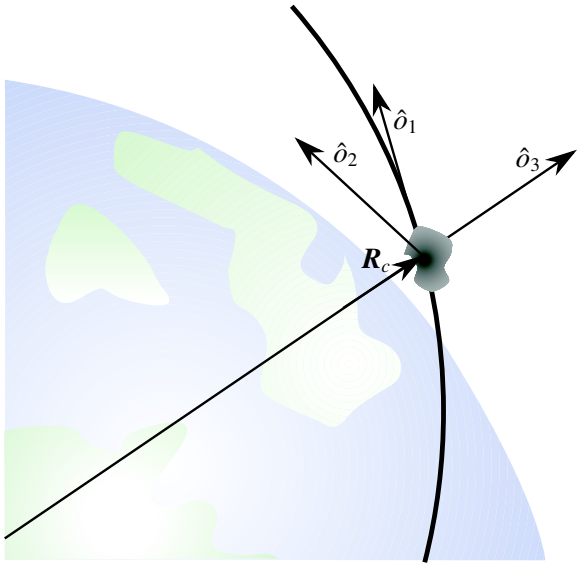


Figure 4.11: Spacecraft in Circular Orbit

4.4.3 Small Departure Motion about Equilibrium Attitudes

Assume a spacecraft $\mathcal{B}$ is in a circular orbit $\mathcal{O}$ about Earth. The orbit frame orientation is defined through the unit vectors $\hat{\boldsymbol{o}}_1$, $\hat{\boldsymbol{o}}_2$ and $\hat{\boldsymbol{o}}_3$ as shown in Figure 4.11. Let the vector $\boldsymbol{R}_c = R_c\hat{\boldsymbol{o}}_3$ define the spacecraft position measured from the Earth's center. The gravity gradient torque was found to be zero whenever the principal body axis where aligned with the orbit frame axis. Therefore, we choose for the spacecraft body fixed axes $\{\hat{\boldsymbol{b}}\}$ to be nominally aligned with the orbit frame $\{\hat{\boldsymbol{o}}\}$. There are 24 possible orientations for a rigid spacecraft to have its principal axes aligned with another reference frame. We choose for each $\hat{\boldsymbol{b}}_i$ vector to be in the $\hat{\boldsymbol{o}}_i$ direction.

Since only small spacecraft rotations about the $\{\hat{\boldsymbol{o}}\}$ frame are considered, the $(3-2-1)$ Euler angles $(\psi, \theta, \phi)$are chosen to describe the relative spacecraft attitude to the orbit frame. The orbit frame angular velocity vector relative to the inertial frame is given by

$$\boldsymbol{\omega}_{\mathcal{O}/\mathcal{N}} = \Omega\hat{\boldsymbol{o}}_2 \tag{4.139}$$

where the magnitude Ω is given by Kepler's equation to be[2, 12]

$$\Omega^2 = \frac{GM_e}{R_c^3} \tag{4.140}$$

The relative angular velocity vector $\boldsymbol{\omega}_{\mathcal{B}/\mathcal{O}}$ is written in terms of the yaw, pitch

and roll rates using Eq. (3.55) as

$$
{}^{\mathcal{B}}\boldsymbol{\omega}_{\mathcal{B}/\mathcal{O}} = \begin{bmatrix} -\sin\theta & 0 & 1 \\ \sin\phi\cos\theta & \cos\phi & 0 \\ \cos\phi\cos\theta & -\sin\phi & 0 \end{bmatrix} \begin{pmatrix} \dot{\psi} \\ \dot{\theta} \\ \dot{\phi} \end{pmatrix} \tag{4.141}
$$

The spacecraft angular velocity vector relative to the inertial frame is

$$
\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}} = \boldsymbol{\omega}_{\mathcal{B}/\mathcal{O}} + \boldsymbol{\omega}_{\mathcal{O}/\mathcal{N}} \tag{4.142}
$$

Using Eq. (3.33), the direction cosine matrix $[BO]$ which relates the $\mathcal{O}$ frame to the $\mathcal{B}$ frame is written in terms of the $(3-2-1)$ Euler angles as

$$
[BO] = \begin{bmatrix} c\theta c\psi & c\theta s\psi & -s\theta \\ s\phi s\theta c\psi - c\phi s\psi & s\phi s\theta s\psi + c\phi c\psi & s\phi c\theta \\ c\phi s\theta c\psi + s\phi s\psi & c\phi s\theta s\psi - s\phi c\psi & c\phi c\theta \end{bmatrix} \tag{4.143}
$$

Expressing $\hat{\boldsymbol{o}}_2$ in terms of $\hat{\boldsymbol{b}}_i$ and substituting Eqs. (4.139) and (4.141) into Eq. (4.142), the spacecraft body angular velocity vector is expressed as

$$
{}^{\mathcal{B}}\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}} = {}^{\mathcal{B}}\begin{pmatrix} \dot{\phi} - \sin\theta\dot{\psi} + \Omega\cos\theta\sin\psi \\ \sin\phi\cos\theta\dot{\psi} + \cos\phi\dot{\theta} + \Omega\left(\sin\phi\sin\theta\sin\psi + \cos\phi\cos\psi\right) \\ \cos\phi\cos\theta\dot{\psi} - \sin\phi\dot{\theta} + \Omega\left(\cos\phi\sin\theta\sin\psi - \sin\phi\cos\psi\right) \end{pmatrix} \tag{4.144}
$$

From here on the angular velocity $\boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}$ is abbreviated as $\boldsymbol{\omega}$. Eq. (4.144) can be rewritten to yield the Euler angle rates in terms of the body angular velocity $\boldsymbol{\omega}$ and the orbital rate Ω. Using the $[B(\psi,\theta,\phi)]$ matrix definition in Eq. (3.56) we find

$$
\begin{pmatrix} \dot{\psi} \\ \dot{\theta} \\ \dot{\phi} \end{pmatrix} = [B(\psi,\theta,\phi)]\boldsymbol{\omega} - \frac{\Omega}{\cos\theta}\begin{pmatrix} \sin\theta\sin\psi \\ \cos\theta\cos\psi \\ \sin\psi \end{pmatrix} \tag{4.145}
$$

It is rather surprising that the algebra reduces to the relatively simple form of Eq. (4.145). Using the Euler parameter vector $\boldsymbol{\beta}$ or the MRP vector $\boldsymbol{\sigma}$ as the relative attitude coordinates to the orbit frame, even simpler attitude coordinate rate expressions are found. Using Eq. (3.105), the Euler parameter rates are simply

$$
\dot{\boldsymbol{\beta}} = \frac{1}{2}[B(\boldsymbol{\beta})]\boldsymbol{\omega} - \frac{\Omega}{2}\begin{pmatrix} -\beta_2 \\ \beta_3 \\ \beta_0 \\ -\beta_1 \end{pmatrix} = \frac{1}{2}[B(\boldsymbol{\beta})]\boldsymbol{\omega} - \frac{\Omega}{2}\boldsymbol{g}(\boldsymbol{\beta}) \tag{4.146}
$$

Note that the vector $\boldsymbol{g}(\boldsymbol{\beta})$ is perpendicular to $\boldsymbol{\beta}$. This makes intuitively sense since all valid Euler parameters sets must lie on the four-dimensional unit hypersphere surface. Finding the MRP rates with the orbital motion included is greatly simplified using the identity

$$
[B(\boldsymbol{\sigma})][BO(\boldsymbol{\sigma})] = [B(\boldsymbol{\sigma})]^T \tag{4.147}
$$

where the matrix $[B(\boldsymbol{\sigma})]$ is defined in Eq. (3.150) and $[BO(\boldsymbol{\sigma})]$ is the direction cosine matrix relating the orbit frame $\mathcal{O}$ to the body frame $\mathcal{B}$ in terms of the MRP vector $\boldsymbol{\sigma}$. This relationship is developed in Appendix D. The desired MRP rates are then given by

$$\dot{\boldsymbol{\sigma}} = \frac{1}{4}[B(\boldsymbol{\sigma})]\boldsymbol{\omega} - \frac{\Omega}{4}\begin{pmatrix} 2\left(\sigma_1\sigma_2 + \sigma_3\right) \\ 2\sigma_2^2 + 1 - \sigma^2 \\ 2\left(\sigma_2\sigma_3 - \sigma_1\right) \end{pmatrix} \tag{4.148}$$

Linearizing Eq. (4.144) about zero yaw, pitch and roll angles while considering Ω to be large we find

$$\boldsymbol{\omega} \approx \ {}^{\mathcal{B}}\begin{pmatrix} \dot{\phi} + \Omega\psi \\ \dot{\theta} + \Omega \\ \dot{\psi} - \Omega\phi \end{pmatrix} \tag{4.149}$$

The angular acceleration vector $\dot{\boldsymbol{\omega}}$ is then given by

$$\dot{\boldsymbol{\omega}} \approx \ {}^{\mathcal{B}}\begin{pmatrix} \ddot{\phi} + \Omega\dot{\psi} \\ \ddot{\theta} \\ \ddot{\psi} - \Omega\dot{\phi} \end{pmatrix} \tag{4.150}$$

since $\dot{\Omega}$ is zero for a circular orbit. Before we can write out the linearized equations of motion, the gravity gradient torque vector $\boldsymbol{L}_G$ in Eq. (4.134) still needs to be linearized. The center of mass position vector $\boldsymbol{R}_c$ is given in $\mathcal{O}$ frame components as

$${}^{\mathcal{O}}\boldsymbol{R}_c = \ {}^{\mathcal{O}}\begin{pmatrix} 0 \\ 0 \\ R_c \end{pmatrix} \tag{4.151}$$

After using Eq. (4.143) to map $\boldsymbol{R}_c$ into $\mathcal{B}$ frame components, the position vector is written as

$${}^{\mathcal{B}}\begin{pmatrix} R_{c_1} \\ R_{c_2} \\ R_{c_3} \end{pmatrix} = \ {}^{\mathcal{B}}\begin{pmatrix} -\sin\theta \\ \sin\phi\cos\theta \\ \cos\phi\cos\theta \end{pmatrix} R_c \tag{4.152}$$

Substituting these R_{c_i} into Eq. (4.134), the gravity gradient torque $\boldsymbol{L}_G$ is expressed in terms of the $(3-2-1)$ Euler angels as

$${}^{\mathcal{B}}\boldsymbol{L}_G = \frac{3}{2}\Omega^2 \ {}^{\mathcal{B}}\begin{pmatrix} \left(I_{33} - I_{22}\right)\cos^2\theta\sin 2\phi \\ -\left(I_{11} - I_{33}\right)\cos\phi\sin 2\theta \\ -\left(I_{22} - I_{11}\right)\sin\phi\sin 2\theta \end{pmatrix} \tag{4.153}$$

Note that the gravity torque vector $\boldsymbol{L}_G$ does not explicitly depend on the yaw angle ψ. Linearizing Eq. (4.153) yields

$${}^{\mathcal{B}}\boldsymbol{L}_G \approx 3\Omega^2 \ {}^{\mathcal{B}}\begin{pmatrix} \left(I_{33} - I_{22}\right)\phi \\ -\left(I_{11} - I_{33}\right)\theta \\ 0 \end{pmatrix} \tag{4.154}$$

It is interesting to note that the linearized $\boldsymbol{L}_G$ does have any torque components about the third body axis $\hat{\boldsymbol{b}}_3$. Further, for the pitch and roll components of $\boldsymbol{L}_G$ to be stabilizing, we find that $I_{22} > I_{33}$ and $I_{22} > I_{11}$ must be true. Therefore I_{22} must be the largest principal inertia of the spacecraft for these small gravity gradient torque induced oscillations to be stable. After substituting Eqs. (4.149), (4.150) and (4.154) into (4.33), the equations of motion about each body axis are expressed as

$$I_{11}\left(\ddot{\phi}+\Omega\dot{\psi}\right) = -\left(I_{33}-I_{22}\right)\left(\dot{\theta}+\Omega\right)\left(\dot{\psi}-\Omega\phi\right)+3\Omega^2\left(I_{33}-I_{22}\right)\phi \quad (4.155)$$

$$I_{22}\ddot{\theta} = -\left(I_{11}-I_{33}\right)\left(\dot{\psi}-\Omega\phi\right)\left(\dot{\phi}+\Omega\psi\right)-3\Omega^2\left(I_{11}-I_{33}\right)\theta \quad (4.156)$$

$$I_{33}\left(\ddot{\psi}-\Omega\dot{\phi}\right) = -\left(I_{22}-I_{11}\right)\left(\dot{\phi}+\Omega\psi\right)\left(\dot{\theta}+\Omega\right) \quad (4.157)$$

After neglecting the higher order terms, the linearized spacecraft equations of motion can be decoupled into the pitch and roll / yaw modes. The linearized pitch equation is

$$\ddot{\theta}+3\Omega^2\left(\frac{I_{11}-I_{33}}{I_{22}}\right)\theta = 0 \quad (4.158)$$

which is the dynamical equivalent of a simple spring-mass system. It is immediately clear from linear control theory that for the pitch mode to be stable

$$I_{11} \geq I_{33} \quad (4.159)$$

must be true. The coupled roll-yaw equations of motion are written as

$$\begin{pmatrix}\ddot{\phi}\\ \ddot{\psi}\end{pmatrix}+\begin{bmatrix}0 & \Omega\left(1-k_Y\right)\\ \Omega\left(k_R-1\right) & 0\end{bmatrix}\begin{pmatrix}\dot{\phi}\\ \dot{\psi}\end{pmatrix}+\begin{bmatrix}4\Omega^2k_Y & 0\\ 0 & \Omega^2k_R\end{bmatrix}\begin{pmatrix}\phi\\ \psi\end{pmatrix}=0 \quad (4.160)$$

where the inertia ratios k_R and k_Y are defined as

$$k_R = \frac{I_{22}-I_{11}}{I_{33}} \quad (4.161)$$

$$k_Y = \frac{I_{22}-I_{33}}{I_{11}} \quad (4.162)$$

To determine stability conditions of the roll-yaw motion, the roots λ_i of the characteristic equation of Eq. (4.160) must be investigated. The characteristic equation of Eq. (4.160) is given by

$$\begin{vmatrix}\lambda^2+4\Omega^2k_Y & \lambda\Omega\left(1-k_Y\right)\\ \lambda\Omega\left(k_R-1\right) & \lambda^2+\Omega^2k_R\end{vmatrix}=0 \quad (4.163)$$

which can be expanded to

$$\lambda^4+\lambda^2\Omega^2\left(1+3k_Y+k_Yk_R\right)+4\Omega^4k_Yk_R = 0 \quad (4.164)$$

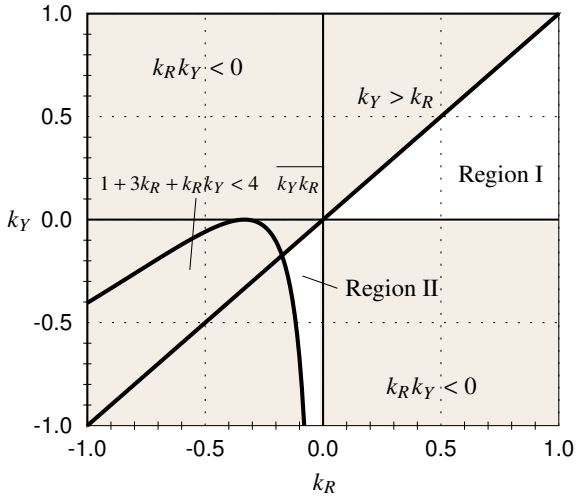


Figure 4.12: Linearized Gravity Gradient Spacecraft Stability Regions

The roll-yaw equations of motion in Eq. (4.160) are stable if none of the roots λ_i have any positive real parts. Note that the characteristic equations is quadratic in λ^2 and can be solved using the quadratic solution formula. No root λ_i^2 can be positive since the corresponding set $\lambda_{i_1} = +\sqrt{\lambda_i^2}$ and $\lambda_{i_2} = -\sqrt{\lambda_i^2}$ would contain a real, positive root. To guarantee that all λ_i^2 terms are negative and real, it is necessary and sufficient that

$$1 + 3k_R + k_Y k_R > 4\sqrt{k_Y k_R} \tag{4.165}$$

$$k_R k_Y > 0 \tag{4.166}$$

These two stability conditions have to be satisfied along with the pitch motion stability condition in Eq. (4.159). This condition is expressed in terms of the inertia ratios k_R and k_Y as

$$k_Y < k_R \tag{4.167}$$

All three stability conditions are shown in Figure 4.12. The unstable regions are shaded while the stable regions I and II are white.

Since all four roots of the characteristic equation in Eq. (4.164) are imaginary, only neutral stability of the *linearized* system is guaranteed. The actual nonlinear system may or may not be stable. It turns out the triangular region I represents the truly stable region, while the small white region II in the third quadrant is unstable if damping effects are included.[1] To prove this rigorously the dynamics of the center manifold would have to be studied which is beyond the scope of this book. The stability conditions in Eqs. (4.166) and (4.167) for region I can be written directly in terms of the principal spacecraft inertias I_{ii} as

$$I_{22} \geq I_{11} \geq I_{33} \tag{4.168}$$

Therefore, for the spacecraft attitude assumed at the beginning of this section to be stable in the presence of gravity gradient torques, the pitch axis inertia must be largest and the yaw axis inertia the smallest. If the spacecraft is aligned with the $\mathcal{O}$ frame, then its only angular velocity is $\boldsymbol{\omega}_{\mathcal{O}/\mathcal{N}} = \Omega\hat{\boldsymbol{o}}_2 = \Omega\hat{\boldsymbol{b}}_2$. As was shown in Eq. (4.64), having a pure spin about the largest moment of inertia corresponds to *a minimum kinetic energy* condition. The neutrally stable region II would correspond to having I_{22} be less than I_{11} and I_{33}. This indicates that the spacecraft is nominally rotating about the axis of least inertia which is a *maximum kinetic energy* state. As is shown in Figure 4.7, in the presence of damping this spin will degrade in the presence of damping to a pure spin about the axis of maximum inertia (i.e. minimum kinetic energy state). Gravity gradient satellites are therefore typically long and skinny structures flying in an "upright" attitude relative to the local horizon.

Problems

4.1 Starting with Eq. (4.9) and using Eqs. (4.10) through (4.12), verify Eq. (4.13).

4.2 Find the moment of inertia matrix of a box with side lengths $2a$, $2b$ and $2c$. The cube material has a unit density. The cube center is at the cartesian coordinate system origin and all its sides are aligned perpendicular to coordinate axes.

4.3 Let the unit axis of the rigid body coordinate frame $\mathcal{B} : \{\hat{\boldsymbol{b}}_1, \hat{\boldsymbol{b}}_2, \hat{\boldsymbol{b}}_3\}$ be given in terms of inertial frame $\mathcal{N}$ components as

$$\hat{\boldsymbol{b}}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \hat{\boldsymbol{b}}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad \hat{\boldsymbol{b}}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

The inertia matrix in terms of $\mathcal{B}$ frame components is given by

$$^{\mathcal{B}}[I] = \begin{bmatrix} 15 & 0 & 0 \\ 0 & 11 & 5 \\ 0 & 5 & 16 \end{bmatrix}$$

a) Find the rotation matrix $[C]$ that will map the $\mathcal{B}$ frame into a new frame $\mathcal{F}$ such that $^{\mathcal{F}}[I]$ is diagonal.

b) What are the principal inertias of this rigid body.

c) What are the principal body axis expressed in $\mathcal{N}$ frame components.

4.4 A rigid body at an orientation of $\psi = 5^{\text{ffi}}$, $\theta = 10^{\text{ffi}}$ and $\phi = -3^{\text{ffi}}$ has a total mass $M = 100$ kg and an inertia matrix

$$[I] = \begin{bmatrix} 34 & 1 & 6 \\ 1 & 15 & 3 \\ 6 & 3 & 10 \end{bmatrix} kgm^2$$

The center of mass is moving at $5m/s$ and the (3-2-1) Euler angle yaw, pitch and roll rates are $\dot{\psi} = -1^{\text{ffi}}/s$, $\dot{\theta} = 1^{\text{ffi}}/s$ and $\dot{\phi} = 4^{\text{ffi}}/s$. Find the total kinetic energy of this rigid body.

4.5 A solid disk with mass m and radius r is rolling under the influence of a constant gravity field inside a cylinder of radius L as shown in Figure 4.13.

a) Find the angular momentum vector $\boldsymbol{H}_O$ relative to the cylinder center O.

b) Find the equations of motion of the disk.

c) What is the natural frequency of the motion assuming that θ is small?

d) Given an initial angular position $\theta(0)$ and rate $\dot{\theta}(0)$, find the angular velocity $\dot{\theta}$ when $\theta = 0°$.

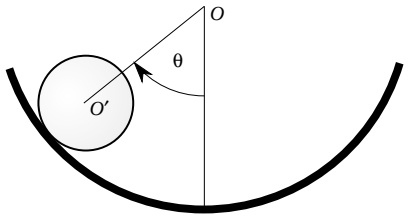


Figure 4.13: Solid Disk Rolling inside a Cylinder

4.6 A slender rod of length L and mass m is standing vertically on a smooth, level surface. After it is slightly disturbed, the rod will fall on the ground.

a) Find the differential equations of motion of the rod where the angle θ defines the orientation of the rod.

b) Find a relationship between the angular rate $\dot{\theta}$ and the orientation angle θ.

4.7 A slender rod of length L and mass m is standing vertically on a rough, level surface with a friction coefficient μ. After it is slightly disturbed, the rod will start to rotate towards the ground. Find a relationship between the friction coefficient μ and the rod orientation angle θ where the rod starts to slip.

4.8 A rigid link of mass M and length L is attached to the ceiling as shown in Figure 4.14. A mass m is attached to the lower end of the link. Find the differential equations of motion of this pendulum system and its natural frequency.

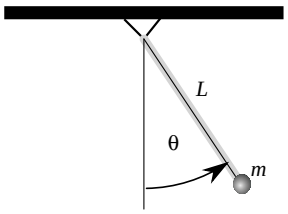


Figure 4.14: Rigid Link Pendulum with Mass Attached at End

4.9 The principal inertias of a rigid satellite are given by

$$I_1 = 210kgm^2 \quad I_2 = 200kgm^2 \quad I_3 = 118kgm^2$$

At time t_0 the body angular velocity vector is $\boldsymbol{\omega} = (0.2, 0.15, -0.18)^T rad/s$. Numerically solve the resulting torque-free motion for 30 seconds and plot the resulting attitude in terms of the (3-2-1) Euler angles.

4.10 A solid cylinder of mass m, radius a, and length l is pivoted about a transverse axis (B-B') through its center of mass as shown in Figure 4.15. The axis (A-A') rotates with a constant angular velocity Ω. Assume $l > \sqrt{3}a$.

a) Find the frequency ω_n of small oscillations about $\theta = \frac{\pi}{2}$.

b) What is the angular velocity $\dot{\theta}^*$ when $\theta = \frac{\pi}{2}$, if the cylinder is released from $\theta = 0$ with a very small positive value of $\dot{\theta}_0$? Determine $\dot{\theta}^*$ as a function of m, a, l and Ω.

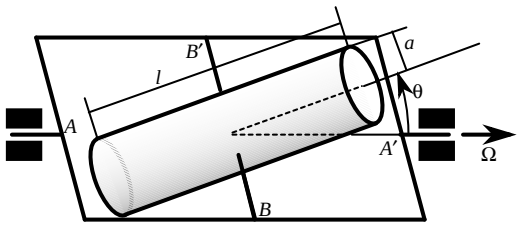


Figure 4.15: Solid Cylinder in a Two Hinge Gyroscope

4.11 ♣ Consider the free rotational motion of an axially symmetric rigid body with $I_a = 2I_t$, where I_a is the axially moment of inertia and I_t is the transverse moment of inertia.

a) What is the largest possible value of the angle between $\boldsymbol{\omega}$ and $\boldsymbol{H}$? *Hint: Consider the angular momentum vector* $\boldsymbol{H}$ *fixed and vary the kinetic energy* T.

b) Find the critical value of kinetic energy which results in the largest angle between $\boldsymbol{\omega}$ and $\boldsymbol{H}$.

4.12 ♣ A vertical shaft is driving two grinding wheels by rotating at a constant angular rate Ω as shown in Figure 4.16. Each grinding wheel and its support shaft have a mass m with the center of mass point located a distance L away from the hinge point. Their inertia about their axis of symmetry is I_s and the transverse inertia is I_t. Due to the level of the floor, the support shafts are raised by an able α. Assume the grinding wheels are rolling without slip.

a) Find the total angular momentum vector of the system about the hinge point O.

b) For a given fixed Ω, how strongly does each grinding wheel push against the wall?

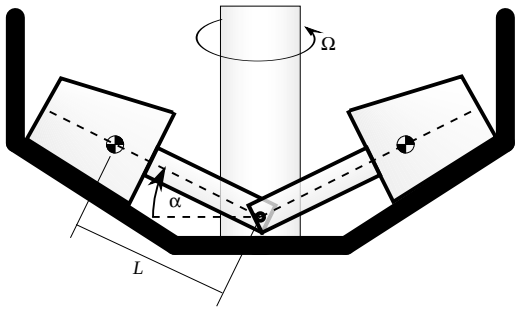


Figure 4.16: Two Grinder Wheels Rolling Rolling about a Driving Shaft

4.13 ♣ An axially symmetric space vehicle with $I_t/I_a = 10$ is undergoing a general torque-free motion. The angle between the angular momentum vector $\boldsymbol{H}$ and the axis of symmetry is 45^{ffi}. At some instant during the motion, symmetrically placed masses are moved slowly toward the axis of symmetry by internal forces. At the end of this process, the rotational kinetic energy is found to be three times its former value, whereas I_a is halved and I_t is 80 percent of its original size. Determine the final angle between $\boldsymbol{H}$ and the symmetry axis.

4.14 A vertical shaft is rotating at a constant angular rate Ω. At its lower end it has a shaft attached to it through a pin connection. A disk is connected to this shaft and is free to spin about the shaft axis. The moment of inertia of the disk/shaft system is I_s and the transverse inertia is I_t. The center of mass of the disk/shaft system is located a distance L away from the hinge point as shown in Figure 4.17. What is the necessary angular velocity vector of the disk relative to the shaft that will maintain a constant angle α?

4.15 Verify the gravity gradient stability conditions in Eqs. (4.165) through (4.167).

Bibliography

[1] Junkins, J. L. and Turner, J. D., *Optimal Spacecraft Rotational Maneuvers*, Elsevier Science Publishers, Amsterdam, Netherlands, 1986.

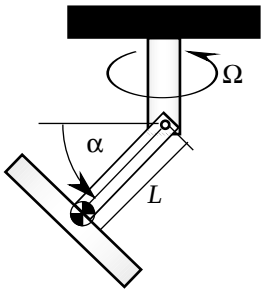


Figure 4.17: Rotating Spinning Disk at Constant Inclination

[2] Wiesel, W. E., *Spaceflight Dynamics*, McGraw-Hill, Inc., New York, 1989.

[3] Oh, H. S. and Vadali, S. R., "Feedback Control and Steering Laws for Spacecraft Using Single Gimbal Control Moment Gyros," *Journal of the Astronautical Sciences*, Vol. 39, No. 2, 1991, pp. 183–203.

[4] Hoelscher, B. R. and Vadali, S. R., "Optimal Open-Loop and Feedback Control Using Single Gimbal Control Moment Gyroscopes," *Journal of the Astronautical Sciences*, Vol. 42, No. 2, 1994, pp. 189–206.

[5] Krishnan, S. and Vadali, S. R., "An Inverse-Free Technique for Attitude Control of Spacecraft Using CMGs," *Acta Astronautica*, Vol. 39, No. 6, 1997, pp. 431–438.

[6] Bedrossian, N. S., *Steering Law Design for Redundant Single Gimbal Control Moment Gyro Systems*, M.S. Thesis, Mechanical Engineering, Massachusetts Institute of Technology, Boston, MA, Aug. 1987.

[7] Schaub, H., Robinett, R. D., and Junkins, J. L., "Globally Stable Feedback Laws for Near-Minimum-Fuel and Near-Minimum-Time Pointing Maneuvers for a Landmark-Tracking Spacecraft," *Journal of the Astronautical Sciences*, Vol. 44, No. 4, 1996, pp. 443–466.

[8] Ford, K. and Hall, C. D., "Flexible Spacecraft Reorientations Using Gimbaled Momentum Wheels," *AAS/AIAA Astrodynamics Specialist Conference*, Sun Valley, Idaho, August 1997, Paper No. 97-723.

[9] Schaub, H., R.Vadali, S., and Junkins, J. L., "Feedback Control Law for Variable Speed Control Moment Gyroscopes," *8th AAS/AIAA Space Flight Mechanics Meeting*, Monterey, California, Feb. 9–11 1998, Paper No. AAS 98-140.

[10] Greenwood, D. T., *Principles of Dynamics*, Prentice-Hall, Inc, Englewood Cliffs, New Jersey, 2nd ed., 1988.

[11] Oh, H. S., Vadali, S. R., and Junkins, J. L., "On the Use of the Work-Energy Rate Principle for Designing Feedback Control Laws," *AIAA Journal of Guidance, Control and Dynamics*, Vol. 15 No. 1, 1992, pp. 272–277.

[12] Battin, R. H., *An Introduction to the Mathematics and Methods of Astrodynamics*, AIAA Education Series, New York, 1987.

Chapter Five

Generalized Methods of Analytical Dynamics

During the mid-19th century, a family of fundamental developments were introduced, led by Lagrange, Hamilton, and Jacobi. These results provided a unifying perspective on analytical mechanics and also stimulated fundamental advances in allied mathematical sub-fields such as variational calculus, differential equations, and topology. The most central developments are embodied in elegant and powerful methods for deriving differential equations of motion by taking gradients of scalar functions they introduced (e.g. the *Lagrangian* and the *Hamiltonian*, closely related to the mechanical kinetic and potential energies of the system), relationships of mechanical system motion to variational principles (e.g. *D'Alembert's Principle* and *Hamilton's Principle*), and efficient methods for accommodating constraints and constraint forces. Collectively, these insights amounted to a revolution in analysis of dynamical systems, even given that their starting point was the summation of the monumental works of Newton, Gauss and Euler. This chapter and the following one provides the most fundamental aspects of these classical developments; we start with Newtonian/Eulerian principles and utilize a system of particles as a conceptual representation for a large class of systems. We introduce virtual and related variational arguments leading to D'Alembert's Principle, Lagrange's Equations, and Hamilton's Principle. Finally, we generalize these particle mechanics results to establish the corresponding developments applicable to systems idealized as collections of particles, rigid bodies, and distributed parameter systems. Examples are utilized throughout this discussion to illustrate the ideas and provide some insights into their utility.

5.1 Generalized Coordinates

Consider the familiar problem of a particle moving relative to an inertially fixed Cartesian coordinate frame. With reference to Fig. 5.1, we introduce

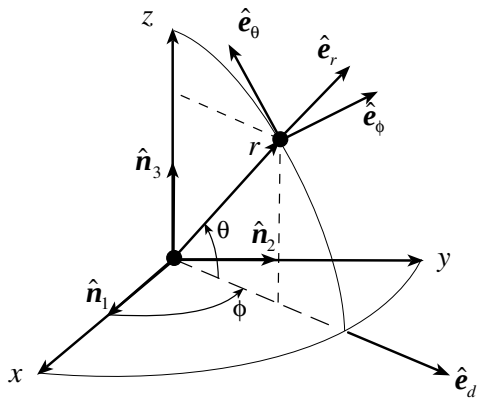


Figure 5.1: Generalized Coordinates of Point P

three classical coordinate choices to locate point P relative to Point O, viz.: Cartesian Coordinates (x, y, z), Spherical Coordinates (r, ϕ, θ), and Cylindrical Coordinates(d, ϕ, z), with the following three corresponding vector representations of the inertial position, velocity and acceleration:

Cartesian Coordinates and $\{\hat{\boldsymbol{n}}_1, \hat{\boldsymbol{n}}_2, \hat{\boldsymbol{n}}_3\}$ vector components:

$$\begin{aligned} \boldsymbol{r} &= x\hat{\boldsymbol{n}}_1 + y\hat{\boldsymbol{n}}_2 + z\hat{\boldsymbol{n}}_3 \\ \dot{\boldsymbol{r}} &= \dot{x}\hat{\boldsymbol{n}}_1 + \dot{y}\hat{\boldsymbol{n}}_2 + \dot{z}\hat{\boldsymbol{n}}_3 \\ \ddot{\boldsymbol{r}} &= \ddot{x}\hat{\boldsymbol{n}}_1 + \ddot{y}\hat{\boldsymbol{n}}_2 + \ddot{z}\hat{\boldsymbol{n}}_3 \end{aligned} \tag{5.1}$$

Spherical Coordinates and $\{\hat{\boldsymbol{e}}_r, \hat{\boldsymbol{e}}_\phi, \hat{\boldsymbol{e}}_\theta\}$ vector components:

$$\begin{aligned} \boldsymbol{r} &= r\hat{\boldsymbol{e}}_r \\ \dot{\boldsymbol{r}} &= \dot{r}\hat{\boldsymbol{e}}_r + r\dot{\theta}\hat{\boldsymbol{e}}_\theta + r\dot{\phi}\cos\theta\hat{\boldsymbol{e}}_\phi \\ \ddot{\boldsymbol{r}} &= (\ddot{r} - r\dot{\theta}^2 - r\dot{\phi}^2\cos^2\theta)\hat{\boldsymbol{e}}_r \\ &\quad + (r\ddot{\phi}\cos\theta + 2\dot{r}\dot{\phi}\cos\theta - 2r\dot{\phi}\dot{\theta}\cos\theta)\hat{\boldsymbol{e}}_\phi \\ &\quad + (r\ddot{\theta} + 2\dot{r}\dot{\theta} + r\dot{\phi}^2\sin\theta\cos\theta)\hat{\boldsymbol{e}}_\theta \end{aligned} \tag{5.2}$$

Cylindrical Coordinates and $\{\hat{\boldsymbol{e}}_d, \hat{\boldsymbol{e}}_\phi, \hat{\boldsymbol{n}}_3\}$ vector components:

$$\begin{aligned} \boldsymbol{r} &= d\hat{\boldsymbol{e}}_d + z\hat{\boldsymbol{n}}_3 \\ \dot{\boldsymbol{r}} &= \dot{d}\hat{\boldsymbol{e}}_d + d\dot{\phi}\hat{\boldsymbol{e}}_\phi + \dot{z}\hat{\boldsymbol{n}}_3 \\ \ddot{\boldsymbol{r}} &= \ddot{d}\hat{\boldsymbol{e}}_d + (d\ddot{\phi} + 2\dot{d}\dot{\phi} - d\dot{\phi}^2)\hat{\boldsymbol{e}}_\phi + \ddot{z}\hat{\boldsymbol{n}}_3 \end{aligned} \tag{5.3}$$

From inspection of the geometry in Fig. 5.1, we can readily establish the corresponding family of six coordinate transformations:

Transformations to Cartesian Coordinates:

$$x(r,\phi,\theta) = r\cos\theta\cos\phi, \qquad x(d,\phi,z) = d\cos\phi \tag{5.4a}$$
$$y(r,\phi,\theta) = r\cos\theta\sin\phi, \qquad y(d,\phi,z) = d\sin\phi \tag{5.4b}$$
$$z(r,\phi,\theta) = r\sin\theta, \qquad z(d,\phi,z) = z \tag{5.4c}$$

Spherical Transformations:

$$r(x,y,z) = \sqrt{x^2+y^2+z^2}, \qquad r(d,\phi,z) = \sqrt{d^2+z^2} \tag{5.5a}$$
$$\phi(x,y,z) = \tan^{-1}(y/x), \qquad \phi(d,\phi,z) = \phi \tag{5.5b}$$
$$\theta(x,y,z) = \sin^{-1}(z/\sqrt{x^2+y^2+z^2}), \qquad \theta(d,\phi,z) = \sin^{-1}(z/\sqrt{d^2+z^2}) \tag{5.5c}$$

Cylindrical Transformations:

$$d(x,y,z) = \sqrt{x^2+y^2}, \qquad d(r,\phi,\theta) = r\cos\theta \tag{5.6a}$$
$$\phi(x,y,z) = \tan^{-1}(y/x), \qquad \phi(r,\phi,\theta) = \phi \tag{5.6b}$$
$$z(x,y,z) = z, \qquad z(r,\phi,\theta) = r\sin\theta \tag{5.6c}$$

Thus, even in this simple and most familiar example, we see that that an infinity of coordinate choices are possible. Depending upon the objectives being pursued in any given problem, any of these coordinate choices may be appropriate. It is clear that the details of most traditional analyses, such as formulating the differential equations of motion, are affected by the coordinates selected, since expressions for all kinematical and physical quantities depend on the coordinate choice. You can verify that the kinetic energy $T = m\dot{\boldsymbol{r}}\cdot\dot{\boldsymbol{r}}/2$, for the three above coordinate choices has the following three corresponding functional forms:

$$\begin{aligned} T(x,y,z,\dot{x},\dot{y},\dot{z}) &= m(\dot{x}^2+\dot{y}^2+\dot{z}^2)/2 \\ T(r,\phi,\theta,\dot{r},\dot{\phi},\dot{\theta}) &= m(\dot{r}^2+r^2\dot{\theta}^2+r^2\dot{\phi}^2\cos^2\theta)/2 \\ T(d,\phi,z,\dot{d},\dot{\phi},\dot{z}) &= m(\dot{d}^2+d^2\dot{\phi}^2+\dot{z}^2)/2 \end{aligned} \tag{5.7}$$

Lagrange, in thinking about the above and analogous issues, was apparently the first to ask the question: "Can one develop a *universal* form of the differential equations of motion, as a function of the system kinetic energy and unspecified *generalized* coordinates, i.e., $T(q_1,q_2,...,q_n,\dot{q}_1,\dot{q}_2,...,\dot{q}_n)$, that holds for all infinity of possible coordinate choices, and for particle motions, rigid body motions, translations, rotations, deformational vibrations... ?" The answer to this open ended, multi-faceted question is a qualified *yes*. The immortal developments that follow were introduced, mainly by Lagrange, in his quest to address these and related issues. In the process of re-tracing some of the work of Lagrange et al, we will find important branch points to concepts that go far beyond the scope of the above question. At the heart of these developments, it is evident that the various vector descriptions for position ($\boldsymbol{r}$), velocity ($\dot{\boldsymbol{r}}$), and acceleration ($\ddot{\boldsymbol{r}}$) ,

e.g., Eqs. (5.1) – (5.3), are alternative descriptions of mathematical representations of the same physical quantities, and of course, the corresponding forms of kinetic energy, e.g., Eq. (5.7), are likewise alternate mathematical representations for the same physical quantity. Thus the most important key to obtaining generalized (universal) forms for the equations of motion, for example, from variation of a generic function for kinetic energy $T(q_1, q_2, ..., q_n, \dot{q}_1, \dot{q}_2, ..., \dot{q}_n)$, is to consider from the onset a broad class of systems (for both "body models" and forces) and a brad class of admissible coordinate choices. While broad generality necessarily introduces a level of abstraction in the formulation, the ensuing analysis is of bearable complexity and well justified by the powerful *generalized* results obtained therefrom.

5.2 D'Alembert's Principle

Here we derive from Newton's second law an alternative formalism for developing equations of motion, this formalism will be seen to have an advantage that certain *virtually non-working* forces can be ignored. The most important role of D'Alembert's Principle, however, is that it is a stepping stone leading to Lagrange's Equations, Hamilton's Principle, and other variational principles in analytical dynamics.

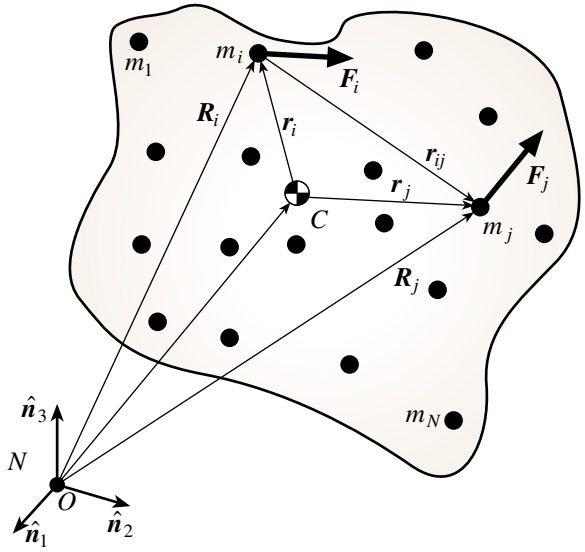


Figure 5.2: A System of N Particles

5.2.1 Virtual Displacements and Virtual Work

We consider a system of N particles, with the ith particle having mass m_i. With reference to Fig. 5.2, we locate m_i with an inertial position vector $\boldsymbol{R}_i$. We consider the total force vector acting on m_i to be segregated into two summed sub-sets of forces as

$$\boldsymbol{F}_i = \boldsymbol{f}_{c_i} + \boldsymbol{f}_i \tag{5.8}$$

where $\boldsymbol{f}_{c_i}$ is the vector sum of all *virtually non-working constraint forces* forces (as explained below) acting on m_i, and $\boldsymbol{f}_i = \boldsymbol{F}_i - \boldsymbol{f}_{c_i}$ is the vector sum of all other forces acting on m_i. We will see that the constraint forces ($\boldsymbol{f}_{c_i}$) can be eliminated from the analysis and this is an advantageous feature common to all of the methods of generalized mechanics. In order to accomplish the elimination of the constraint forces, we introduce the concept of *virtual displacement* $\delta\boldsymbol{R}_i$. A virtual displacement, in the most general context, is an instantaneous differential displacement *for the sake of analysis*. The virtual displacement $\delta(\cdot)$ of a dynamical motion variable $(\cdot)$ is closely related to the *first variation* of coordinates in variational calculus. We discuss subtle differences between virtual displacements and first variations in the developments of this chapter and especially in chapter 6. In dynamics problems where constraints are present, the most frequently used subset of virtual displacements $\delta\boldsymbol{R}_i$ are *consistent virtual displacements* which locate differentially displaced neighboring positions $\boldsymbol{R}_i + \delta\boldsymbol{R}_i$ for m_i that satisfy the constraint equations. In general, these virtual displacements are otherwise independently variable at each instant of time, and do not necessarily locate a family of points on a smooth neighboring trajectory (although this is an important special case). If the constraints acting on the system are smooth differential functions of $\boldsymbol{R}_i(t, q_1, q_2, ..., q_n)$, then admissible $\boldsymbol{R}_i$ are constrained to lie on a smooth *holonomic* (function of *position coordinates only*) constraint surface $\psi(t, q_1, q_2, ..., q_n)$, and we see that *admissible or consistent* virtual displacements $\delta\boldsymbol{R}_i$ locate points in a tangent plane, whose normal can be obtained by taking the gradient $\nabla\psi$ of the constraint surface. This idea is illustrated in general by Fig. 5.3. Note that the differential displacement $d\boldsymbol{R}_i = \dot{\boldsymbol{R}}_i(t)dt$ is tangent to a particular trajectory, whereas the consistent virtual displacement $\delta\boldsymbol{R}_i$ is an arbitrary differential displacement to *any* neighboring point in the tangent plane of feasible displacements. Thus the virtual displacements are not necessarily tangent to any solution trajectory, but they are required to locate neighboring differentially displaced points satisfying the constraints, at some arbitrary and unspecified time t in the motion. Ignoring friction, the constraint force $\boldsymbol{f}_{c_i}$ is always normal to the constraint surface (i.e., in the direction of $\nabla\psi$), and therefore can be written as

$$\boldsymbol{f}_{c_i} = \lambda\nabla\psi \tag{5.9}$$

where the scalar λ is a *Lagrange multiplier*. Friction and all forces (other than Eq. (5.9)) are accounted for in $\boldsymbol{f}_i$ of Eq. (5.8)

The *Virtual Work* δW is an abstract idea analogous to mechanical work, but associated with the instantaneous virtual displacements. The virtual work

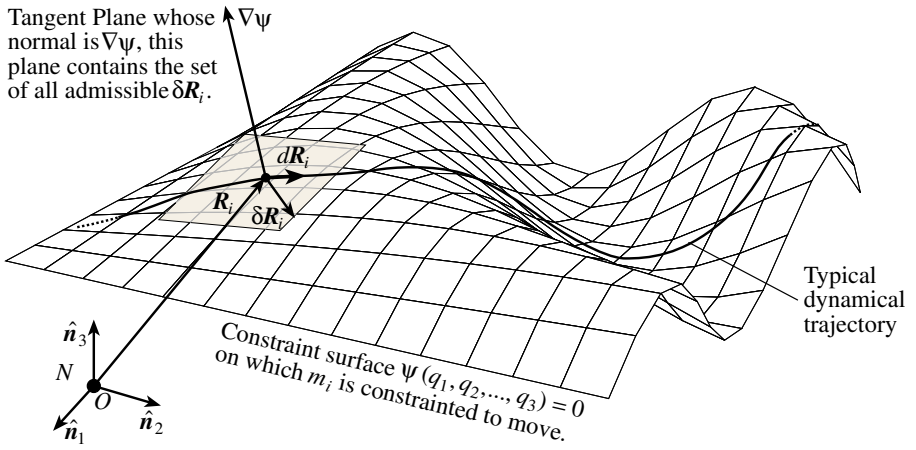


Figure 5.3: Particle Moving on a Holonomic Constraint Surface

done on m_i as a consequence of virtual displacement $\delta \boldsymbol{R}_i$ is defined as

$$\delta W_i \equiv \boldsymbol{F}_i \cdot \delta \boldsymbol{R}_i \tag{5.10}$$

Observe that the constraint force $\boldsymbol{f}_{c_i} = \lambda \nabla \psi$ is normal to the plane containing all infinity of admissible virtual displacements $\delta \boldsymbol{R}_i$, and this can be stated as the orthogonality condition:

$$\delta W_{c_i} = \boldsymbol{f}_{c_i} \cdot \delta \boldsymbol{R}_i = 0 \tag{5.11}$$

Thus the virtual work done by the normal constraint force associated with holonomic constraints is zero. Note, substituting Eq. (5.8), and making use of Eq. (5.9), we find that the virtual work on m_i reduces to

$$\delta W_i = \boldsymbol{f}_i \cdot \delta \boldsymbol{R}_i \tag{5.12}$$

We define the total virtual work to be the sum of the δW_i, so that

$$\delta W = \sum_{i=1}^{N} \boldsymbol{F}_i \cdot \delta \boldsymbol{R}_i = \sum_{i=1}^{N} \boldsymbol{f}_i \cdot \delta \boldsymbol{R}_i \tag{5.13}$$

5.2.2 Classical Developments of D'Alembert's Principle

From Newton's second law for the motion of m_i, we know $\boldsymbol{F}_i = m_i \ddot{\boldsymbol{R}}_i$, so using Eq. (5.8), we can write

$$\boldsymbol{f}_{c_i} + \boldsymbol{f}_i - m_i \ddot{\boldsymbol{R}}_i = 0, \qquad for\ i = 1, 2, ..., N \tag{5.14}$$

Upon taking the dot product of Eq. (5.14) with an arbitrary virtual displacement $\delta \boldsymbol{R}_i$ and summing over all N particles, we find the most general form of

D'Alembert's Principle to be

$$\delta W - \sum_{i=1}^{N} m_i \ddot{\boldsymbol{R}}_i \cdot \delta \boldsymbol{R}_i = 0 \tag{5.15}$$

We can put Eq. (5.15) in a more convenient form by recognizing that $\boldsymbol{R}_i = \boldsymbol{R}_i(t, q_1, q_2, ..., q_n)$, so that we can consider $\delta \boldsymbol{R}_i$ to be *generated* by a set of independent virtual variations in the $q_i s$ through

$$\delta \boldsymbol{R}_i = \sum_{j=1}^{n} \frac{\partial \boldsymbol{R}_i}{\partial q_j} \delta q_j \tag{5.16}$$

As a consequence, the virtual work can be written from Eq. (5.13) as

$$\delta W = \sum_{j=1}^{n} Q_j \delta q_j \tag{5.17}$$

where the n *generalized forces* Q_j are defined as a function of the N *virtually working forces* $\boldsymbol{f}_i$ as

$$Q_j \equiv \sum_{i=1}^{N} \boldsymbol{f}_i \cdot \frac{\partial \boldsymbol{R}_i}{\partial q_j} \tag{5.18}$$

Using Eqs. (5.16)–(5.18), D'Alembert's Principle of Eq. (5.15) is brought to the form

$$\sum_{j=1}^{n} \left[Q_j - \sum_{i=1}^{N} m_i \ddot{\boldsymbol{R}}_i \cdot \frac{\partial \boldsymbol{R}_i}{\partial q_j} \right] \delta q_j = 0 \tag{5.19}$$

Now, since the δq_j are independent virtual variations, they may be chosen independently and arbitrarily, so that the only conclusion possible from Eq. (5.19) is that each $[\cdot]$ term must independently vanish. This gives the most famous form of D'Alembert's Principle as

$$\sum_{i=1}^{N} m_i \ddot{\boldsymbol{R}}_i \cdot \frac{\partial \boldsymbol{R}_i}{\partial q_j} = Q_j \quad for\, j = 1, 2, ..., n \tag{5.20}$$

These equations are generally a coupled system of n second order differential equations, as will be illustrated by several examples below. First we consider a modification of Eq. (5.20) which facilitates derivation of the generalized forces and also makes connections with the notations of Kane, Moon, et al.

Observe that the position vector $\boldsymbol{R}_k = \boldsymbol{R}_k(t, q_1, q_2, \ldots, q_n)$ can be differentiated, using the chain rule, to obtain the expression for velocity

$$\boldsymbol{V}_i = \dot{\boldsymbol{R}}_i = \frac{\partial \boldsymbol{R}_i}{\partial t} + \sum_{k=1}^{N} \frac{\partial \boldsymbol{R}_i}{\partial q_k} \dot{q}_k, \quad i = 1, 2, ..., N \tag{5.21}$$

From this equation, we can immediately see the following important identity (known as the "cancellation of dots identity")

$$\frac{\partial \boldsymbol{V}_i}{\partial \dot{q}_k} = \frac{\partial \dot{\boldsymbol{R}}_i}{\partial \dot{q}_k} = \frac{\partial \boldsymbol{R}_i}{\partial q_k} \equiv \boldsymbol{v}_{ik} \tag{5.22}$$

Note that the time derivatives of the generalized coordinates ($\dot{q}_k$) *always appear linearly in the inertial velocities* $\boldsymbol{V}_i$. The quantities $\boldsymbol{v}_{ik}$ of Eq. (5.22) are simply the vector coefficients of $\dot{q}_k$, so Eq. (5.21) can be re-written as

$$\boldsymbol{V}_i = \dot{\boldsymbol{R}}_i = \frac{\partial \boldsymbol{R}_i}{\partial t} + \sum_{k=1}^{N} \dot{q}_k \boldsymbol{v}_{ik}, \quad i = 1, 2, ..., N \tag{5.23}$$

The vectors $\boldsymbol{v}_{ik}$ are obviously important kinematic quantities, they have been given various names such as "partial velocities" (Kane et al)[1] and "tangent vectors" (Lesser, Moon, et al)[?, 2] . We adopt Kane's partial velocity label, because partial velocity is descriptive of the definition $\boldsymbol{v}_{ik} \equiv \frac{\partial \dot{\boldsymbol{R}}_i}{\partial \dot{q}_k}$. Whatever we choose to call them, the n vectors $\{\boldsymbol{v}_{i1}, \boldsymbol{v}_{i2}, \ldots, \boldsymbol{v}_{in}\}$ form a vector basis for the inertial velocity $\boldsymbol{V}_i$ of the ith mass m_i, and for the case that time does not appear explicitly, the $\dot{q}_k$ are the coefficients that linearly combine the basis vectors $\boldsymbol{v}_{ik}$ to give the velocity vector $\boldsymbol{V}_i$. The general case is given by Eq. (5.23).

As a consequence of the truth that the inertial velocities must be formed en route to determining the inertial accelerations, we can simply record the vectors $\boldsymbol{v}_{ik}$ as they are generated in deriving the velocity-level kinematic description of the system. We can now re-write D'Alembert's Principle of Eq. (5.20) and the generalized force of Eq. (5.18) as

$$\sum_{i=1}^{N} m_i \ddot{\boldsymbol{R}}_i \cdot \boldsymbol{v}_{ij} = Q_j \quad \text{for } j = 1, 2, ..., n \tag{5.24}$$

and

$$Q_j \equiv \sum_{i=1}^{N} \boldsymbol{f}_i \cdot \boldsymbol{v}_{ij} \tag{5.25}$$

or, we can combine Eqs. (5.18) and (5.25) write D'Alembert's Principle in the form[2]

$$\sum_{i=1}^{N} [\boldsymbol{f}_i - m_i \dot{\boldsymbol{V}}_i] \cdot \boldsymbol{v}_{ij} = 0 \quad \text{for } j = 1, 2, ..., n \tag{5.26}$$

The above developments can be illustrated by the following example.

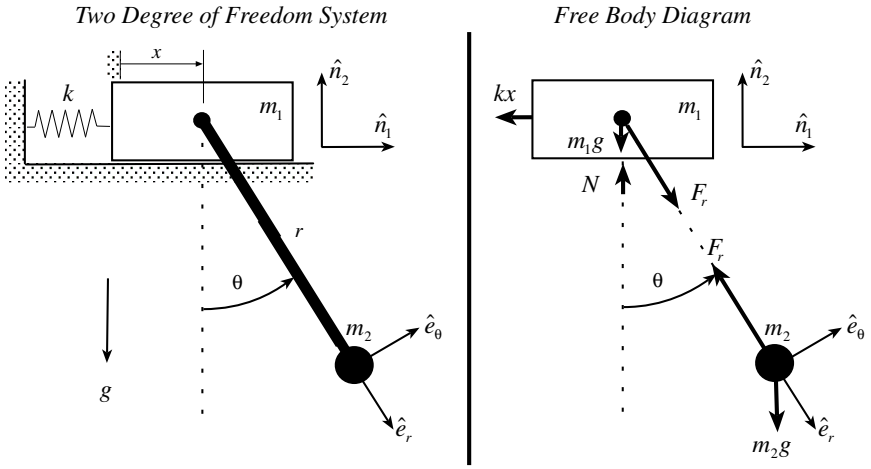


Figure 5.4: Classical Cart - Pendulum System

Example 5.1:
With reference to Fig. 5.4, we develop this system's equations of motion using (i) Newton's Laws and (ii) then D'Alembert's Principle. First we set down the kinematic equations as follows:

Kinematics of m_1:

$$\boldsymbol{R}_1 = x\hat{\boldsymbol{n}}_1, \quad \dot{\boldsymbol{R}}_1 = \boldsymbol{V}_1 = \dot{x}\hat{\boldsymbol{n}}_1, \quad \ddot{\boldsymbol{R}}_1 = \dot{\boldsymbol{V}}_1 = \ddot{x}\hat{\boldsymbol{n}}_1 \tag{5.27}$$

Kinematics of m_2:

$$\begin{aligned}
\boldsymbol{R}_2 &= x\hat{\boldsymbol{n}}_1 + r\hat{\boldsymbol{e}}_r = (x + r\sin\theta)\hat{\boldsymbol{n}}_1 + (-r\cos\theta)\hat{\boldsymbol{n}}_2 \\
\dot{\boldsymbol{R}}_2 &= \boldsymbol{V}_2 = \dot{x}\hat{\boldsymbol{n}}_1 + r\dot{\theta}\hat{\boldsymbol{e}}_\theta = (\dot{x} + r\dot{\theta}\cos\theta)\hat{\boldsymbol{n}}_1 + (r\dot{\theta}\sin\theta)\hat{\boldsymbol{n}}_2 \\
\ddot{\boldsymbol{R}}_2 &= \dot{\boldsymbol{V}}_2 = \ddot{x}\hat{\boldsymbol{n}}_1 - r\dot{\theta}^2\hat{\boldsymbol{e}}_r + r\ddot{\theta}\hat{\boldsymbol{e}}_\theta \\
&= (\ddot{x} - r\dot{\theta}^2\sin\theta + r\ddot{\theta}\cos\theta)\hat{\boldsymbol{n}}_1 + (r\dot{\theta}^2\cos\theta + r\ddot{\theta}\sin\theta)\hat{\boldsymbol{n}}_2 \\
&= (\ddot{x}\sin\theta - r\dot{\theta}^2)\hat{\boldsymbol{e}}_r + (\ddot{x}\cos\theta + r\ddot{\theta})\hat{\boldsymbol{e}}_\theta
\end{aligned} \tag{5.28}$$

Differential Equations Derived via Newton's Laws:
Making use of Newton's second law, we have the vector equations of motion

$$m_i\ddot{\boldsymbol{R}}_i = m_i\dot{\boldsymbol{V}}_i = \boldsymbol{F}_i \tag{5.29}$$

Referring to the free body diagram on the right hand side of Fig. 5.4, and making use of Eqs. (5.27) and (5.28) to obtain

$$\begin{aligned}
m_1\ddot{x} &= -kx + F_r\sin\theta \\
0 &= N - m_1 g - F_r\cos\theta
\end{aligned} \tag{5.30}$$

and for the m_2 equations, taking components of Eqs. (5.29) in the $\{\boldsymbol{e}_r, \boldsymbol{e}_\theta\}$ basis

$$\begin{aligned} m_2(\ddot{x}\sin\theta - r\dot{\theta}^2) &= -F_r + m_2 g\cos\theta \\ m_2(\ddot{x}\cos\theta + r\ddot{\theta}) &= -m_2 g\sin\theta \end{aligned} \tag{5.31}$$

Solving the first of Eqs. (5.31) for the constraint force (pendulum tension) F_r, we obtain

$$F_r = m_2 g\cos\theta - m_2(\ddot{x}\sin\theta - r\dot{\theta}^2) \tag{5.32}$$

Which, substituting into the first of Eqs. (5.30) and the second of Eqs. (5.31) eliminates the constraint force F_r and leads to the pair of differential equations that govern the system dynamics:

$$\begin{aligned} (m_1 + m_2\sin^2\theta)\ddot{x} - m_2 r\dot{\theta}^2\sin\theta &= -kx + m_2 g\sin\theta\cos\theta \\ (m_2\cos\theta)\ddot{x} + (m_2 r)\ddot{\theta} &= -m_2 g\sin\theta \end{aligned} \tag{5.33}$$

Differential Equations Derived via D'Alembert's Principle:
We will make use of Eqs. (5.12) - (5.20) to derive the equations of motion via a path that does not require us to first introduce the constraint forces (N, F_r), then eliminate them. For the initial developments we will also not make use of the partial velocity ideas, but rather directly differentiate the position vectors to obtain the terms needed in the classical D'Alembert's Principle equations of motion. We see that the gradient of the inertial position vectors with respect to (x, θ) is needed, the needed partial derivatives can be obtained directly from $\boldsymbol{R}_i(x, \theta)$ as

$$\begin{aligned} \frac{\partial \boldsymbol{R}_1}{\partial x} &= \hat{\boldsymbol{n}}_1 \qquad & \frac{\partial \boldsymbol{R}_1}{\partial \theta} &= 0 \\ \frac{\partial \boldsymbol{R}_2}{\partial x} &= \hat{\boldsymbol{n}}_1 \qquad & \frac{\partial \boldsymbol{R}_2}{\partial \theta} &= r\frac{\partial \hat{\boldsymbol{e}}_r}{\partial \theta} = r\hat{\boldsymbol{e}}_\theta = r(\cos\theta\hat{\boldsymbol{n}}_1 + \sin\theta\hat{\boldsymbol{n}}_2) \end{aligned} \tag{5.34}$$

Using these, we obtain the generalized forces from Eqs. (5.18) as

$$\begin{aligned} Q_x &= \boldsymbol{F}_1\cdot\frac{\partial \boldsymbol{R}_1}{\partial x} + \boldsymbol{F}_2\cdot\frac{\partial \boldsymbol{R}_2}{\partial x} \\ &= [(-kx)\hat{\boldsymbol{n}}_1]\cdot[\hat{\boldsymbol{n}}_1] + [-mg\boldsymbol{n}_2]\cdot[\hat{\boldsymbol{n}}_1] \\ &= -kx \\ Q_\theta &= \boldsymbol{F}_1\cdot\frac{\partial \boldsymbol{R}_1}{\partial \theta} + \boldsymbol{F}_2\cdot\frac{\partial \boldsymbol{R}_2}{\partial \theta} \\ &= [(-kx)\hat{\boldsymbol{n}}_1]\cdot[0] + [mg\cos\theta\hat{\boldsymbol{e}}_r - mg\sin\theta\hat{\boldsymbol{e}}_\theta]\cdot[r\hat{\boldsymbol{e}}_\theta] \\ &= -mgr\sin\theta \end{aligned} \tag{5.35}$$

We are now prepared to develop the differential equations of motion using D'Alembert's Principle in the form of Eqs. (5.20) as follows:

$$\begin{aligned} m_1\ddot{\boldsymbol{R}}_1\cdot\frac{\partial \boldsymbol{R}_1}{\partial x} + m_2\ddot{\boldsymbol{R}}_2\cdot\frac{\partial \boldsymbol{R}_2}{\partial x} &= Q_x \\ m_1\ddot{\boldsymbol{R}}_1\cdot\frac{\partial \boldsymbol{R}_1}{\partial \theta} + m_2\ddot{\boldsymbol{R}}_2\cdot\frac{\partial \boldsymbol{R}_2}{\partial \theta} &= Q_\theta \end{aligned} \tag{5.36}$$

Substitution of Eqs. (5.35) and (5.34) leads to the system of differential equations:

$$\begin{aligned}(m_1+m_2)\ddot{x}+(m_2r\cos\theta)\ddot{\theta}-m_2r\dot{\theta}^2\sin\theta&=-kx\\(m_2r\cos\theta)\ddot{x}+(m_2r^2)\ddot{\theta}&=-mgr\sin\theta\end{aligned}\tag{5.37}$$

The above developments could be accelerated modestly by making use of the so called virtual power[2] form of D'Alembert's Principle (Eqs. (5.26)), and by collecting the partial velocities $\boldsymbol{v}_{ij}$ from the velocity level kinematics. Adopting the notation $q_1=x, q_2=\theta$, then Eqs. (5.26) specializes to

$$\begin{aligned}[\boldsymbol{f}_1-m_1\dot{\boldsymbol{V}}_1]\cdot\boldsymbol{v}_{11}+[\boldsymbol{f}_2-m_2\dot{\boldsymbol{V}}_1]\cdot\boldsymbol{v}_{21}&=0\\[\boldsymbol{f}_1-m_1\dot{\boldsymbol{V}}_2]\cdot\boldsymbol{v}_{12}+[\boldsymbol{f}_2-m_2\dot{\boldsymbol{V}}_1]\cdot\boldsymbol{v}_{22}&=0\end{aligned}\tag{5.38}$$

where from the velocity-level kinematics we see

$$\begin{aligned}\boldsymbol{V}_1=\dot{\boldsymbol{R}}_1=\dot{x}\hat{\boldsymbol{n}}_1,\quad&\rightarrow\quad\boldsymbol{v}_{11}=\hat{\boldsymbol{n}}_1,\,\boldsymbol{v}_{12}=0\\\boldsymbol{V}_2=\dot{\boldsymbol{R}}_2=\dot{x}\hat{\boldsymbol{n}}_1+r\dot{\theta}\hat{\boldsymbol{e}}_\theta,\quad&\rightarrow\quad\boldsymbol{v}_{21}=\hat{\boldsymbol{n}}_1,\,\boldsymbol{v}_{22}=r\dot{\theta}\hat{\boldsymbol{e}}_\theta\end{aligned}\tag{5.39}$$

Direct substitution of Eqs. (5.39) into Eqs. (5.38), along with $\boldsymbol{f}_1=-kx\hat{\boldsymbol{n}}_1$ and $\boldsymbol{f}_2=-mg\hat{\boldsymbol{n}}_2$, immediately verifies Eqs. (5.37). For many degree of freedom systems, the systematic notation of the virtual power formulation offers some advantages. The most important advantage is to recognize that one does not need to return to the position vector to take the position partial derivatives in the classical version (5.20), these can be simply replaced by the partial velocities, by virtue of Eqs. (5.22) which are already available as a consequence of having derived the velocity expressions in the form of Eqs. (5.23) en route to the also required acceleration vectors.

Discussion:

Comparing Eqs. (5.37) to those (Eqs. (5.33)) obtained from Newton's second law, we see that these equations have different forms. The more elegant form of Eqs. (5.37) is preferred due to the symmetry of the acceleration coefficients (the elements of the "mass matrix"). Both sets of equations are correct and it is easy to re-arrange Eqs. (5.33) via linear combinations of the two equations to obtain the form of Eqs. (5.37); this development is left as an exercise. An implicit question arises: How can we guarantee that we obtain the symmetric form of Eqs. (5.37) directly from Newton's laws? The answer can be verified, for this special case, by re-working the Newton's law developments and insisting that all acceleration and force vectors be projected such that all acceleration and force vector components are taken in a common reference frame [i.e., use either $(\hat{\boldsymbol{e}}_r,\hat{\boldsymbol{e}}_\theta)$ or $(\hat{\boldsymbol{n}}_1,\hat{\boldsymbol{n}}_2)$ unit vectors exclusively, rather than the mixed pattern used to obtain Eqs. (5.33)] before writing the two sets of component equations from Newton's second law. More insight on this issue can be obtained from the subsequent developments of this chapter.

5.2.3 Holonomic Constraints

The above developments implicitly assume that the generalized coordinates are independent. It often occurs that the coordinates are not independent. In the simplest case, the redundancy of the coordinates arise because of constraining algebraic relationships of the form

$$\psi_k(t, q_1, q_2, \dots, q_n) = 0 \qquad k = 1, 2, \dots, m \tag{5.40}$$

We note that any velocity-dependent constraint that cannot be integrated to obtain the above form would not qualify as *holonomic*, and obviously, all inequality constraints must be considered non-holonomic. If time does not explicitly appear, then this special case of holonomic constraints are said to be *rheonomic*. For most cases, we restrict attention to the case that the $\psi_k(t, q_1, q_2, \dots, q_n)$ are continuous and differentiable with respect to all arguments.

Consider briefly the special case of $m = 1$, then $\psi(q_1, q_2, \dots, q_n) = 0$ constitutes a *constraint surface* on which the admissible trajectories lie, and as a consequence, the coordinates $\{q_1, q_2, \dots, q_n\}$ are not independent. There are two obvious approaches to dealing with the constraint: (i) Solve the constraint equation for any one of the coordinates as a function of the other $n-1$ *q's* which may now be considered independent, or (ii) Replace the constraint surface by an equivalent *constraint force* that effectively causes the motion to remain on the constraint surface. These two approaches are illustrated by the following example.

Example 5.2: Let us study the simple pendulum shown in Fig. 5.5. Consider

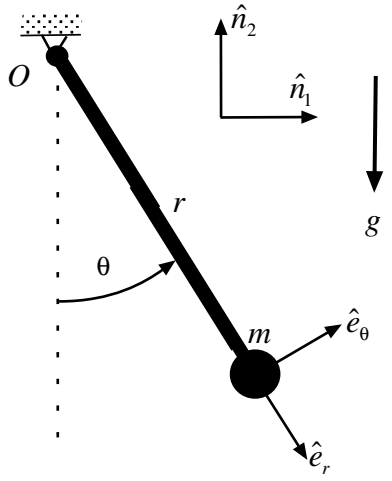


Figure 5.5: Classical Pendulum

the redundant coordinates (r, θ). The position, velocity, and acceleration vectors are given by

$$\boldsymbol{R} = r\hat{\boldsymbol{e}}_r, \quad \dot{\boldsymbol{R}} = \dot{r}\hat{\boldsymbol{e}}_r + r\dot{\theta}\hat{\boldsymbol{e}}_\theta, \quad \ddot{\boldsymbol{R}} = (\ddot{r} - r\dot{\theta}^2)\hat{\boldsymbol{e}}_r + (r\ddot{\theta} + 2\dot{r}\dot{\theta})\hat{\boldsymbol{e}}_\theta \tag{5.41}$$

We consider two approaches for imposing the holonomic (and rheonomic) constraint $r = R = constant$, or

$$\psi(r, \theta) = r - R = 0 \tag{5.42}$$

Algebraic Constraint Elimination

In this approach, the constraint equation of Eq. (5.42) is trivially solvable for $r = R$, and the derived constraint conditions that $\dot{r} = 0, \ddot{r} = 0$ are imposed on the kinematics equations of Eq. (5.41) to obtain

$$\boldsymbol{R} = R\hat{\boldsymbol{e}}_r, \qquad \dot{\boldsymbol{R}} = R\dot{\theta}\hat{\boldsymbol{e}}_\theta, \qquad \ddot{\boldsymbol{R}} = (R\dot{\theta}^2)\hat{\boldsymbol{e}}_r + (R\ddot{\theta})\hat{\boldsymbol{e}}_\theta \tag{5.43}$$

leaving only θ as an independent coordinate. We can now apply D'Alembert's Principle of Eq. (5.20) to generate the differential equations as follows

$$\begin{aligned} m\ddot{\boldsymbol{R}} \cdot \frac{\partial \boldsymbol{R}}{\partial \theta} &= Q_\theta \\ mR^2\ddot{\theta} = -mgR\sin\theta \quad &\rightarrow \quad \ddot{\theta} = -\frac{g}{R}\sin\theta \end{aligned} \tag{5.44}$$

Constraint Force via Lagrange Multipliers

In this approach, we observe that the pendulum is physically constrained to move on a circular constraint surface and the associated force must be normal to this surface. While the direction is known, the magnitude is not. Thus the constraint force associated with $\psi(r, \theta) = r - R = 0$ is written as

$$F_c = \lambda \nabla \psi = \lambda \hat{\boldsymbol{e}}_r \tag{5.45}$$

where the unknown scalar λ is a *Lagrange Multiplier*.

The total force acting on mass m is $-mg\hat{\boldsymbol{n}}_2 + \lambda\hat{\boldsymbol{e}}_r$ so that D'Alembert's Principle of Eq. (5.20), considering both r and θ as generalized coordinates is

$$\begin{aligned} m\ddot{\boldsymbol{R}} \cdot \frac{\partial \boldsymbol{R}}{\partial r} &= Q_r \\ m\ddot{\boldsymbol{R}} \cdot \frac{\partial \boldsymbol{R}}{\partial \theta} &= Q_\theta \end{aligned} \tag{5.46}$$

which gives

$$\begin{aligned} m[(\ddot{r} - r\dot{\theta}^2)\hat{\boldsymbol{e}}_r + (r\ddot{\theta} + 2\dot{r}\dot{\theta})\hat{\boldsymbol{e}}_\theta] \cdot (\hat{\boldsymbol{e}}_r) &= [-mg\hat{\boldsymbol{n}}_2 + \lambda\hat{\boldsymbol{e}}_r] \cdot (\hat{\boldsymbol{e}}_r) = mg\cos\theta + \lambda \\ m[(\ddot{r} - r\dot{\theta}^2)\hat{\boldsymbol{e}}_r + (r\ddot{\theta} + 2\dot{r}\dot{\theta})\hat{\boldsymbol{e}}_\theta] \cdot (r\hat{\boldsymbol{e}}_\theta) &= [-mg\hat{\boldsymbol{n}}_2 + \lambda\hat{\boldsymbol{e}}_r] \cdot (r\hat{\boldsymbol{e}}_\theta) = -mgr\sin\theta \end{aligned} \tag{5.47}$$

Imposing $r = constant = R$ and carrying out the implied algebra, these simplify to the final result

$$\begin{aligned} \lambda &= -m(g\cos\theta + R\dot{\theta}^2) \\ \ddot{\theta} &= -\frac{g}{R}\sin\theta \end{aligned} \tag{5.48}$$

So, we see in this example that D'Alembert's Principle (when used with a redundant coordinate description of the motion, and imposing the holonomic constraint conditions at the end) generates the constrained equations

of motion and determines the Lagrange multiplier (which in this case has the interpretation of the negative of the constraint force F_r) as a function of the coordinates and their derivatives.

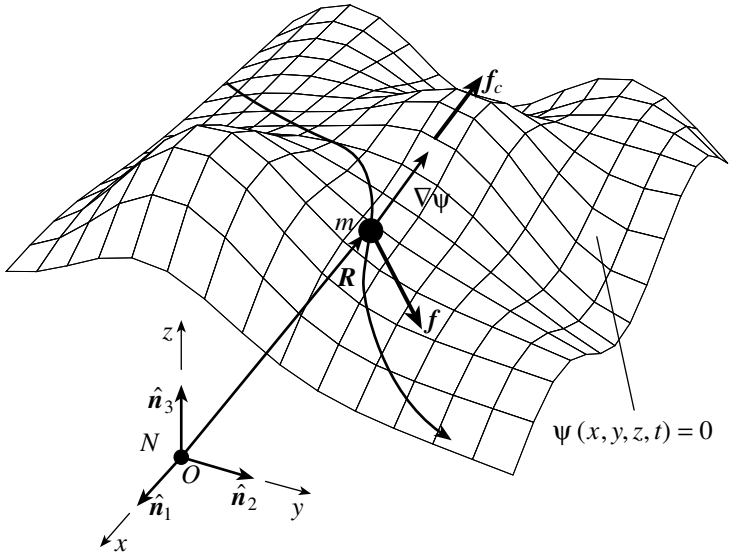


Figure 5.6: Particle Moving on a Holonomic Constraint Surface

The above developments can be viewed from several perspectives and generalized for the case of m constraints and n generalized coordinates. Perhaps it would be instructive to first consider the case of three rectangular coordinates (x, y, z) and one constraint, with Newton's second law used to develop the equations of motion. Consider Figure 5.6, Newton's second law provides the equation of motion

$$\boldsymbol{f} + \boldsymbol{f}_c = m\ddot{\boldsymbol{R}} \tag{5.49}$$

where $\boldsymbol{f}_c$ is the constraint force normal to the smooth holonomic constraint surface

$$\psi(x, y, z, t) = 0 \tag{5.50}$$

and $\boldsymbol{f}$ is the vector sum of all other forces not normal to the constraint surface. Since $\psi(x, y, z, t) = 0$ is assumed differentiable, then from Eq. (5.50), we have the derived equation

$$\frac{d\psi}{dt} = \frac{\partial\psi}{\partial x}\dot{x} + \frac{\partial\psi}{\partial y}\dot{y} + \frac{\partial\psi}{\partial z}\dot{z} + \frac{\partial\psi}{\partial t} = 0 \tag{5.51}$$

The above condition should be viewed as the time derivative of the constraint at any/all points along the path $(x(t), y(t), z(t), \dot{x}(t), \dot{y}(t), \dot{z}(t))$. Alternatively, we can consider the differential of ψ along the path, which also must vanish.

$$d\psi = \frac{\partial \psi}{\partial x}dx + \frac{\partial \psi}{\partial y}dy + \frac{\partial \psi}{\partial z}dz + \frac{\partial \psi}{\partial t}dt = 0 \tag{5.52}$$

Conceptually, notice that the differential change $d\psi$ along the path is different from the virtual change $\delta\psi$:

$$\delta\psi = \frac{\partial \psi}{\partial x}\delta x + \frac{\partial \psi}{\partial y}\delta y + \frac{\partial \psi}{\partial z}\delta z = 0 \tag{5.53}$$

Note in Eq. (5.53), $(\delta x, \delta y, \delta z)$ are arbitrary admissible virtual displacements that locate all infinity of points lying in the local tangent plane whose normal is $\nabla\psi$ (see Figure 5.3), whereas (dx, dy, dz) are the particular differential displacements along the path from $\{x(t), y(t), z(t)\}$ to $\{x(t+dt), y(t+dt), z(t+dt)\}$. From another perspective, Eq. (5.53) can be viewed as the condition that admissible virtual displacements must satisfy. If t is not explicitly contained in ψ, then $\{dx, dy, dz\}$ are obviously a special case of $\{\delta x, \delta y, \delta z\}$. Following the same argument leading to Eq. (5.9), we know $\boldsymbol{f}_c$ is proportional to $\nabla\psi$

$$\boldsymbol{f}_c = \lambda\nabla\psi = \lambda\frac{\partial \psi}{\partial x}\hat{\boldsymbol{n}}_1 + \lambda\frac{\partial \psi}{\partial y}\hat{\boldsymbol{n}}_2 + \lambda\frac{\partial \psi}{\partial z}\hat{\boldsymbol{n}}_3 \tag{5.54}$$

and thus the equations of motion of Eq. (5.14) become

$$\begin{aligned} m\ddot{x} &= f_x + \lambda\frac{\partial \psi}{\partial x} \\ m\ddot{y} &= f_y + \lambda\frac{\partial \psi}{\partial y} \\ m\ddot{z} &= f_z + \lambda\frac{\partial \psi}{\partial z} \end{aligned} \tag{5.55}$$

$$\psi(x, y, z, t) = 0 \tag{5.56}$$

We note that Eqs. (5.55) and (5.56) provide three differential equations and one algebraic equation — four equations involving four unknowns $x(t)$, $y(t)$, $z(t)$ and $\lambda(t)$.

Example 5.3: We return to the simple pendulum of Figure 5.5 and consider the alternative choice of rectangular coordinates (x, y). In lieu of the polar coordinate representation of kinematics in Eq. (5.43), we have

$$\begin{aligned} \boldsymbol{R} &= x\hat{\boldsymbol{n}}_1 + y\hat{\boldsymbol{n}}_2 \\ \dot{\boldsymbol{R}} &= \dot{x}\hat{\boldsymbol{n}}_1 + \dot{y}\hat{\boldsymbol{n}}_2 \\ \ddot{\boldsymbol{R}} &= \ddot{x}\hat{\boldsymbol{n}}_1 + \ddot{y}\hat{\boldsymbol{n}}_2 \end{aligned} \tag{5.57}$$

From Eq. (5.55) we have three equations with three unknowns $x(t)$, $y(t)$ and $\lambda(t)$ as

$$\begin{aligned} m\ddot{x} &= \lambda(2x) \\ m\ddot{y} &= -my + \lambda(2y) \end{aligned} \tag{5.58}$$

$$x^2 + y^2 - L^2 = 0 \tag{5.59}$$

We can eliminate λ by taking two time derivatives of Eq. (5.59) as

$$2x\ddot{x} + 2y\ddot{y} + 2\dot{x}^2 + 2\dot{y}^2 = 0 \tag{5.60}$$

and solving for $(\ddot{x}, \ddot{y})$ as a function of λ from Eq. (5.58) as

$$\begin{aligned} \ddot{x} &= \lambda\left(\frac{2x}{m}\right) \\ \ddot{y} &= -g + \lambda\left(\frac{2y}{m}\right) \end{aligned} \tag{5.61}$$

Substitution of $(\ddot{x}, \ddot{y})$ into Eq. (5.60) and making use of $r^2 = x^2 + y^2$, we have

$$\lambda = \frac{m}{2r^2}\left[yg - (\dot{x}^2 + \dot{y}^2)\right] \tag{5.62}$$

And finally, substitution of Eq. (5.62) into Eq. (5.58), we have

$$\begin{aligned} m\ddot{x} &= \quad + \frac{mx}{r^2}\left[yg - (\dot{x}^2 + \dot{y}^2)\right] \\ m\ddot{y} &= -g + \frac{my}{r^2}\left[yg + (\dot{x}^2 + \dot{y}^2)\right] \end{aligned} \tag{5.63}$$

Notice either of the equations in Eq. (5.63) could be solved, together with either $x = \pm\sqrt{r^2 - y^2}$ or $y = \pm\sqrt{r^2 - x^2}$. These equations are "sufficiently ugly" that we note the un-surprising truth, comparing Eqs. (5.62) and (5.63) with the familiar/elegant Eqs. (5.48), a judicious coordinate choice is frequently of vital importance! In this case (r, θ) is vastly superior to (x, y), because $r = R =$ constant and $\theta(t)$ directly describes all feasible motions consistent with the constraint. We note — for small motions near $(r, \theta) = (R, 0)$ and $(x, y) = (0, -R)$ both reduce to essentially the same linear system

$$\ddot{\theta} = -\frac{g}{R}\theta \quad \text{and/or} \quad \ddot{x} = -\frac{g}{R}x \tag{5.64}$$

and thus we note that coordinate selection is often more forgiving for small (linear) motions than for large (nonlinear) motions.

We now consider the case that two constraints ψ_i exist:

$$\psi_k(x, y, z, t) = 0 \qquad \text{for } k = 1, 2 \tag{5.65}$$

The constraint force $\boldsymbol{f}_c$ in Eq. (5.54) is the vector sum of the two constraint forces as

$$\boldsymbol{f}_c = \lambda_1 \nabla\psi_1 + \lambda_2 \nabla\psi_2 \tag{5.66}$$

and the Newtonian equations of motion become

$$\begin{aligned} m\ddot{x} &= f_x + \lambda_1 \frac{\partial \psi_1}{\partial x} + \lambda_2 \frac{\partial \psi_2}{\partial x} \\ m\ddot{y} &= f_y + \lambda_1 \frac{\partial \psi_1}{\partial y} + \lambda_2 \frac{\partial \psi_2}{\partial y} \\ m\ddot{z} &= f_z + \lambda_1 \frac{\partial \psi_1}{\partial z} + \lambda_2 \frac{\partial \psi_2}{\partial z} \end{aligned} \tag{5.67}$$

and we have the two algebraic equations of Eq. (5.65) providing five equations in terms of the five unknowns

$$\{x(t), y(t), z(t), \lambda_1(t), \lambda_2(t)\} \tag{5.68}$$

Example 5.4: We note that the above developments all hold for a certain class of holonomic constraints $\psi(x, y, z, t) = 0$, for which

$$\dot{\psi} = 0 = \frac{\partial \psi}{\partial t} + \frac{\partial \psi}{\partial x}\dot{x} + \frac{\partial \psi}{\partial y}\dot{y} + \frac{\partial \psi}{\partial z}\dot{z} \tag{5.69}$$

We will see subsequently that the above developments can be generalized to consider a class of non-holonomic constraints

$$B(x, y, z) + A_1(x, y, z)\dot{x} + A_2(x, y, z)\dot{y} + A_3(x, y, z)\dot{z} = 0 \tag{5.70}$$

Such constraints, which depend on velocity linearly, are known as *Pfaffian* constraints. Notice, if a function $\psi(x, y, z, t)$ exists such that

$$\begin{aligned} B(x, y, z) &= \frac{\partial \psi}{\partial t} \\ A_1(x, y, z) &= \frac{\partial \psi}{\partial x}, \quad A_2(x, y, z) = \frac{\partial \psi}{\partial y}, \quad A_3(x, y, z) = \frac{\partial \psi}{\partial z} \end{aligned} \tag{5.71}$$

then Eqs. (5.70) are said to be *integrable* to the holonomic constraint $\psi(x, y, z, t) = 0$. Performing *partial integration*, it is easy to test directly to see if a Pfaffian form non-holonomic constraint is integrable to a corresponding holonomic constraint. Notice the following example. Suppose a spherical pendulum shown in Figure 5.7 is acted upon by a motor torque which exactly enforces the motion constraint:

$$\dot{\theta} = \frac{t(\cos\phi - \dot{\phi})}{2 + \sin\phi} \tag{5.72}$$

or, we have the Pfaffian form

$$(-\cos\phi)t + (2 + \sin\phi)\dot{\theta} + (t)\dot{\phi} = 0 \tag{5.73}$$

Comparing Eqs. (5.70) and (5.73), we have

$$\begin{aligned} B(\phi, \theta) &= -(\cos\phi)t \\ A_1(\phi, \theta) &= 2 + \sin\phi \\ A_2(\phi, \theta) &= t \end{aligned} \tag{5.74}$$

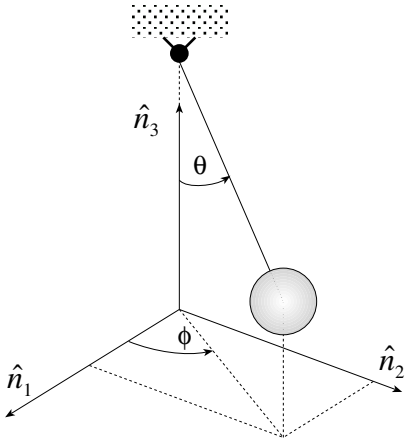


Figure 5.7: Illustration of Motor Driven Spherical Pendulum

To see if the non-holonomic constrain of Eq. (5.73) is integrable, we conjecture the existence of $\psi(\phi,\theta,t)$ such that

$$\frac{\partial\psi}{\partial t} = B = -(\cos\phi)t \quad \rightarrow \quad \psi = -\frac{1}{2}(\cos\phi)t^2 + f_1(\phi,\theta) \tag{5.75}$$

$$\frac{\partial\psi}{\partial\theta} = A_1 = 2+\sin\phi \quad \rightarrow \quad \psi = (2+\sin\phi)\theta + f_2(\phi,t) \tag{5.76}$$

$$\frac{\partial\psi}{\partial\phi} = A_2 = t \quad \rightarrow \quad \psi = \phi t + f_3(\theta,t) \tag{5.77}$$

If ψ exists, there must be a valid choice for the *functions of partial integration* $f_1(\phi,\theta)$, $f_2(\phi,t)$ and $f_3(\theta,t)$ such that the same function $\psi(\phi,\theta,t)$ is obtained from Eqs. (5.75), (5.76) and (5.77). In this case we conjecture observing Eqs. (5.76) and (5.77) that

$$f_2(\phi,t) = \phi t \tag{5.78}$$

$$f_3(\theta,t) = (2\sin\phi)\theta \quad \rightarrow \quad \text{inconsistent, depends on } \phi! \tag{5.79}$$

Now, comparing Eqs. (5.75) and (5.76), we conjecture

$$f_1(\phi,\theta) = (2+\sin\phi)\theta \tag{5.80}$$

$$f_2(\phi,t) = -\frac{1}{2}(\cos\phi)t^2 \quad \rightarrow \quad \text{inconsistent with Eq. (5.78)} \tag{5.81}$$

Finally, comparing Eqs. (5.75) and (5.77), we conjecture

$$f_1(\phi,\theta) = \phi t \quad \rightarrow \quad \text{inconsitent with Eq. (5.80)} \tag{5.82}$$

$$f_3(\theta,t) = -\frac{1}{2}(\cos\phi)t^2 \quad \rightarrow \quad \text{inconsistent, depends on } \phi \tag{5.83}$$

Any of the four inconsistency results obtained for the functions of partial integration are sufficient that $\psi(\phi,\theta,t)$ does not exist. Therefore the Pfaffian form of Eq. (5.73) is a non-integrable non-holonomic constraint.

5.2.4 Newtonian Constrained Dynamics of N Particles

The above developments are now generalized to N particles subject to m constraints. As the most fundamental choice of coordinates, we could choose the inertial Cartesian set of coordinates

$$\{\boldsymbol{q}\} \equiv \{q_1, q_2, q_3, \ldots, q_n\} \equiv \{x_1, y_1, z_1; x_2, y_2, z_2; \ldots; x_n, y_n, z_n\} \tag{5.84}$$

where $n = 3N$. More generally, $\{\boldsymbol{q}\}$ is any set of $n = 3N$ generalized coordinates.

Consider a set of m Pfaffian non-holonomic constraints of the form

$$\sum_{j=1}^{n} A_{kj}\dot{q}_j + B_k = 0 \qquad k = 1, 2, \ldots, m \tag{5.85}$$

or, in differential form

$$\sum_{j=1}^{n} A_{kj}dq_j + B_k dt = 0 \qquad k = 1, 2, \ldots, m \tag{5.86}$$

where $A_{kj} = A_{kj}(q_1, \ldots, q_n, t)$ and $B_k = B_k(q_1, \ldots, q_n, t)$. Some, and occasionally all, of the constraints may be integrable to obtain holonomic constraints of the form

$$\psi_j(q_1, \ldots, q_n, t) = 0 \qquad j = 1, 2, \ldots, n \tag{5.87}$$

Conversely, given smooth holonomic constraints of the form of Eq. (5.87) can always be differentiated to obtain Eqs. (5.85) and (5.86). The reverse is true only for integrable constraints (see example 5.4). The equations of motion for the N particles, for the case that $\{\boldsymbol{q}\}$ is the set of $n = 3N$ Cartesian inertial coordinates is

$$M_j\ddot{q}_j = f_j + f_{c_j} = f_j + \sum_{k=1}^{m} A_{kj}\lambda_k \qquad j = 1, 2, \ldots, n \tag{5.88}$$

where

$$\{M_1, M_2, M_3; M_4, M_5, M_6; \ldots; M_{n-2}, M_{n-1}, M_n\} \equiv \{m_1, m_1, m_1; m_2, m_2, m_2; \ldots; m_N, m_N, m_N\}$$

and

$$\sum_{j=1}^{n} A_{kj}\dot{q}_j + B_k = 0 \qquad k = 1, 2, \ldots, m \tag{5.89}$$

Eqs. (5.85) and (5.88) provide $n + m$ equations in the $n + m$ unknowns

$$\{q_1, q_2, \ldots, q_n; \lambda_1, \lambda_2, \ldots, \lambda_m\}$$

The Eqs. (5.88) and (5.89) constitute a set of *Differential-Algebraic Equations* (DAEs). Prior to developing the analogous constrained dynamics results for the generalized methods that follow from D'Alembert's Principle, we digress to consider a version of the Lagrange Multiplier Rule for parameter optimization.

5.2.5 Lagrange Multiplier Rule for Constrained Optimization

Consider the problem of extremizing (maximizing or minimizing) the smooth, twice differentiable function

$$\phi(q_1, q_2, \ldots, q_n) \tag{5.90}$$

subject to satisfying m equality constraints of the form

$$\psi_j(q_1, q_2, \ldots, q_n) = 0 \qquad j = 1, 2, \ldots, m < n \tag{5.91}$$

where $\psi(q_1, q_2, \ldots, q_n)$ is continuous and at least once differentiable.

An important result, due to Lagrange, is that the necessary conditions for extremizing Eq. (5.90) subject to Eq. (5.91) are identical to the necessary conditions for extremizing the *augmented function*

$$\Phi(q_1, \ldots, q_n; \lambda_1, \ldots, \lambda_m) \equiv \phi(q_1, \ldots, q_n) + \sum_{j=1}^{m} \lambda_j \psi_j(q_1, \ldots, q_n) \tag{5.92}$$

where $\{\lambda_1, \ldots, \lambda_m\}$ are Lagrange multipliers, and, as developed below, the necessary conditions for extremizing Φ is

$$\left.\frac{\partial \Phi}{\partial q_j}\right|_{\hat{\boldsymbol{q}}, \hat{\boldsymbol{\lambda}}} = 0 \qquad j = 1, 2, \ldots, n \tag{5.93}$$

$$\left.\frac{\partial \Phi}{\partial \lambda_j}\right|_{\hat{\boldsymbol{q}}, \hat{\boldsymbol{\lambda}}} \equiv \psi_j(\hat{q}_1, \ldots, \hat{q}_n) = 0 \qquad j = 1, 2, \ldots, m \tag{5.94}$$

Eqs. (5.93) and (5.94) provide $n + m$ algebraic equations to be solved for the "stationary points"

$$\hat{\boldsymbol{q}} \equiv (\hat{q}_1, \ldots, \hat{q}_n)^T \qquad \hat{\boldsymbol{\lambda}} \equiv \left(\hat{\lambda}_1, \ldots, \hat{\lambda}_m\right)^T \tag{5.95}$$

The existence of stationary points $(\hat{\boldsymbol{q}}, \hat{\boldsymbol{\lambda}})$ that satisfy Eqs. (5.93) and (5.94) is not guaranteed; there may be one solution, no solution, or multiple solutions. Furthermore, additional analysis is required to discern whether the point $(\hat{\boldsymbol{q}}, \hat{\boldsymbol{\lambda}})$ represents a local minimum, maximum, or generalized inflection point.

Implicit in the concept (and proof) of the Lagrange multiplier rule is the idea of locally "constrained differential variations." For example, suppose we wish to minimize a function of three variables

$$\phi(x, y, z) \tag{5.96}$$

subject to one equality constraint

$$\psi(x, y, z) = 0 \tag{5.97}$$

For arbitrary virtual displacements $(\delta x, \delta y, \delta z)$, the virtual differential change in ϕ is

$$\delta\phi = \frac{\partial\phi}{\partial x}\delta x + \frac{\partial\phi}{\partial y}\delta y + \frac{\partial\phi}{\partial z}\delta z \tag{5.98}$$

Similarly, at an arbitrary admissible point (x, y, z) that satisfies Eq. (5.97), the virtual change of ψ is

$$\delta\psi = \frac{\partial\psi}{\partial x}\delta x + \frac{\partial\psi}{\partial y}\delta y + \frac{\partial\psi}{\partial z}\delta z \tag{5.99}$$

For variations to be admissible, we require $\delta\psi = 0$, because the differential variations $(\delta x, \delta y, \delta z)$ must be locally consistent with $\psi(x+\delta x, y+\delta y, z+\delta z) = 0$. Since, for all the infinity of points that satisfy the constraint $\psi(x, y, z) = 0$, to minimize $\phi(x, y, z)$, we seek the particular stationary point(s) $\hat{x}, \hat{y}, \hat{z})$ that satisfy

$$\delta\phi = 0 = \frac{\partial\phi}{\partial x}\delta x + \frac{\partial\phi}{\partial y}\delta y + \frac{\partial\phi}{\partial z}\delta z \tag{5.100}$$

$$\delta\psi = 0 = \frac{\partial\psi}{\partial x}\delta x + \frac{\partial\psi}{\partial y}\delta y + \frac{\partial\psi}{\partial z}\delta z \tag{5.101}$$

In the absence of the constraint of Eq. (5.97), we could argue that $(\delta x, \delta y, \delta z)$ are arbitrary. This gives the familiar necessary conditions for an un-constrained minimum $\frac{\partial\phi}{\partial x} = 0; x \to y, z$. However, $(\delta x, \delta y, \delta z)$ cannot be taken arbitrarily, due to the condition of Eq. (5.101). Following Lagrange, we can locally eliminate any of the three variations $(\delta x, \delta y, \delta z)$ to enforce Eq. (5.101), e.g.

$$\delta z = -\left(\frac{1}{\frac{\partial\psi}{\partial z}}\right)\left(\frac{\partial\psi}{\partial x}\delta x + \frac{\partial\psi}{\partial y}\delta y\right) \tag{5.102}$$

Thus, for all infinity of differential variations $(\delta x, \delta y)$, δz from Eq. (5.102) guarantees (so long as $\frac{\partial\psi}{\partial z} \neq 0$) that Eq. (5.101) is satisfied — i.e., $(\delta x, \delta y, \delta z)$ lie in the tangent plane whose normal to $\nabla\psi(x, y, z)$ with $\psi(x, y, z) = 0$. Substituting Eq. (5.102) into Eq. (5.100) gives

$$\delta\phi = \left[\frac{\partial\phi}{\partial x} - \left(\frac{\frac{\partial\phi}{\partial z}}{\frac{\partial\psi}{\partial z}}\right)\frac{\partial\psi}{\partial x}\right]\delta x + \left[\frac{\partial\phi}{\partial y} - \left(\frac{\frac{\partial\phi}{\partial z}}{\frac{\partial\psi}{\partial z}}\right)\frac{\partial\psi}{\partial y}\right]\delta y = 0 \tag{5.103}$$

Thus we have "differentially eliminated" δz. Since $(\delta x, \delta y)$ must be consistent with $\psi(x, y, z) = 0$ and $\delta\psi(x, y, z) = 0$, then Eq. (5.103) can be interpreted as the "constrained variation" of ϕ, along the curve of intersection of $\phi(x, y, z)$ with $\psi(x, y, z) = 0$. Since all infinity of arbitrary $(\delta x, \delta y)$ can now be admitted (while δz from Eq. (5.102) guarantees satisfaction of Eq. (5.101)), we can argue that

the coefficients of $(\delta x, \delta y)$ in Eq. (5.103) must vanish as the necessary conditions:

$$\begin{aligned} \frac{\partial \phi}{\partial x} - \left(\frac{\frac{\partial \phi}{\partial z}}{\frac{\partial \psi}{\partial z}} \right) \frac{\partial \psi}{\partial x} = 0 = f_1(x,y,z) \\ \frac{\partial \phi}{\partial y} - \left(\frac{\frac{\partial \phi}{\partial z}}{\frac{\partial \psi}{\partial z}} \right) \frac{\partial \psi}{\partial y} = 0 = f_2(x,y,z) \\ \psi(x,y,z) = 0 = f_3(x,y,z) \end{aligned} \tag{5.104}$$

The equations in (5.104) provide three algebraic equations containing three unknowns (x, y, z). The stationary points $(\hat{x}, \hat{y}, \hat{z})$ that satisfy Eq. (5.104) must be further evaluated to confirm which points are local maxima, minima, and/or saddle points. Upon evaluating all stationary points, the global minimum can be discerned as the smallest of the local minima, assuming at least one local minimum stationary point is found.

The above necessary conditions are not unique, because, instead of differentially eliminating δz, we could have chosen to eliminate δx or δy. These lead to the two alternate forms of the constrained necessary conditions. For δy eliminated with $(\delta x, \delta z)$ arbitrary we find:

$$\begin{aligned} \frac{\partial \phi}{\partial x} - \left(\frac{\frac{\partial \phi}{\partial y}}{\frac{\partial \psi}{\partial y}} \right) \frac{\partial \psi}{\partial x} = 0 \\ \frac{\partial \phi}{\partial z} - \left(\frac{\frac{\partial \phi}{\partial y}}{\frac{\partial \psi}{\partial y}} \right) \frac{\partial \psi}{\partial z} = 0 \\ \psi(x,y,z) = 0 \end{aligned} \tag{5.105}$$

and, for δx eliminated with $(\delta y, \delta z)$ arbitrary, we find:

$$\begin{aligned} \frac{\partial \phi}{\partial y} - \left(\frac{\frac{\partial \phi}{\partial x}}{\frac{\partial \psi}{\partial x}} \right) \frac{\partial \psi}{\partial y} = 0 \\ \frac{\partial \phi}{\partial z} - \left(\frac{\frac{\partial \phi}{\partial x}}{\frac{\partial \psi}{\partial x}} \right) \frac{\partial \psi}{\partial z} = 0 \\ \psi(x,y,z) = 0 \end{aligned} \tag{5.106}$$

Lagrange noticed this lack of uniqueness and "automated" the derivation of all possibilities by introducing a free multiplier parameter λ, and set

$$\begin{aligned} \delta\phi + \lambda\delta\psi = \\ \left(\frac{\partial \phi}{\partial x} + \lambda \frac{\partial \psi}{\partial x} \right) \delta x + \left(\frac{\partial \phi}{\partial y} + \lambda \frac{\partial \psi}{\partial y} \right) \delta y + \left(\frac{\partial \phi}{\partial z} + \lambda \frac{\partial \psi}{\partial z} \right) \delta z = 0 \end{aligned} \tag{5.107}$$

As before, it "isn't fair" to set the three ( ) terms to zero using the argument that $(\delta x, \delta y, \delta z)$ are arbitrary and independent — but since λ is arbitrary, we can set

any one of the three ( ) terms to zero to determine λ (and thereby eliminate one of the three of $(\delta x, \delta y, \delta z)$) — since the remaining two of $(\delta x, \delta y, \delta z)$ can be chosen arbitrarily, all three ( ) terms must vanish. This argument leads to the following four equations as the constrained necessary conditions:

$$\begin{aligned} \frac{\partial \phi}{\partial x} + \lambda \frac{\partial \psi}{\partial x} &= 0 \\ \frac{\partial \phi}{\partial y} + \lambda \frac{\partial \psi}{\partial y} &= 0 \\ \frac{\partial \phi}{\partial z} + \lambda \frac{\partial \psi}{\partial z} &= 0 \\ \psi(x, y, z) &= 0 \end{aligned} \tag{5.108}$$

It is easy to verify that all three sets of constrained necessary conditions (Eqs. (5.104), (5.105) or (5.106)) are implicit in Eq. (5.108), depending upon which equation is used to solve for λ and then eliminating λ in the other equations. More generally, the equations in (5.108) provide four equations to determine the four unknowns (x, y, z, λ).

Lagrange noticed that the above necessary conditions could be obtained by taking the gradient with respect to (x, y, z, λ) of the *augmented function*

$$\Phi \equiv \phi(x, y, z) + \lambda \psi(x, y, z) \tag{5.109}$$

The Lagrange multiplier rule is given in Eqs. (5.90), (5.91) and (5.92), and is proved by a straight-forward extension of the above developments.

Example 5.5: Assume we would like to minimize

$$\phi(x, y) = x^2 + y^2 \tag{5.110}$$

subject to

$$\psi(x, y,) = (x - 5)^2 + (y - 5)^2 - 1 = 0 \tag{5.111}$$

Geometrically, we seek the point on the circle of Eq. (5.111) that is nearest the origin. We form the augmented function

$$\Phi = x^2 + y^2 + \lambda \left[(x - 5)^2 + (y - 5)^2 - 1 \right] \tag{5.112}$$

Following the Lagrange multiplier rule, the necessary conditions are

$$\frac{\partial \Phi}{\partial x} = 2x + 2\lambda(x - 5) = 0 \tag{5.113}$$

$$\frac{\partial \Phi}{\partial y} = 2y + 2\lambda(y - 5) = 0 \tag{5.114}$$

$$\frac{\partial \Phi}{\partial \lambda} = (x - 5)^2 + (y - 5)^2 - 1 = 0 \tag{5.115}$$

From Eqs. (5.113) and (5.114), we solve for (x, y) as a function of λ as

$$x = \frac{5\lambda}{1 + \lambda} \qquad y = \frac{5\lambda}{1 + \lambda} \tag{5.116}$$

Observing Eq. (5.115), we form

$$\begin{aligned} x - 5 &= 5\left(\frac{\lambda - 1}{\lambda + 1}\right) \\ y - 5 &= 5\left(\frac{\lambda - 1}{\lambda + 1}\right) \end{aligned} \tag{5.117}$$

Substituting Eq. (5.117) into Eq. (5.115) we obtain

$$\left(\frac{\lambda - 1}{\lambda + 1}\right) = \pm\frac{1}{5\sqrt{2}} \tag{5.118}$$

These, in turn, give from Eq. (5.117) the stationary points

$$\begin{aligned} \hat{x} &= 5 \pm \frac{1}{\sqrt{2}} \\ \hat{y} &= 5 \pm \frac{1}{\sqrt{2}} \end{aligned} \tag{5.119}$$

It is easy to verify from Eq. (5.110), the global constrained minimum is at the stationary point

$$(\hat{x}, \hat{y}) = \left(5 - \frac{1}{\sqrt{2}}, 5 - \frac{1}{\sqrt{2}}\right) \tag{5.120}$$

and

$$\hat{\phi} = \hat{x}^2 + \hat{y}^2 = \left(5\sqrt{2} - 1\right)^2 \tag{5.121}$$

and similarly $(\hat{x}, \hat{y}) = (5 + \frac{1}{\sqrt{2}}, 5 + \frac{1}{\sqrt{2}})$, locates the global maximum of $\hat{\phi} = (5\sqrt{2} + 1)^2$.

5.3 Lagrangian Dynamics

As presented above, D'Alembert's principle offers a fundamental advantage over Newton's second law, in that the internal forces and all other virtually non-working constraint forces can be simply ignored in developing the equations of motion. On the other hand, the vector kinematic algebraic overhead associated with Newton's second law and D'Alembert's Principle is essentially identical, since both require vector kinematics to be taken through the acceleration level. In this section, we develop the first of several classical formulations (Lagrange's Equations) which require only velocity level vector kinematics. For the developments so far in this chapter, we use the system of particles model for the system; these developments will subsequently be generalized to accommodate rigid bodies, systems of rigid bodies, and general collections of particles, rigid bodies, and distributed parameter systems.

5.3.1 Minimal Coordinate Systems and Unconstrained Motion

We begin by writing the D'Alembert's Principle form of the system differential equations of motion from Eqs. (5.18), (5.20) as

$$\sum_{i=1}^{N} m_i \ddot{\boldsymbol{R}}_i \cdot \frac{\partial \boldsymbol{R}_i}{\partial q_j} = \sum_{i=1}^{N} \boldsymbol{f}_i \cdot \frac{\partial \boldsymbol{R}_i}{\partial q_j} \quad for\, j = 1, 2, ..., n$$

which using cancellation of dots identity of Eqs. (5.22) become

$$\sum_{i=1}^{N} m_i \ddot{\boldsymbol{R}}_i \cdot \frac{\partial \dot{\boldsymbol{R}}_i}{\partial \dot{q}_j} = \sum_{i=1}^{N} \boldsymbol{f}_i \cdot \frac{\partial \dot{\boldsymbol{R}}_i}{\partial \dot{q}_j} \quad for\, j = 1, 2, ..., n \tag{5.122}$$

Lagrange was apparently the first to recognize that differential equations closely related to Eqs. (5.122) could be generated using position and velocity coordinate gradients of energy functions. We verify these classical developments, beginning with the definition of Kinetic energy for a system of N particles:

$$T = \frac{1}{2} \sum_{i=1}^{N} m_i \dot{\boldsymbol{R}}_i \cdot \dot{\boldsymbol{R}}_i \tag{5.123}$$

We observe that the partial derivatives of T with respect to $(q_j, \dot{q}_j)$ are

$$\frac{\partial T}{\partial q_j} = \sum_{i=1}^{N} m_i \dot{\boldsymbol{R}}_i \cdot \frac{\partial \dot{\boldsymbol{R}}_i}{\partial q_j} \quad \text{and} \quad \frac{\partial T}{\partial \dot{q}_j} = \sum_{i=1}^{N} m_i \dot{\boldsymbol{R}}_i \cdot \frac{\partial \dot{\boldsymbol{R}}_i}{\partial \dot{q}_j} \tag{5.124}$$

Now consider the following developments:

$$\begin{aligned}
\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) &= \sum_{i=1}^{N} m_i \ddot{\boldsymbol{R}}_i \cdot \frac{\partial \dot{\boldsymbol{R}}_i}{\partial \dot{q}_j} + \sum_{i=1}^{N} m_i \dot{\boldsymbol{R}}_i \cdot \frac{d}{dt}\left(\frac{\partial \dot{\boldsymbol{R}}_i}{\partial \dot{q}_j}\right) \\
&= \sum_{i=1}^{N} (\boldsymbol{f}_i + \boldsymbol{f}_{ci}) \cdot \frac{\partial \dot{\boldsymbol{R}}_i}{\partial \dot{q}_j} + \sum_{i=1}^{N} m_i \dot{\boldsymbol{R}}_i \cdot \frac{d}{dt}\left(\frac{\partial \boldsymbol{R}_i}{\partial q_j}\right) \\
&= \sum_{i=1}^{N} \boldsymbol{f}_i \cdot \frac{\partial \dot{\boldsymbol{R}}_i}{\partial \dot{q}_j} + \sum_{i=1}^{N} m_i \dot{\boldsymbol{R}}_i \cdot \left(\frac{\partial \dot{\boldsymbol{R}}_i}{\partial q_j}\right) \\
&= \sum_{i=1}^{N} \boldsymbol{f}_i \cdot \frac{\partial \dot{\boldsymbol{R}}_i}{\partial \dot{q}_j} + \frac{\partial T}{\partial q_j}
\end{aligned} \tag{5.125}$$

From the above and Eqs. (5.122), we establish the following elegant result (*Lagrange's Equations*)

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = \sum_{i=1}^{N} \boldsymbol{f}_i \cdot \frac{\partial \dot{\boldsymbol{R}}_i}{\partial \dot{q}_j} \equiv Q_j, \quad \text{for} \quad j = 1, 2, \ldots, n \tag{5.126}$$

This is the most fundamental version of Lagrange's Equations, and these equations are amongst the most important results in analytical dynamics. We will subsequently seek alternative paths to these equations, via Hamilton's Principle, for example. Notice from Eqs. (5.126), there results a system of n second order differential equations of the form

$$f_j(t; q_1, \dots, q_n; \dot{q}_1, \dots, \dot{q}_n; \ddot{q}_1, \dots, \ddot{q}_n) = Q_j, \quad j = 1, 2, \dots, n \tag{5.127}$$

that are generated by simply differentiating the system kinetic energy $T(t; q_j; \dot{q}_j)$ with respect to the chosen set of generalized coordinates $(q_j, \dot{q}_j,\ j = 1, \dots, n)$ and summing the dot product of the virtually working forces with the partial velocities $\boldsymbol{v}_{ik} \equiv \frac{\partial \dot{\boldsymbol{R}}_i}{\partial \dot{q}_k}$. While we have have developed these equations in the context of a system of particles, we will see that an appropriate definition of kinetic energy results in these equations applying to systems of rigid bodies and particles. Also of significance, we will find that these apply to some degree of approximation, upon introducing appropriate *spatial discretization methods* such as the Ritz method or the finite element method,[3] to approximate the dynamics of distributed parameter systems such as vibrating flexible structures. Also implicit in these equations is the assumption that, if constraints are present, they are simple algebraic holonomic constraints which have been kinematically eliminated to establish a *minimal coordinate* description of the system, i.e., constraints have been enforced in the kinematic description of the system, so the generalized coordinates $\{q_1, q_2, \dots, q_n\}$ *must be independent* in Eqs. (5.126). We consider three examples to introduce the reader to the process of applying Lagrange's Equations.

Example 5.6: With reference to Fig. 5.8, the radius is decreasing at a constant rate, determine angular velocity as a function of time. The general expression for velocity is $\dot{\boldsymbol{R}} = \dot{r}\hat{\boldsymbol{e}}_r + r\dot{\theta}\hat{\boldsymbol{e}}_\theta$. Upon imposing the constraint that $\dot{r} = -c = constant$, then only θ remains as a generalized coordinate and imposing this kinematic constraint, the velocity is $\dot{\boldsymbol{R}} = -c\hat{\boldsymbol{e}}_r + (r_o - ct)\dot{\theta}\hat{\boldsymbol{e}}_\theta$. Thus the kinetic energy has the specific structure: $T = \frac{1}{2}m[c^2 + (r_o - ct)^2\dot{\theta}^2]$. From Lagrange's equations in the form

$$\frac{d}{dt}(\frac{\partial T}{\partial \dot{\theta}}) - \frac{\partial T}{\partial \theta} = (F_r\hat{\boldsymbol{e}}_r) \cdot \frac{\partial \dot{\boldsymbol{R}}}{\partial \dot{\theta}} = Q_\theta$$

we have

$$\frac{d}{dt}[m(r_o - ct)^2\dot{\theta}] - [0] = [0]$$

from which

$$m(r_o - ct)^2\dot{\theta} = \text{constant} = mr_o^2\dot{\theta}_o$$

giving the desired result

$$\dot{\theta} = (\frac{r_o}{r_o - ct})^2\dot{\theta}_o$$

Discussion: Observe that the kinetic energy does not depend on θ and $Q_\theta = 0$, and a consequence, $\frac{d}{dt}(\frac{\partial T}{\partial \dot{\theta}}) = 0$ and $p_\theta = \frac{\partial T}{\partial \dot{\theta}} = constant$. Whenever

the derivative $\dot{q}_j$ of a coordinate q_j appears in T, but not the coordinate itself, then such coordinates are called *cyclic*. If the generalized force Q_j is zero, then the corresponding generalized conjugate momentum $p_j = \frac{\partial T}{\partial \dot{q}_j}$ is a *constant of the motion*. When the observation is made that a coordinate is cyclic and $Q_j = 0$, then the corresponding generalized momentum can immediately be set to a constant, thereby by-passing the formal algebra of Lagrange's equations. In this case, $p_\theta = mr^2\dot{\theta} = constant$ has the physical interpretation of angular momentum conservation.

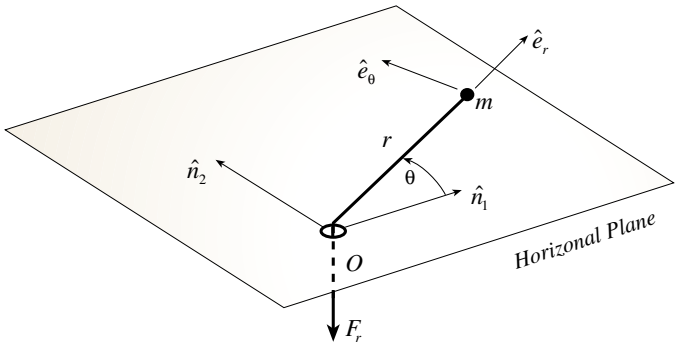


Figure 5.8: Particle on Table with Constantly Decreasing Radius

---

Example 5.7: Consider two particles sliding on a frictionless horizontal plane as shown in Figure 5.9. The particles are connected by a linear spring whose un-stretched length is R. The form of the equations of motion depend upon

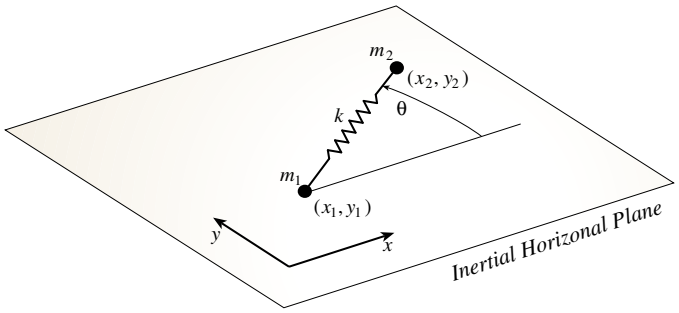


Figure 5.9: Two Particles Sliding on a fixed Inertial Plane

the coordinates chosen, also, judicious coordinates often reveal easier insight into the system behavior. In particular, the same system may have cyclic coordinates with one coordinate choice, but not for other coordinate choices.

To illustrate, consider the two Cartesian coordinate choices:

set 1	(x_1, y_1, x_2, y_2)	Inertial Coordinates
set 2	$(x_c, y_c, x_{12}, y_{12})$	Mass Center Motion and Relative Motion

where

$$\begin{aligned} x_c &= \frac{1}{m}\left(m_1 x_1 + m_2 x_2\right), \quad m = m_1 + m_2 \\ y_c &= \frac{1}{m}\left(m_1 y_1 + m_2 y_2\right) \end{aligned} \tag{5.128}$$

$$x_{12} = x_2 - x_1, \qquad y_{12} = y_2 - y_1 \tag{5.129}$$

The spring force has a magnitude $k(r-d)$ where $r^2 = x_{12}^2 + y_{12}^2$. The direction of the force is along the instantaneous line of centers, oriented by angle θ. Note $\cos\theta = x_{12}/r$, $\sin\theta = y_{12}/r$. Thus Newton's laws lead to the equations of motion for set 1 as

$$\begin{aligned} m_1\ddot{x}_1 &= \quad k(r-d)\frac{x_{12}}{r}, & m_1\ddot{y}_1 &= \quad k(r-d)\frac{y_{12}}{r} \\ m_2\ddot{x}_2 &= -k(r-d)\frac{x_{12}}{r}, & m_2\ddot{y}_2 &= -k(r-d)\frac{y_{12}}{r} \end{aligned} \tag{5.130}$$

with $r = \sqrt{x_{12}^2 + y_{12}^2}$. For the special case of $d = 0$, these simplify to the linear system

$$\begin{aligned} m_1\ddot{x}_1 &= \quad k(x_2 - x_1), & m_1\ddot{y}_1 &= \quad k(y_2 - y_1) \\ m_2\ddot{x}_2 &= -k(x_2 - x_1), & m_2\ddot{y}_2 &= -k(y_2 - y_1) \end{aligned} \tag{5.131}$$

For the set 2 choice of coordinates, use Eqs. (5.128) and (5.129) to verify that the general dynamical equations are for the case where $d \neq 0$

$$\begin{aligned} m\ddot{x}_c &= 0, & \left(\frac{m_1 m_2}{m_1 + m_2}\right)\ddot{x}_{12} &= -k(r-d)\frac{x_{12}}{r} \\ m\ddot{y}_c &= 0, & \left(\frac{m_1 m_2}{m_1 + m_2}\right)\ddot{y}_{12} &= -k(r-d)\frac{y_{12}}{r} \end{aligned} \tag{5.132}$$

For the special case where $d = 0$, the equations of motion become

$$\begin{aligned} m\ddot{x}_c &= 0, & \left(\frac{m_1 m_2}{m_1 + m_2}\right)\ddot{x}_{12} &= -kx_{12} \\ m\ddot{y}_c &= 0, & \left(\frac{m_1 m_2}{m_1 + m_2}\right)\ddot{y}_{12} &= -ky_{12} \end{aligned} \tag{5.133}$$

With the set 2 coordinate choice, we see in all cases that the mass center moves in a straight line, and the relative motion dynamics un-couples from the mass center motion. For $d = 0$, we see that the relative motion in the x and y direction also un-couples and becomes simple harmonic motion with the natural frequency $\omega = \sqrt{k/\frac{m_1 m_2}{m_1 + m_2}}$.

You can verify that the same differential equations in Eq. (5.130) and (5.133) result if Lagrange's equations are utilized for coordinate set 1. The kinetic energy T is given by

$$T = \frac{1}{2}\left(m_1\dot{x}_1^2 + m_1\dot{y}_1^2 + m_2\dot{x}_2^2 + m_2\dot{y}_2^2\right) \tag{5.134}$$

while the potential energy V is expressed as

$$V = V(x_1, y_1, x_2, y_2) = \frac{1}{2}k(r-d)^2 \tag{5.135}$$

with $r = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2}$. For coordinate set 1, the solution is straight forward. However, for coordinate set 2, you will need to express the energies as functions of the set 2 coordinates as

$$\begin{aligned} T &= T(x_c, y_c, x_{12}, y_{12}) \\ V &= V(x_c, y_c, x_{12}, y_{12}) \end{aligned} \tag{5.136}$$

and for this you need the coordinate transformations:

$$\begin{aligned} x_i &= f_i(x_c, y_c, x_{12}, y_{12}) \qquad i = 1,2 \\ y_i &= g_i(x_c, y_c, x_{12}, y_{12}) \qquad i = 1,2 \end{aligned} \tag{5.137}$$

These are obtained by inverting Eqs. (5.128) and (5.129). You will find (x_c, y_c) do not appear explicitly in T and V. Therefore (x_c, y_c) are cyclic and $\dot{x}_c =$ constant, $\dot{y}_c =$ constant are immediately obvious. These steps are left as an exercise.

5.3.2 Lagrange's Equations for Conservative Forces

Recall that the generalized forces can in general be written as $Q_j = \sum_{i=1}^{N} \boldsymbol{f}_i \cdot \frac{\partial \boldsymbol{R}_i}{\partial q_j}$, and for the case of conservative forces, $\boldsymbol{f}_i = -\frac{\partial V}{\partial \boldsymbol{R}_i}$, so the generalized conservative forces can be written as

$$Q_j = \sum_{i=1}^{N} \boldsymbol{f}_i \cdot \frac{\partial \boldsymbol{R}_i}{\partial q_j} = -\sum_{i=1}^{N} \frac{\partial V}{\partial \boldsymbol{R}_i} \cdot \frac{\partial \boldsymbol{R}_i}{\partial q_j} = -\frac{\partial V}{\partial q_j} \quad j = 1, 2, \ldots, n \tag{5.138}$$

We introduce the definition of the *Lagrangian* function:

$$\mathcal{L} = \mathcal{L}(t, q_1, \ldots, q_n; \dot{q}_1, \ldots, \dot{q}_n) \equiv T - V \tag{5.139}$$

Since $V = V(t, q_1, \ldots, q_n)$, it follows that $\frac{\partial \mathcal{L}}{\partial \dot{q}_j} = \frac{\partial T}{\partial \dot{q}_j}$ and $\frac{\partial \mathcal{L}}{\partial q_j} = \frac{\partial T}{\partial q_j} - \frac{\partial V}{\partial q_j}$, then for the case that all forces are conservative, then Eqs. (5.127) assume the most famous form of Lagrange's equations

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j}\right) - \frac{\partial \mathcal{L}}{\partial q_j} = 0, \quad \text{for} \quad j = 1, 2, \ldots, n \tag{5.140}$$

For many elementary conservative systems, the potential and kinetic energy can be simply written with a minimum of derivations; for these cases, Eqs. (5.140) do not require derivation of any generalized forces and are therefore especially attractive. In more general circumstances for which there are both conservative and nonconservative forces which do virtual work, then Eqs. (5.140) are replaced by the more general form

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j}\right) - \frac{\partial \mathcal{L}}{\partial q_j} = Q_{ncj} = \sum_{i=1}^{N} \boldsymbol{f}_{nci} \cdot \frac{\partial \dot{\boldsymbol{R}}_i}{\partial \dot{q}_j}, \quad \text{for} \quad j = 1, 2, \ldots, n \qquad (5.141)$$

where $\boldsymbol{f}_{nci}$ is the nonconservative force acting on m_i. Of course, the special case of non-conservative forces which act normal to the vector $\frac{\partial \boldsymbol{R}_i}{\partial q_j} \equiv \frac{\partial \dot{\boldsymbol{R}}_i}{\partial \dot{q}_j}$ has $Q_{nc} = 0$ and this special class of non-conservative systems' equations of motion are also given by Eq. (5.140). The utility of these forms of Lagrange's equations are illustrated in the following examples.

Example 5.8: With reference to Fig. 5.10, the linear spring pendulum is con-

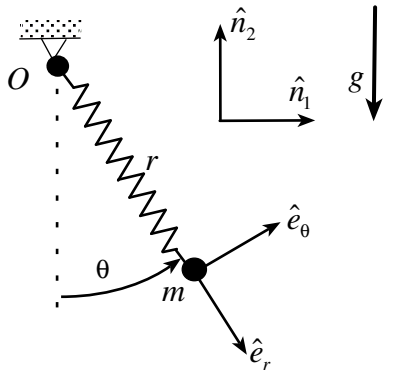


Figure 5.10: Classical Spring Pendulum

sidered (nominal unstretched spring length (r_o), linear spring constant (k)). The objective is to use the version of Lagrange's Eqs. (5.141) to efficiently develop the equations of motion. It is evident that the only virtually working forces are the spring force and gravity, and that there are two generalized coordinates (r, θ). The position and velocity vectors are given by

$$\boldsymbol{R} = r\hat{\boldsymbol{e}}_r, \quad \dot{\boldsymbol{R}} = \dot{r}\hat{\boldsymbol{e}}_r + r\dot{\theta}\hat{\boldsymbol{e}}_\theta$$

Thus the kinetic energy is

$$T = \frac{1}{2} m \dot{\boldsymbol{R}} \cdot \dot{\boldsymbol{R}} = \frac{1}{2} m(\dot{r}^2 + r^2\dot{\theta}^2)$$

From inspection of Fig. 5.10,, it is evident that the potential energy function is

$$V = mgr(1 - \cos\theta) + \frac{1}{2}k(r - r_o)^2$$

so the Lagrangian function is

$$\mathcal{L} = T - V = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - mgr(1-\cos\theta) - \frac{1}{2}k(r-r_o)^2$$

$$m\ddot{r} + k(r-r_o) - mr\dot{\theta}^2 - mg\cos\theta = 0$$
$$mr^2\ddot{\theta} + mgr\sin\theta + 2mr\dot{r}\dot{\theta} = 0$$

To appreciate the efficiency of the above path to obtain the equations of motion, it is instructive to repeat the solution using Eqs. (5.126), including the formulation of generalized forces

$$Q_r = [-k(r-r_o)\hat{\boldsymbol{e}}_r - mg\hat{\boldsymbol{n}}_2]\cdot\frac{\partial\dot{\boldsymbol{R}}}{\partial\dot{r}} = \ldots = -k(r-r_o) + mg\cos\theta$$
$$Q_\theta = [-k(r-r_o)\hat{\boldsymbol{e}}_r - mg\hat{\boldsymbol{n}}_2]\cdot\frac{\partial\dot{\boldsymbol{R}}}{\partial\dot{\theta}} = \ldots = -mgr\sin\theta$$

arising from the spring force and the gravity force. When conservative forces that have easily established potential energy functions are present, then Eqs. (5.141) clearly results in significant reductions in the algebra associated with derivation of the generalized force functions. Obviously, we can ignore all internal and constraint forces, as well as all conservative forces, and consider only the nonconservative virtually working forces when determining the generalized force functions. In this case, all virtually working forces are conservative, so it is not necessary to formulate any generalized forces (they are implicitly accounted for by being included in the potential energy function).

Example 5.9: With reference to Fig. 5.11, we generalize the earlier example

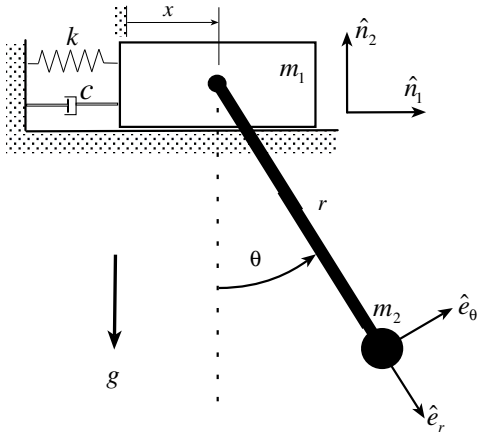


Figure 5.11: Damped Cart - Pendulum System

solved using Newton's laws and D'Alembert's Principle, to include the dashpot linear damper with associated damping force $(-c\dot{x}\hat{\boldsymbol{n}}_1)$. We note that, among the virtually working forces, only the damping force is nonconservative. This means that the generalized forces are easily determined as

$$Q_x = [-c\dot{x}\hat{\boldsymbol{n}}_1] \cdot \frac{\partial \dot{\boldsymbol{R}}_1}{\partial \dot{x}} = -c\dot{x}$$
$$Q_\theta = 0$$

$$\dot{\boldsymbol{R}}_1 = \boldsymbol{V}_1 = \dot{x}\hat{\boldsymbol{n}}_1$$
$$\dot{\boldsymbol{R}}_2 = \boldsymbol{V}_2 = \dot{x}\hat{\boldsymbol{n}}_1 + r\dot{\theta}\hat{\boldsymbol{e}}_\theta$$

The kinetic and potential energies are

$$T = \frac{1}{2}(m_1 + m_2)\dot{x}^2 + \frac{1}{2}m_2[\dot{x}^2 + r^2\dot{\theta}^2 + 2\dot{x}r\dot{\theta}\cos\theta]$$
$$V = \frac{1}{2}kx^2 + m_2 g(1 - \cos\theta)$$

It is easy to verify that Eqs. (5.141) immediately give

$$(m_1 + m_2)\ddot{x} + (m_2 r\cos\theta)\ddot{\theta} - m_2 r\dot{\theta}^2\sin\theta = -kx - c\dot{x}$$
$$(m_2 r\cos\theta)\ddot{x} + (m_2 r^2)\ddot{\theta} = -mgr\sin\theta$$

This process is an elegant alternative to the previous development of the equations of motion via Newton's laws or D'Alembert's Principle, especially with regard to requiring only velocity-level vector kinematics.

5.3.3 Redundant Coordinate Systems and Constrained Motion

Here we extend Lagrange's equations to consider redundant coordinates subject to Pfaffian non-holonomic constraints. Recall our path to Lagrange's equations for a system of N particles.

$$\text{Newton's 2nd Law:} \quad \boldsymbol{f}_i + \boldsymbol{f}_{c_i} - m_i\ddot{\boldsymbol{R}}_i = 0 \tag{5.142}$$
$$\text{Virtual Work:} \quad \delta W_i = (\boldsymbol{f}_i + \boldsymbol{f}_{c_i} - m\ddot{\boldsymbol{R}}_i)\cdot\delta\boldsymbol{R}_i = 0 \tag{5.143}$$
$$\text{Consistent Virtual Displacements:} \quad \delta W_{c_i} = \boldsymbol{f}_{c_i}\cdot\delta\boldsymbol{R}_i = 0 \tag{5.144}$$
$$\text{Virtual Work Becomes:} \quad \delta W_i = (\boldsymbol{f}_i - m\ddot{\boldsymbol{R}}_i)\cdot\delta\boldsymbol{R}_i = 0 \tag{5.145}$$
$$\text{Total Virtual Work:} \quad \delta W = \sum_{i=1}^{N}(\boldsymbol{f}_i - m_i\ddot{\boldsymbol{R}}_i)\cdot\delta\boldsymbol{R}_i = 0 \tag{5.146}$$
$$\text{Introduce Generalized Coordinates:} \quad \boldsymbol{R}_i = \boldsymbol{R}_i(q_1, \ldots, q_n, t) \tag{5.147}$$

We made us of $\boldsymbol{R}_i = \boldsymbol{R}_i(q_1, \ldots, q_n, t)$ to write

$$\delta\boldsymbol{R}_i = \sum_{j=1}^{n}\frac{\partial\boldsymbol{R}_i}{\partial q_j}\delta q_j \tag{5.148}$$

so that the virtual work δW of Eq. (5.146) can be re-written as

$$\delta W = \sum_{j=1}^{n} \left(\sum_{i=1}^{N} \boldsymbol{f}_i \cdot \frac{\partial \boldsymbol{R}_i}{\partial q_j} - \sum_{i=1}^{N} m_i \ddot{\boldsymbol{R}}_i \cdot \frac{\partial \boldsymbol{R}_i}{\partial q_j} \right) \delta q_j = 0 \tag{5.149}$$

We previously defined the generalized force

$$Q_j \equiv \sum_{i=1}^{N} \boldsymbol{f}_i \cdot \frac{\partial \boldsymbol{R}_i}{\partial q_j} \qquad j = 1, 2, \ldots, n \tag{5.150}$$

and we proved the identity

$$\sum_{i=1}^{N} m_i \ddot{\boldsymbol{R}}_i \cdot \frac{\partial \boldsymbol{R}_i}{\partial q_j} \equiv \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} \qquad j = 1, 2, \ldots, n \tag{5.151}$$

that that Eq. (5.149) becomes

$$\delta W = \sum_{j=1}^{n} \left[Q_j + \frac{\partial T}{\partial q_j} - \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) \right] \delta q_j = 0 \tag{5.152}$$

Now, in the previous developments of section 5.3.1, we assumed n was the number of degrees of freedom (which implicitly means all holonomic constraints have been eliminated, and that $\{q_1, \ldots, q_n\}$ are a minimal set of independent coordinates). For this case, we argued that the virtual displacements δq_j could be chosen arbitrarily — Eq. (5.152) can only be satisfied if each bracketed $[\ ]_j$ coefficient of δq_j must vanish independently — leading to the most familiar forms of Lagrange's equations.

However, if $\{q_1, \ldots, q_n\}$ are *not* independent, and constraints are present, then we cannot set the coefficient of δq_j in Eq. (5.152) to zero. In particular, consider the case of m differential non-holonomic constraints of the Pfaffian form

$$\sum_{j=1}^{n} A_{kj} \dot{q}_j + B_k = 0 \qquad k = 1, 2, \ldots, m \tag{5.153}$$

or, along an trajectory, we have the differential constraint

$$\sum_{j=1}^{n} A_{kj} dq_j + B_k dt = 0 \qquad k = 1, 2, \ldots, m \tag{5.154}$$

For instantaneous virtual displacements consistent with these constraints in Eq. (5.154), we have

$$\sum_{j=1}^{n} A_{kj} \delta q_j = 0 \qquad k = 1, 2, \ldots, m \tag{5.155}$$

Any m of the n δq_j's could be solved from the algebraic equations in (5.155) as a function of the remaining $(n-m)$ δq_j's analogous to the above developments for the Lagrange multiplier rule. These m equations could be used in Eq. (5.152) to eliminate m δq_j's. Recollection of the terms would give virtual work as a constrained variation function of $(n-m)$ δq_j's which can be chosen independently. Again, analogous to the development of the Lagrange multiplier rule, it is easy to argue that the following Lagrange multiplier rule "automates" all possible differential constraint eliminations:

$$\delta W = \sum_{j=1}^{n} \left[Q_j + \sum_{k=1}^{m} \lambda_k A_{kj} + \frac{\partial T}{\partial q_j} - \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) \right] \delta q_j = 0 \tag{5.156}$$

Leading to the constrained version of Lagrange's equations of motion

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = Q_j + \sum_{k=1}^{m} \lambda_k A_{kj} \qquad j = 1, 2, \ldots, n \tag{5.157}$$

$$\sum_{j=1}^{n} A_{kj}\dot{q}_j + B_k = 0 \qquad k = 1, 2, \ldots, m \tag{5.158}$$

constituting an $(n+m)$ differential-algebraic system of equations in the $(n+m)$ unknowns

$$\{q_1(t), \ldots, q_n(t), \lambda_1(t), \ldots, \lambda_m(t)\}$$

We mention that a more convenient version of Lagrange's equation, in lieu of Eq. (5.157), can be written as a generalization of Eq. (5.141) as

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j}\right) - \frac{\partial \mathcal{L}}{\partial q_j} = Q_{nc_j} + \sum_{k=1}^{m} \lambda_k A_{kj} \qquad j = 1, 2, \ldots, n \tag{5.159}$$

Where $\mathcal{L} = T - V$, $V = V(q_1, \ldots, q_n)$ is the potential energy function, $Q_j = -\frac{\partial V}{\partial q_j}$, $\boldsymbol{f}_i = \boldsymbol{f}_{nc_i} - \nabla V$, and $Q_{nc_j} = \sum_{i=1}^{N} \boldsymbol{f}_{nc_i} \cdot \frac{\partial \boldsymbol{R}_i}{\partial q_j}$.

Example 5.10: Consider the particle sliding in a rotating tube as shown in Figure 5.12. The forces are shown in the free-body diagram. F_θ is the normal reaction force due to interaction of the mass m with the tube wall. First, we develop the equations of motion using Newton's second law

$$\boldsymbol{F} = m\ddot{\boldsymbol{R}} \tag{5.160}$$

From kinematics we find that

$$\ddot{\boldsymbol{R}} = (\ddot{r} - r\dot{\theta}^2)\hat{\boldsymbol{e}}_r + (r\ddot{\theta} + 2\dot{r}\dot{\theta})\hat{\boldsymbol{e}}_\theta$$

and introducing the kinematic constraint $\dot{\theta} = \Omega$ =constant, we have

$$\ddot{\boldsymbol{R}} = (\ddot{r} - r\Omega^2)\hat{\boldsymbol{e}}_r + (2\dot{r}\Omega)\hat{\boldsymbol{e}}_\theta$$

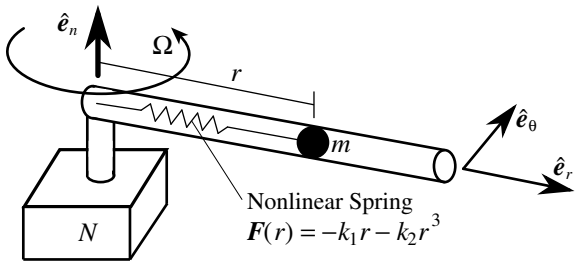


Figure 5.12: Particle Moving in a Rotating Tube

From the free-body diagram, we have

$$\boldsymbol{F} = -(k_1 r + k_2 r^3)\hat{\boldsymbol{e}}_r + (F_\theta)\hat{\boldsymbol{e}}_\theta \tag{5.161}$$

so the equations of motion are

$$-(k_1 r + k_2 r^3)\hat{\boldsymbol{e}}_r + F_\theta \hat{\boldsymbol{e}}_\theta = m\left[(\ddot{r} - r\Omega^2)\hat{\boldsymbol{e}}_r + (2\dot{r}\Omega)\hat{\boldsymbol{e}}_\theta\right] \tag{5.162}$$

leading to the scalar differential equation

$$m\ddot{r} - m\Omega^2 r = -k_1 r - k_2 r^3 \tag{5.163}$$

and the constraint force

$$F_\theta = 2m\dot{r}\Omega \tag{5.164}$$

Secondly, we consider a first form of Lagrange's equations. As the first case, let us consider a minimal coordinate formulation (following the developments of section 5.3.1. The general position and velocity vectors for a particle in a moving plane, as a function of polar coordinates, is given by

$$\boldsymbol{R} = r\hat{\boldsymbol{e}}_r \qquad \dot{\boldsymbol{R}} = (\dot{r})\hat{\boldsymbol{e}}_r + (r\dot{\theta})\hat{\boldsymbol{e}}_\theta \tag{5.165}$$

Imposing $\dot{\theta} = \Omega =$constant gives

$$\dot{\boldsymbol{R}} = \dot{r}\hat{\boldsymbol{e}}_r + r\Omega\hat{\boldsymbol{e}}_\theta \tag{5.166}$$

The kinetic and potential energy functions are

$$T = \frac{1}{2}m\dot{\boldsymbol{R}}\cdot\dot{\boldsymbol{R}} = \frac{1}{2}m(\dot{r}^2 + r^2\Omega^2) \tag{5.167}$$

$$V = \frac{1}{2}k_1 r^2 + \frac{1}{4}k_2 r^4 \tag{5.168}$$

The Lagrangian L is given by

$$\mathcal{L} = T - V = \frac{1}{2}m\dot{r}^2 + \left(-\frac{k_1}{2} + \frac{m\Omega^2}{2}\right)r^2 - \frac{1}{4}k_2 r^4 \tag{5.169}$$

Using Lagrange's equations in the form of Eqs. (5.141), we see that all forces acting on m are conservative, except for $F_\theta\hat{\boldsymbol{e}}_\theta$. However, $F_\theta\hat{\boldsymbol{e}}_\theta$ does no virtual

work (because $(F_\theta \hat{\boldsymbol{e}}_\theta) \cdot (\delta r \hat{\boldsymbol{e}}_r) = 0$) — therefore $Q_{nc} = 0$ and the equations of motion from Eqs. (5.141) reduce to

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j}\right) - \frac{\partial \mathcal{L}}{\partial q_j} = 0 \qquad j = 1, 2, \ldots, n$$

For the one-dimensional dynamical system considered in this present example this yields

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{r}}\right) - \frac{\partial \mathcal{L}}{\partial r} = 0$$

and specifically, substituting Eq. (5.169) gives

$$m\ddot{r} + (k_1 - m\Omega^2)r^2 + k_2 r^3 = 0 \tag{5.170}$$

Both Eqs. (5.163) and (5.170) give the identical differential equation of motion, which can be written as

$$\ddot{r} + (\alpha_1^2 - \Omega^2)r^2 + \alpha_2^2 r^3 = 0 \tag{5.171}$$

with

$$\alpha_1^2 \equiv \frac{k_1}{m} \qquad \alpha_2^2 \equiv \frac{k_2}{m} \tag{5.172}$$

Thirdly, we can use a redundant coordinate set (r, θ), and make use of the developments presented in the current section, namely Eqs. (5.158) and (5.159). The kinetic energy, for general planar motion, is

$$T = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) \tag{5.173}$$

and the Lagrangian is

$$\mathcal{L} = T - V = \frac{1}{2}\dot{r}^2 + \frac{1}{2}mr^2\dot{\theta}^2 - \frac{k_1}{2}r^2 - \frac{k_2}{4}r^4 \tag{5.174}$$

The constraint $\dot{\theta} = \Omega =$ constant is written as the Pfaffian form

$$\dot{\theta} - \Omega = 0 \tag{5.175}$$

We note that Eq. (5.175) is the particular case of Eq. (5.158). Taking $(q_1, q_2) = (r, \theta)$, we have by direct inspection:

$$A_{11} = 0 \qquad A_{12} = 1 \qquad B_1 = \Omega \tag{5.176}$$

The equations of motion follow from Eq. (5.159) as

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{r}}\right) - \frac{\partial \mathcal{L}}{\partial r} &= \lambda(0) \\ \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{\theta}}\right) - \frac{\partial \mathcal{L}}{\partial \theta} &= \lambda(1) \end{aligned} \tag{5.177}$$

Carrying out the differentiations, by substituting Eq. (5.174) into (5.177) we have

$$m\ddot{r} - mr\dot{\theta}^2 + k_1 r + k_2 r^3 = 0 \tag{5.178}$$

$$mr^2\ddot{\theta} + 2mr\dot{r}\dot{\theta} + 0 = \lambda \tag{5.179}$$

Imposing the constraint of Eq. (5.175) on Eqs. (5.178) and (5.179), we obtain

$$m\ddot{r} + (k_1 - m\Omega^2)r + k_2 r^3 = 0 \tag{5.180}$$

$$\lambda = 2mr\dot{r}\Omega \equiv rF_\theta \tag{5.181}$$

We see that Eq. (5.180) is identical to Eqs. (5.170) and (5.171), whereas λ is the moment causes by the constraint force $F_\theta = 2m\dot{r}\Omega$.

5.3.4 Vector-Matrix Form of the Lagrangian Equations of Motion

We note that the Lagrangian equations of motion in Eq. (5.159) and the system of Pfaffian constraints can be written in the vector-matrix form

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{q}}}\right) - \frac{\partial \mathcal{L}}{\partial \boldsymbol{q}} = \boldsymbol{Q}_{nc} + [A]^T\boldsymbol{\lambda} \tag{5.182}$$

and

$$[A]\dot{\boldsymbol{q}} + \boldsymbol{B} = 0 \tag{5.183}$$

where

$$\mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) = T(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) - V(\boldsymbol{q}, t)$$

$$\boldsymbol{Q}_{nc} = \text{col}\left(\sum_{i=1}^{N} \boldsymbol{f}_{nc_i} \cdot \frac{\partial \boldsymbol{R}_i}{\partial q_1}, \sum_{i=1}^{N} \boldsymbol{f}_{nc_i} \cdot \frac{\partial \boldsymbol{R}_i}{\partial q_2}, \dots, \sum_{i=1}^{N} \boldsymbol{f}_{nc_i} \cdot \frac{\partial \boldsymbol{R}_i}{\partial q_n}\right)$$

$$\equiv \text{col}\left(\sum_{i=1}^{N} \boldsymbol{f}_{nc_i} \cdot \frac{\partial \dot{\boldsymbol{R}}_i}{\partial \dot{q}_1}, \sum_{i=1}^{N} \boldsymbol{f}_{nc_i} \cdot \frac{\partial \dot{\boldsymbol{R}}_i}{\partial \dot{q}_2}, \dots, \sum_{i=1}^{N} \boldsymbol{f}_{nc_i} \cdot \frac{\partial \dot{\boldsymbol{R}}_i}{\partial \dot{q}_n}\right)$$

$$\boldsymbol{q} \equiv \begin{pmatrix} q_1 & q_2 & \cdots & q_n \end{pmatrix}^T$$

$$\boldsymbol{\lambda} \equiv \begin{pmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_m \end{pmatrix}^T$$

$$[A] = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix} = [A(\boldsymbol{q})]$$

$$\boldsymbol{B} = \begin{pmatrix} B_1 & B_2 & \cdots B_n \end{pmatrix}^T = \boldsymbol{B}(\boldsymbol{q})$$

Once a vector basis for expressing positions and velocities has been selected, more explicit Lagrange's equations of motion can be written. Starting with the velocity expression in the vector form of

$$\dot{\boldsymbol{R}}_i = \frac{\partial \boldsymbol{R}_i}{\partial t} + \sum_{k=1}^{n} \boldsymbol{v}_{ik}\dot{\boldsymbol{q}}_k \qquad i = 1, 2, \ldots, N \tag{5.184}$$

Letting ${}^{\mathcal{F}}\{\boldsymbol{R}_i\}$ denote a matrix of components of $\boldsymbol{R}_i$ in frame $\mathcal{F}$ and similarly for vectors ${}^{\mathcal{F}}\left\{\frac{\boldsymbol{R}_i}{\partial t}\right\}$ and ${}^{\mathcal{F}}\{\boldsymbol{v}_{ik}\}$, then the vector equations of Eqs. (5.184) can be written as the corresponding matrix equations

$${}^{\mathcal{F}}\{\dot{\boldsymbol{R}}_i\} = {}^{\mathcal{F}}\left\{\frac{\boldsymbol{R}_i}{\partial t}\right\} + \sum_{k=1}^{n} {}^{\mathcal{F}}\{\boldsymbol{v}_{ik}\}\dot{q}_k \qquad i = 1, 2, \ldots, N \tag{5.185}$$

or

$${}^{\mathcal{F}}\{\dot{\boldsymbol{R}}_i\} = {}^{\mathcal{F}}\left\{\frac{\boldsymbol{R}_i}{\partial t}\right\} + [{}^{\mathcal{F}}V_i]\{\dot{\boldsymbol{q}}\} \qquad i = 1, 2, \ldots, N \tag{5.186}$$

where

$$\dot{\boldsymbol{q}} = \begin{pmatrix} \dot{q}_1 & \dot{q}_2 \cdots \dot{q}_n \end{pmatrix}^T$$
$$[{}^{\mathcal{F}}V_i] = \begin{bmatrix} {}^{\mathcal{F}}\{\boldsymbol{V}_{i1}\} & {}^{\mathcal{F}}\{\boldsymbol{V}_{i2}\} & \cdots & {}^{\mathcal{F}}\{\boldsymbol{V}_{in}\} \end{bmatrix}$$

We elect to delete "$\mathcal{F}$" and simply understand that all vector's components are taken in some reference frame, thus we write Eq. (5.186) as

$$\{\dot{\boldsymbol{R}}_i\} = \left\{\frac{\boldsymbol{R}_i}{\partial t}\right\} + [V_i]\{\dot{\boldsymbol{q}}\} \qquad i = 1, 2, \ldots, N \tag{5.187}$$

The 3×3 partial velocity matrices $[V_i]$ are important, because they directly parameterize the system mass matrix, as will be seen. Note the vector version of kinetic energy, for a system of particles, is written as

$$T = \frac{1}{2}\sum_{i=1}^{N} m_i \dot{\boldsymbol{R}}_i \cdot \dot{\boldsymbol{R}}_i \tag{5.188}$$

The corresponding matrix expression for T is

$$T = \frac{1}{2}\sum_{i=1}^{N} m_i \{\dot{\boldsymbol{R}}_i\}^T \{\dot{\boldsymbol{R}}_i\} \tag{5.189}$$

Substitution of Eqs. (5.187) into Eq. (5.189) gives

$$T_2 = \frac{1}{2}\dot{\boldsymbol{q}}^T [M]\dot{\boldsymbol{q}} \tag{5.190}$$

and

$$T = \frac{1}{2}\dot{\boldsymbol{q}}^T[M]\dot{\boldsymbol{q}} + T_1 + T_0 \tag{5.191}$$

where the symmetric positive definite system mass matrix $[M] = [M]^T$ is defined as an explicit function of the $[V_i]$ matrices:

$$[M(\boldsymbol{q})] = \sum_{i=1}^{N} m_i [V_i]^T [V_i] \tag{5.192}$$

The remaining two energy components T_1 and T_0 are defined as

$$T_1 = \sum_{i=1}^{N} \left\{ \frac{\partial \boldsymbol{R}_i}{\partial t} \right\}^T [V_i]\dot{\boldsymbol{q}} \qquad \text{(linear in } \dot{\boldsymbol{q}}\text{)} \tag{5.193}$$

$$T_0 = \frac{1}{2}\sum_{i=1}^{N} \left\{ \frac{\partial \boldsymbol{R}_i}{\partial t} \right\}^T \left\{ \frac{\partial \boldsymbol{R}_i}{\partial t} \right\} \qquad \text{(does not contain } \dot{\boldsymbol{q}}\text{)} \tag{5.194}$$

Substitution of Eqs. (5.191) – (5.194) into Eq. (5.182) gives

$$\begin{aligned} \frac{d}{dt}\left([M]\dot{\boldsymbol{q}} + \sum_{i=1}^{N} \left\{ \frac{\partial \boldsymbol{R}_i}{\partial t} \right\}^T [V_i] \right) \\ - \operatorname{col}\left(\frac{1}{2}\dot{\boldsymbol{q}}^T \left[\frac{\partial M}{\partial q_1} \right] \dot{\boldsymbol{q}}, \frac{1}{2}\dot{\boldsymbol{q}}^T \left[\frac{\partial M}{\partial q_2} \right] \dot{\boldsymbol{q}}, \ldots, \frac{1}{2}\dot{\boldsymbol{q}}^T \left[\frac{\partial M}{\partial q_n} \right] \dot{\boldsymbol{q}} \right) \\ + \frac{\partial T_1}{\partial \boldsymbol{q}} + \frac{\partial T_0}{\partial \boldsymbol{q}} + \frac{\partial V}{\partial \boldsymbol{q}} = \boldsymbol{Q}_{nc} + [A]^T\boldsymbol{\lambda} \end{aligned}$$

which is rearranged to the form

$$[M]\ddot{\boldsymbol{q}} + \boldsymbol{G}(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \boldsymbol{Q}_{nc} + [A]^T\boldsymbol{\lambda} \tag{5.195}$$

$$[A]\dot{\boldsymbol{q}} + \boldsymbol{B} = 0 \tag{5.196}$$

where

$$\begin{aligned} \boldsymbol{G}(\boldsymbol{q}, \dot{\boldsymbol{q}}) = [\dot{M}]\dot{\boldsymbol{q}} + \frac{d}{dt}\left(\sum_{i=1}^{N} \left\{ \frac{\partial \boldsymbol{R}_i}{\partial t} \right\}^T [V_i] \right) \\ - \operatorname{col}\left(\frac{1}{2}\dot{\boldsymbol{q}}^T \left[\frac{\partial M}{\partial q_1} \right] \dot{\boldsymbol{q}}, \frac{1}{2}\dot{\boldsymbol{q}}^T \left[\frac{\partial M}{\partial q_2} \right] \dot{\boldsymbol{q}}, \ldots, \frac{1}{2}\dot{\boldsymbol{q}}^T \left[\frac{\partial M}{\partial q_n} \right] \dot{\boldsymbol{q}} \right) \\ - \frac{\partial}{\partial \boldsymbol{q}}(T_0 + T_1 - V) \end{aligned} \tag{5.197}$$

Finally, the nonlinear term can be simplified to the form

$$\boldsymbol{G}(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \tag{5.198}$$

Example 5.11: Will do the five bar mechanism here

Problems

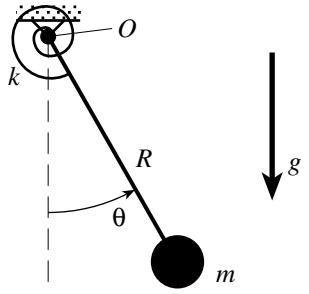


Figure 5.13: Illustration of a Pendulum with a Linear Torsional Spring

5.1

Consider the system shown in Figure 5.13. The pendulum has a weightless rod of length R. In addition to gravity, consider the torsional linear spring (moment $= -k\theta$, spring potential energy $= \frac{1}{2}k\theta^2$). Your tasks are to derive the equations of motion using:

a) Newton's laws

b) Euler's Equations of motion ($\boldsymbol{L} = \dot{\boldsymbol{H}}$), taking moment about pivot point O.

c) A version of Lagrange's equations

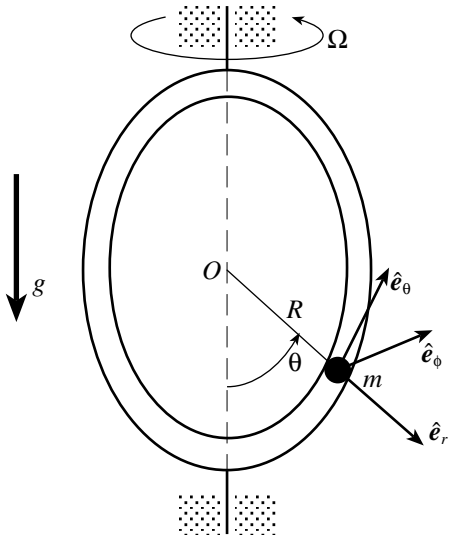


Figure 5.14: Illustration of a Particle Sliding Inside a Frictionless Rotating Circular Tube.

5.2

Consider the particle of mass m sliding with out friction in a rotating circular tube as shown in Figure 5.14. The angular rate $\dot{\phi} = \Omega$ is constant.

a) Formulate the "minimal coordinate" version of the kinetic energy $T(\theta, \dot{\theta})$ and potential energy $V(\theta)$ and derive the equations of motion using Lagrange's method. Find all equilibrium points θ_{eq} for which $\ddot{\theta} = \dot{\theta} = 0$. Investigate the stability of all points as a function of Ω over the range $0 \leq \Omega \leq \infty$. Verify that $\Omega = \sqrt{g/r}$ is a critical frequency where stability properties change.

b) Formulate kinetic and potential energy as a function of the redundant coordinates (r, ϕ, θ)

$$T = T(r, \phi, \theta, \dot{r}, \dot{\phi}, \dot{\theta}) \qquad V = V(r, \phi, \theta)$$

Use the Lagrange multiplier rule and Lagrange's equations in the form

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j}\right) - \frac{\partial \mathcal{L}}{\partial q_j} = Q_j + \sum_{k=1}^{2} \lambda_k \frac{\partial \psi_k}{\partial q_j}$$

with the generalized coordinates

$$(q_1, q_2, q_3) = (r, \phi, \theta)$$

and the holonomic constraints

$$\psi_1(r, \phi, \theta) = r - R = 0$$
$$\psi_2(r, \phi, \theta) = \phi - \Omega t = 0$$

Derive the system equations, determine the Lagrange multipliers (λ_1, λ_2) as functions of $(\theta, \dot{\theta}, \Omega, R, m)$ and eliminate them to verify the equation of motion of a).

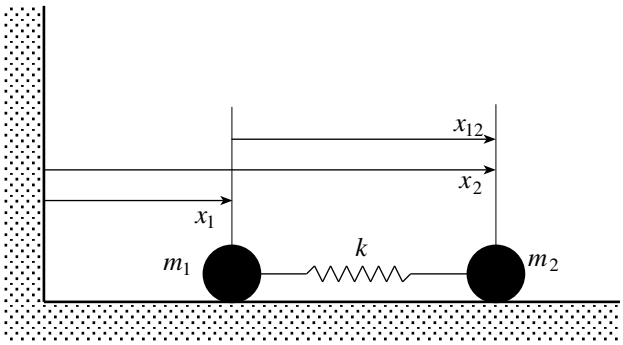


Figure 5.15: Illustration of a Spring Connected Two-Particle System.

5.3

Consider the two-particle system shown in Figure 5.15. The particles move along a straight line on a frictionless plane. The un-stretched spring length is d, so that the force acting on m_1 is $k[(x_2 - x_1) - d] = k(x_{12} - d)$. Consider two sets of generalized coordinates:

$$\begin{aligned} \text{Set I:} \quad & (q_1, q_2) \equiv (x_1, x_2) \\ \text{Set II:} \quad & (q_1, q_2) \equiv (x_c, x_{12}) \end{aligned}$$

with $x_c = \frac{1}{m_1+m_2}(m_1 x_1 + m_2 x_2)$. Your tasks are to:

a) Formulate the equations of motion using

i) Newton's Laws

ii) Lagrange's Equations

b) Starting with the equation of motion for Set I arranged in the matrix form

$$[M] \begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{pmatrix} = \boldsymbol{F}$$

First derive the 2×2 constant transformation matrix $[A]$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = [A] \begin{pmatrix} x_c \\ x_{12} \end{pmatrix}$$

and verify that the part a.i) differential equations are obtained from

$$[A]^T [M][A] \begin{pmatrix} \ddot{x}_c \\ \ddot{x}_{12} \end{pmatrix} = [A]^T \boldsymbol{F}$$

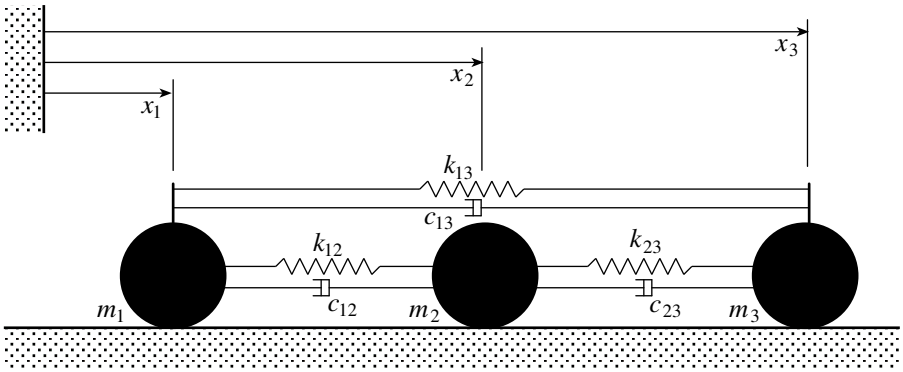


Figure 5.16: Illustration of a Spring and Dashpot Connected Three-Particle System.

5.4

Consider the three-particle system shown in Figure 5.16. The particles move along a straight line on a frictionless inertially fixed plane. The un-stretched lengths of the linear springs are d_{ij} so that the force acting on m_1 is

$$k_{12}\left[(x_2 - x_1) - d_{12}\right] + c_{12}(\dot{x}_2 - \dot{x}_1) + k_{13}\left[(x_3 - x_1) - d_{13}\right] + c_{13}(\dot{x}_3 - \dot{x}_1)$$

Consider the following three sets of generalized coordinates:

$$\begin{aligned} \text{Set I:} \quad & (q_1, q_2, q_3) \equiv (x_1, x_2, x_3) \\ \text{Set II:} \quad & (q_1, q_2, q_3) \equiv (x_{12}, x_{23}, x_3) \\ \text{Set III :} \quad & (q_1, q_2, q_3) \equiv (x_{12}, x_{23}, x_c) \end{aligned}$$

with $x_{ij} = x_j - x_i$, $x_c = \frac{1}{m}(m_1 x_1 + m_2 x_2 + m_3 x_3)$ and $m = m_1 + m_2 + m_3$. Your tasks are as follows:

a) Formulate the equations of motion for each set of generalized coordinates using

   i) Newtons Laws

   ii) Lagrange's Equations

b) Starting with the differential equations of motion derived from the Set I written in the matrix form

$$[M]\ddot{\boldsymbol{x}} = \boldsymbol{F}(\boldsymbol{x}) \qquad \boldsymbol{x} = (x_1, x_2, x_3)^T$$

first derive the 3×3 transformation matrices $[A]$ and $[B]$ such that

$$\begin{aligned} \boldsymbol{x} &= [A]\boldsymbol{y}\boldsymbol{y} = (x_{12}, x_{23}, x_3)^T \\ \boldsymbol{x} &= [B]\boldsymbol{z}\boldsymbol{z} = (x_{12}, x_{23}, x_c)^T \end{aligned}$$

Then introducing $\boldsymbol{x} = [A]\boldsymbol{y}$ and $\boldsymbol{x} = [B]\boldsymbol{z}$ into $[M]\ddot{\boldsymbol{x}} = \boldsymbol{F}(\boldsymbol{x})$ find the transformed differential equations

$$\begin{aligned} ([A]^T[M][A])\ddot{\boldsymbol{y}} &= [A]^T\boldsymbol{F}([A]\boldsymbol{y}) \\ ([B]^T[M][B])\ddot{\boldsymbol{z}} &= [B]^T\boldsymbol{F}([B]\boldsymbol{z}) \end{aligned}$$

and verify that these are identical to the results for coordinate sets II and II in part a).

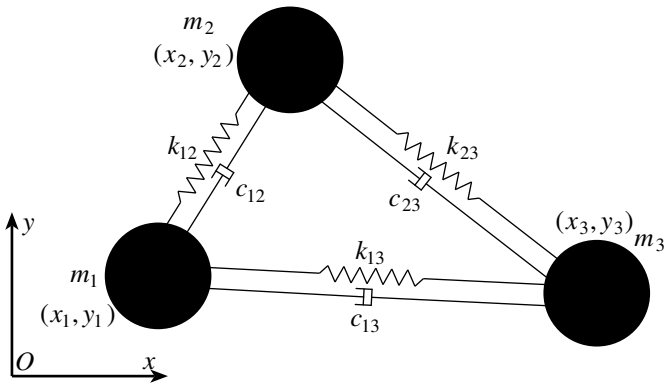


Figure 5.17: Illustration of a Planar Spring and Dashpot Connected Three-Particle System.

5.5 Consider a three-particle system moving on a frictionless, inertially fixed plane as shown in Figure 5.17. Generalize the results of problem 5.4, with the three coordinate choices

$$\begin{aligned}
\text{Set I:} \quad & \boldsymbol{X}^T \equiv \begin{pmatrix} x_1 & y_1 & x_2 & y_2 & x_3 & y_3 \end{pmatrix} \\
\text{Set II:} \quad & \boldsymbol{Y}^T \equiv \begin{pmatrix} x_{12} & y_{12} & x_{23} & y_{23} & x_3 & y_3 \end{pmatrix} \\
\text{Set III :} \quad & \boldsymbol{Z}^T \equiv \begin{pmatrix} x_{12} & y_{12} & x_{23} & y_{23} & x_x & y_x \end{pmatrix}
\end{aligned}$$

where

$$\begin{aligned}
x_{ij} &= x_j - x_i \qquad\qquad y_{ij} = y_j - y_i \\
x_c &= \frac{1}{m}\left(m_1 x_1 + m_2 x_2 + m_3 x_3\right) \\
y_c &= \frac{1}{m}\left(m_1 y_1 + m_2 y_2 + m_3 y_3\right)
\end{aligned}$$

Verify that setting either $y_i \equiv 0$ or $x_i \equiv 0$ one can obtain the differential equations found in problem 5.1.

Bibliography

[1] Kane, T. R. and Levinson, D. A., *Dynamics: Theory and Applications*, McGraw-Hill, Inc., New York, 1985.

[2] Moon, F. S., *Applied Dynamics*, Wiley Interscience, New York, 1998.

[3] Junkins, J. L. and Kim, Y., *Introduction to Dynamics and Control of Flexible Structures*, AIAA Education Series, Washington D.C., 1993.

Chapter Six

Advanced Methods of Analytical Dynamics

text here

6.1 The Hamiltonian Function

text

6.1.1 Some Special Properties of The Hamiltonian

text

6.1.2 Relationship of the Hamiltonian to Total Energy and Work Energy

text

6.1.3 Hamilton's Canonical Equations

text

6.1.4 Hamilton's Principal Function and the Hamilton-Jacobi Equation

text

6.2 Hamilton's Principles

text

6.2.1 Variational Calculus Fundamentals

text

6.2.2 Path Variations versus Virtual Displacements

text

6.2.3 Hamilton's Principles from D'Alembert's Principle

text

6.3 Dynamics of Distributed Parameter Systems

text

6.3.1 Elementary DPS: Newton-Euler Methods

text

6.3.2 Energy Functions for Elastic Rods and Beams

text

6.3.3 Hamilton's Principle Applied for DPS

text

6.3.4 Generalized Lagrange's Equations for Multi-Body DPS

text

Problems

6.1 text here

Chapter Seven

Nonlinear Spacecraft Stability and Control

CONSIDER a spacecraft which is to be reoriented to a new heading. It is possible to prescribe a corresponding trajectory and then derive the required control effort through inverse dynamics which will accomplish the desired maneuver. Such maneuvers are called open-loop maneuvers since no position or velocity feedback is present to indicate how accurately these maneuvers are being accomplished. Naturally, any real system will not be modeled perfectly and unmodeled dynamics and external influences will cause the spacecraft to drift from the desired trajectory or final state; i.e., the inverse solution for the open-loop contains modeling approximations. To guarantee stability of the maneuver, a feedback control law is required. This control law operates on measured updates of the current states and compares them to the where the spacecraft should be at any instant of time. The state errors are then used to modify the control input such that the spacecraft returns to the desired trajectory. Open-loop reference maneuvers are not always required to perform reorientations. In several control laws, spacecraft are reoriented to the new attitude by simply feeding the difference between current and final desired attitude to the control law.

This chapter will develop several control laws for both the regulator problem (maintaining a fixed orientation or configuration) and the trajectory tracking problem, and discuss their stability characteristics. Designing spacecraft attitude control laws combines the skills of rigid body kinematics and kinetics, as well as control methodology. In fact, the proper choice of attitude coordinates can be crucial to the usability of the resulting control law. If large, arbitrary rotations are to be performed, clearly any set of the Euler angle family would be a poor choice due to their small non-singular rotation range. Attitude control laws that make judicious use of various attitude coordinates will be presented.

7.1 Nonlinear Stability Analysis

To design the nonlinear control laws and study their stability, typically Lyapunov's direct method is used. This section will briefly discuss and review the basic concepts involved in nonlinear stability analysis. It is not intended as a complete study of nonlinear stability and control theory. Instead, our goal is to provide enough insight into the essence of Lyapunov stability theory to allow the reader to follow the developments of the attitude control laws and their stability proofs. The reader is assumed to already be familiar with basic linear control concepts. Representative references for linear control theory are references 1 and 2. For a more complete study of nonlinear stability analysis and control theory, the reader is referred to references 3–5.

7.1.1 Stability Definitions

Let $\boldsymbol{x}$ be a generalized state vector, then nonlinear dynamical systems can be written in the form

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, t) \tag{7.1}$$

If the function $\boldsymbol{f}(\boldsymbol{x}, t)$ does not explicitly depend on time, then the dynamical system is said to be *autonomous*. Otherwise the system is said to be *non-autonomous*. A spacecraft unfolding its solar panels at a prescribed rate would yield a non-autonomous dynamical system, since its inertia matrix would be time dependent. Let $\boldsymbol{u}$ be the autonomous feedback control

$$\boldsymbol{u} = \boldsymbol{g}(\boldsymbol{x}) \tag{7.2}$$

then the closed-loop dynamical system is given by

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}) \tag{7.3}$$

To define stability of a dynamical system, the notions of an equilibrium state $\boldsymbol{x}_e$ and nominal reference motion $\boldsymbol{x}_r$ are required.

Definitionn 7.1 (Equilibrium State) *A state vector point $\boldsymbol{x}_e$ is said to be an equilibrium state (or equilibrium point) of a dynamical system described by $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, t)$ at time t_0 if*

$$\boldsymbol{f}(\boldsymbol{x}_e, t) = 0 \qquad \forall\ t > t_0$$

Therefore, once the system reaches the state $\boldsymbol{x}_e$, it will remain there for all time. Equilibrium states can be thought of as the "natural states" of the system. Consider a free-swinging vertical pendulum. Its natural equilibrium state would be being hanging straight down at rest, or perhaps inverted.

Example 7.1: Let us evaluate the equilibrium states of a undamped spring-mass system with a nonlinear spring stiffness.

$$m\ddot{x} + k_1 x + k_2 x^3 = 0$$

The stiffnesses k_1 and k_2 are positive scalar constants. To write these equations in the form of Eq. (7.1), the state vector $\boldsymbol{x}$ is introduced.

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x \\ \dot{x} \end{pmatrix}$$

The dynamical system is then written as

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) = \begin{pmatrix} x_2 \\ -\frac{k_1}{m}x_1 - \frac{k_2}{m}x_1^3 \end{pmatrix}$$

To find all equilibrium states, the function $\boldsymbol{f}(\boldsymbol{x})$ is set equal to the zero vector. The first component immediately indicates that any equilibrium state of this system must have $x_2 = 0$. Setting the second component equal to zero yields three possible roots.

$$x_1 = 0 \quad \text{and} \quad x_1 = \pm\sqrt{-\frac{k_1}{k_2}}$$

Since solely real solutions are of interest for this spring mass system and $k_i > 0$, the only equilibrium state vector is found to be $\boldsymbol{x}_e = (0, 0)^T$.

If the dynamical system is to follow a prescribed motion, then this motion is referred to as the *nominal reference motion* $\boldsymbol{x}_r(t)$. To describe the proximity of one state to another, the notion of neighborhoods is defined.

Definitionn 7.2 (Neighborhood B_δ) *Given $\delta > 0$, a state vector $\boldsymbol{x}(t)$ is said to be in the neighborhood $B_\delta(\boldsymbol{x}_r(t))$ of the state $\boldsymbol{x}_r(t)$ if*

$$||\boldsymbol{x}(t) - \boldsymbol{x}_r(t)|| \ < \ \delta \qquad \Longrightarrow \qquad \boldsymbol{x}(t) \in B_\delta(\boldsymbol{x}_r(t))$$

Since the norm used in the Definition 7.2 is the standard Euclidean norm, neighborhoods can be visualized as being n-dimensional spherical regions (balls) of radius δ around a particular state $\boldsymbol{x}_r(t)$.

A simple form of stability is the concept of a motion simply being bounded (or Lagrange stable) relative to $\boldsymbol{x}_r(t)$. Note that $\boldsymbol{x}(t_0)$ could lie arbitrarily close to $\boldsymbol{x}_r(t_0)$ while $\boldsymbol{x}(t)$ may still deviate from $\boldsymbol{x}_r(t)$. The only stability guarantee made here is that this state vector difference will remain within a finite bound δ.

Definitionn 7.3 (Lagrange Stability) *The motion $\boldsymbol{x}(t)$ is said to be Lagrange stable (or bounded) relative to $\boldsymbol{x}_r(t)$ if there exists a $\delta > 0$ such that*

$$\boldsymbol{x}(t) \in B_\delta(\boldsymbol{x}_r(t)) \qquad \forall \ t > t_0$$

Declaring a motion to be Lyapunov stable (also referred to simply as being stable) is a stronger statement than saying it is Lagrange stable. With stability it is possible to keep the difference between $\boldsymbol{x}(t)$ and $\boldsymbol{x}_r(t)$ arbitrarily small. Let us first define stability relative to $\boldsymbol{x}_r(t)$. Stability relative to an equilibrium point is a special case of this more general setting.

Definitionn 7.4 (Lyapunov Stability) *The motion $\boldsymbol{x}(t)$ is said to be Lyapunov stable (or stable) relative to $\boldsymbol{x}_r(t)$ if for each $\epsilon > 0$ there exists a $\delta(\epsilon) > 0$ such that*

$$\boldsymbol{x}(t_0) \in B_\delta(\boldsymbol{x}_r(t_0)) \Longrightarrow \boldsymbol{x}(t) \in B_\epsilon(\boldsymbol{x}_r(t)) \qquad \forall\ t > t_0$$

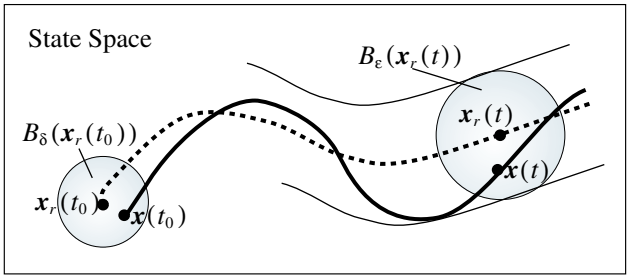


Figure 7.1: Illustration of Lyapunov Stability Definition

In other words, if the state vector $\boldsymbol{x}(t)$ is to remain within any arbitrarily small neighborhood B_ϵ of $\boldsymbol{x}_r(t)$, then there exists a corresponding initial neighborhood $B_\delta(\boldsymbol{x}_r(t_0))$ from which all $\boldsymbol{x}(t)$ must originate. This concept is illustrated in Figure 7.1. If $\boldsymbol{x}(t)$ is not stable, then it is said to be unstable. If the reference motion $\boldsymbol{x}_r(t)$ is an equilibrium state $\boldsymbol{x}_e$, then Definition 7.4 simplifies to:

Definitionn 7.5 (Lyapunov Stability) *The equilibrium state $\boldsymbol{x}_e$ is said to be Lyapunov stable (or stable) if for each $\epsilon > 0$ there exists a $\delta(\epsilon) > 0$ such that*

$$\boldsymbol{x}(t_0) \in B_\delta(\boldsymbol{x}_e) \Longrightarrow \boldsymbol{x}(t) \in B_\epsilon(\boldsymbol{x}_e) \qquad \forall\ t > t_0$$

These stability definitions only guarantee that the motion will remain arbitrarily close to the desired target state, provided that the initial state is close enough to the target state. Nothing is said whether or not the motion will actually converge to the target state.

Example 7.2: Let us analyze the equilibrium point stability of the simple spring-mass system

$$m\ddot{x} + kx = 0$$

using the Lyapunov stability definition. Writing the dynamical system in state space form we get

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & 0 \end{bmatrix} \boldsymbol{x} \tag{7.4}$$

where $\boldsymbol{x} = (x, \dot{x})^T$ and $\boldsymbol{x}_e = (0, 0)^T$. Solving the second order differential equations analytically, the trajectory $x(t)$ is found to be

$$x(t) = A\sin(\omega t + \phi)$$

where $\omega = \sqrt{k/m}$ is the natural frequency, ϕ is the phase angle and A is the oscillation amplitude. The spring stiffness k is assumed to be larger than the mass m, therefore $\omega > 1$. Note that both A and ϕ are determined through the initial conditions. Assume that the all initial states $\boldsymbol{x}(t_0)$ are in a neighborhood B_δ of $\boldsymbol{x}_e$, then

$$||\boldsymbol{x}(0)|| = \sqrt{x(0)^2 + \dot{x}(0)^2} = A\sqrt{1 + (\cos^2\phi)\,(\omega^2 - 1)}$$

Depending on the phase angle, the initial state magnitudes will vary between $A \leq ||\boldsymbol{x}(0)|| \leq A\omega$ since $\omega > 1$. Therefore, if all $||\boldsymbol{x}(0)|| \in B_\delta(\boldsymbol{x}_e)$, then

$$\delta = A\omega$$

The state vector magnitude for $t > 0$ is

$$||\boldsymbol{x}(t)|| = A\sqrt{1 + \cos^2(\omega t + \phi)(\omega^2 - 1)}$$

Since $\omega > 1$, this is bounded from above by

$$||\boldsymbol{x}(t)|| \leq A\omega = \delta$$

For a given $\epsilon > 0$, to guarantee that any trajectory $\boldsymbol{x}(t) \in B_\epsilon(\boldsymbol{x}_e)$, the initial neighborhood size δ must be chosen such that $\delta \leq \epsilon$. Note that in order to prove stability, it was necessary to solve for $x(t)$. For this simple linear system this was possible. However, for general nonlinear system this becomes exceedingly difficult.

A stronger stability statement is to say the motion $\boldsymbol{x}(t)$ is asymptotically stable. In this case the difference between $\boldsymbol{x}(t)$ and $\boldsymbol{x}_r(t)$ will approach zero over time.

Definitionn 7.6 (Asymptotic Stability) *The motion $\boldsymbol{x}(t)$ is asymptotically stable relative to $\boldsymbol{x}_r(t)$ if $\boldsymbol{x}(t)$ is Lyapunov stable and there exists a $\delta > 0$ such that*

$$\boldsymbol{x}(t_0) \in B_\delta(\boldsymbol{x}_r(t_0)) \Longrightarrow \lim_{t\to\infty} \boldsymbol{x}(t) = \boldsymbol{x}_r(t)$$

In other words, there exists a non-empty neighborhood of size δ around $\boldsymbol{x}_r(t_0)$ wherein each $\boldsymbol{x}(t_0)$ results in a motion that asymptotically approaches $\boldsymbol{x}_r(t)$. This result could again be simplified for the case of asymptotic stability of an equilibrium state by setting $\boldsymbol{x}_r(t) = \boldsymbol{x}_e$. Note that asymptotic stability only guarantees that the state error will approach zero, yet it does not predict any specific decay rate.

Definitionn 7.7 (Exponential Stability) *The motion $\boldsymbol{x}(t)$ is said to be exponentially stable relative to $\boldsymbol{x}_r(t)$ if $\boldsymbol{x}(t)$ is asymptotically stable and there exists a $\delta > 0$ and corresponding $\alpha(\delta) > 0$ and $\lambda(\delta) > 0$ such that*

$$\boldsymbol{x}(t_0) \in B_\delta(\boldsymbol{x}_r(t_0)) \Longrightarrow \; ||\boldsymbol{x}(t) - \boldsymbol{x}_r(t)|| \; \leq \; \alpha e^{-\lambda t}||\boldsymbol{x}(t_0) - \boldsymbol{x}_r(t_0)||$$

Therefore exponential stability guarantees that the state errors will decay *at least* at a rate λ. Of all the stability definitions presented, the concept of Lagrange stability is clearly the weakest, while exponential stability is the strongest statement. Unfortunately, proving exponential stability is also the most challenging.

Except for the Lagrange stability definition, all other types of stability defined are referred to as *local stability*. The initial state vector has to be within a certain neighborhood B_δ relative to the desired state vector for stability to be guaranteed. If stability is guaranteed *for any* initial state vector $\boldsymbol{x}(t_0)$, then the system is said to be *globally stable* or stable at large.

Definitionn 7.8 (Global Stability) *The motion $\boldsymbol{x}(t)$ is said to be globally stable (asymptotically stable or exponentially stable) relative to $\boldsymbol{x}_r(t)$ if $\boldsymbol{x}(t)$ is stable (asymptotically stable or exponentially stable) for any initial state vector $\boldsymbol{x}(t_0)$.*

7.1.2 Linearization of Dynamical Systems

For design purposes and to perform stability analysis, many nonlinear dynamical systems are linearized about a nominal reference motion $\boldsymbol{x}_r(t)$ which is defined through the differential equation $\dot{\boldsymbol{x}}_r = \boldsymbol{f}(\boldsymbol{x}_r, \boldsymbol{u}_r)$. This allows for standard linear control techniques and stability theory to be applied. Assume the dynamical system $\boldsymbol{x}(t)$ is defined through $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u})$. The control effort error $\delta\boldsymbol{u}$ is defined as

$$\delta\boldsymbol{u} = \boldsymbol{u} - \boldsymbol{u}_r \tag{7.5}$$

and the state error vector $\delta\boldsymbol{x}$ be defined as

$$\delta\boldsymbol{x} = \boldsymbol{x} - \boldsymbol{x}_r \tag{7.6}$$

The derivative of $\delta\boldsymbol{x}$ is written as

$$\delta\dot{\boldsymbol{x}} = \dot{\boldsymbol{x}} - \dot{\boldsymbol{x}}_r \tag{7.7}$$

Performing a Taylor series expansion of $\boldsymbol{x}$ about $(\boldsymbol{x}_r, \boldsymbol{u}_r)$ we obtain

$$\delta\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}_r, \boldsymbol{u}_r) + \frac{\partial \boldsymbol{f}(\boldsymbol{x}_r, \boldsymbol{u}_r)}{\partial \boldsymbol{x}}\delta\boldsymbol{x} + \frac{\partial \boldsymbol{f}(\boldsymbol{x}_r, \boldsymbol{u}_r)}{\partial \boldsymbol{u}}\delta\boldsymbol{u} + H.O.T - \boldsymbol{f}(\boldsymbol{x}_r, \boldsymbol{u}_r) \tag{7.8}$$

After dropping the higher order terms, this leads to the linearized dynamical system

$$\delta\dot{\boldsymbol{x}} \simeq \frac{\partial \boldsymbol{f}(\boldsymbol{x}_r, \boldsymbol{u}_r)}{\partial \boldsymbol{x}}\delta\boldsymbol{x} + \frac{\partial \boldsymbol{f}(\boldsymbol{x}_r, \boldsymbol{u}_r)}{\partial \boldsymbol{u}}\delta\boldsymbol{u} \tag{7.9}$$

Defining the two Jacobian matrices to be the time-varying matrix functions

$$[A] = \frac{\partial \boldsymbol{f}(\boldsymbol{x}_r, \boldsymbol{u}_r)}{\partial \boldsymbol{x}} \tag{7.10}$$

$$[B] = \frac{\partial \boldsymbol{f}(\boldsymbol{x}_r, \boldsymbol{u}_r)}{\partial \boldsymbol{u}} \tag{7.11}$$

the linearized system is written in the standard form

$$\delta\dot{\boldsymbol{x}} \simeq [A]\delta\boldsymbol{x} + [B]\delta\boldsymbol{u} \tag{7.12}$$

If the nominal reference motion $\boldsymbol{x}_r(t)$ is simply an equilibrium state $\boldsymbol{x}_e$, then the state vector $\boldsymbol{x}$ is typically expressed relative to $\boldsymbol{x}_e$. Therefore, having $\boldsymbol{x} = 0$ implies that the system is at the equilibrium point. The nominal control $\boldsymbol{u}_r$ is zero in this case. The linearized dynamical system about the equilibrium state $\boldsymbol{x}_e$ is expressed as

$$\dot{\boldsymbol{x}} \simeq [A]\boldsymbol{x} + [B]\boldsymbol{u} \tag{7.13}$$

Note that since the original nonlinear system $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u})$ was assumed to be autonomous, the generally time varying matrices $[A]$ and $[B]$ are constant matrices for this case.

Linear stability analysis can now be used on either Eq. (7.12) or (7.13). However, note that any stability claim resulting from this analysis will inherently only be a *local stability* claim. Any stable linear system is inherently globally exponentially stable. However, just because a *linearized* dynamical system is stable does not imply that the nonlinear system will be globally stable. Using linear stability theory on the linearized dynamical system only guarantees that there exists an non-empty neighborhood B_δ about the reference motion $\boldsymbol{x}_r(t_0)$ from which all nonlinear motions $\boldsymbol{x}(t)$ will be stable if $\boldsymbol{x}(t_0) \in B_\delta(\boldsymbol{x}_r(t_0))$.

Example 7.3: Let us find the linearized equations of motion of the nonlinear dynamical system

$$m\ddot{x} + c\dot{x} + k_1 x + k_2 x^3 = 0$$

about some reference motion $x_r(t)$. The given oscillator system has a cubically nonlinear spring with stiffness k_2. To write this second order differential equation in state space form, let us introduce the state vector $\boldsymbol{x}$ as

$$\boldsymbol{x} = \begin{pmatrix} x \\ \dot{x} \end{pmatrix}$$

The dynamical system is then written as the first order differential equation

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) = \begin{pmatrix} \dot{x} \\ -\frac{k_1}{m}x - \frac{k_2}{m}x^3 - \frac{c}{m}\dot{x} \end{pmatrix}$$

Using Eq. (7.10), the Jacobian matrix $[A]$ is found by taking the first partial derivative of $\boldsymbol{f}(\boldsymbol{x})$ with respect to the state vector $\boldsymbol{x}$.

$$[A] = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}} = \begin{bmatrix} 0 & 1 \\ -\frac{k_1}{m} - 3\frac{k_2}{m}x^2 & -\frac{c}{m} \end{bmatrix}$$

The linearized dynamical system about the reference motion $\boldsymbol{x}_r(t)$ is then expressed as

$$\delta\dot{\boldsymbol{x}} \simeq \begin{bmatrix} 0 & 1 \\ -\frac{k_1}{m} - 3\frac{k_2}{m}x_r^2 & -\frac{c}{m} \end{bmatrix} \delta\boldsymbol{x}$$

If the reference motion is simply $x_r(t) = 0$, then the linearized motion about the origin simplifies to

$$\delta\dot{\boldsymbol{x}} \simeq \begin{bmatrix} 0 & 1 \\ -\frac{k_1}{m} & -\frac{c}{m} \end{bmatrix} \delta\boldsymbol{x}$$

Note that the linearized system can be used to establish *local stability* guarantees. To verify that the nonlinear damped oscillator system is indeed globally stable further analysis is needed.

For zero control effort $\boldsymbol{u}$, the linearized system in Eq. (7.13) is stable if no eigenvalues of $[A]$ lie on the right half side of the complex plane (i.e. no eigenvalues with positive real parts). The system is said to be *strictly stable* if all eigenvalues have negative real parts. This guarantees that all states will decay to zero. The system is *marginally stable* if all eigenvalues are on the left half plane and at least one eigenvalue is purely imaginary. The modes corresponding to the imaginary eigenvalues will exhibit an oscillatory, non-decaying motion. The following theorem provides the conditions from which local nonlinear stability can be concluded from linear stability analysis.

Theorem 7.1 (Lyapunov's Linearization Method) *Assume the linearized dynamical system is found to be*

1. *strictly stable, then the nonlinear system is locally asymptotically stable.*
2. *unstable, the the nonlinear system is unstable.*
3. *marginally stable, then one cannot conclude anything about the stability of the nonlinear system without further analysis.*

This theorem is also referred to as *Lyapunov's indirect method*. The theorem makes intuitive sense. If the linearized system is either strictly stable or unstable, then one would expect that a neighborhood would exist where the nonlinear system would also be either stable or unstable. However, if the linearized system is only marginally stable, then the neglected second and higher order terms could render the nonlinear system either stable or unstable.

7.1.3 Lyapunov's Direct Method

Proving stability of nonlinear systems with the basic stability definitions and without resorting to local linear approximations can be quite tedious and difficult. Lyapunov's direct method provides a tool to make rigorous, analytical stability claims of nonlinear systems by studying the behavior of a scalar, energy-like Lyapunov function. A major benefit of this method is that this can be done without having to solve the nonlinear differential equations. To visualize the concept of Lyapunov's direct method, imagine a ball rolling down a steep U-shaped canyon. Having the ball roll down the center of the canyon is assumed to be the nominal reference motion. Initially, the ball is at rest half-way up

the smooth canyon wall. After release, the ball will roll down toward the valley center, overshoot and rise up on the other canyon wall. However, due to the conservation of energy, as long as the other wall is at least as high as the previous one, there is no way the ball can escape the canyon (be unstable). Instead it will roll down the canyon oscillating from wall to wall resulting in a locally stable motion. This example is only locally stable since the ball has to start within the canyon to guarantee stability. If friction drag effects are included in this study, then the oscillations will eventually dampen out and the ball motion will asymptotically track the canyon center. Here stability can be rigorously guaranteed by only looking at the total kinetic and potential energy of the system without having to actually solve for the resulting motion.

To mathematically create a virtual "canyon" around a target state $\boldsymbol{x}_r$, the concept of positive definite functions is important. These functions are zero at the target state (e.g., the canyon floor) and positive away from the target (e.g., the canyon walls).

Definitionn 7.9 (Positive (Negative) Definite Function) *A scalar continuous function $V(\boldsymbol{x})$ is said to be locally positive (negative) definite about $\boldsymbol{x}_r$ if*

$$\boldsymbol{x} = \boldsymbol{x}_r \implies V(\boldsymbol{x}) = 0$$

and there exists a $\delta > 0$ such that

$$\forall\ \boldsymbol{x} \in B_\delta(\boldsymbol{x}_r) \implies V(\boldsymbol{x}) > 0 \quad (V(\boldsymbol{x}) < 0)$$

excluding $\boldsymbol{x} = \boldsymbol{x}_r$. If the above property is true for any state vector $\boldsymbol{x}$, then $V(\boldsymbol{x})$ is said to be globally positive (negative) definite.

If a function is positive definite in a finite region around a target state, then this guarantees that this scalar function has a unique minimum at $\boldsymbol{x}_r$. Since dynamical systems naturally tend toward a state of minimum total energy, the fundamental importance of this method becomes clear. Similarly the concepts of negative definite and semi-definite functions can be defined. A function $V(\boldsymbol{x})$ is negative definite if $-V(\boldsymbol{x})$ is positive definite.

Definitionn 7.10 (Positive (Negative) Semi-Definite Function) *A scalar continuous function $V(\boldsymbol{x})$ is said to be locally positive (negative) semi-definite about $\boldsymbol{x}_r$ if*

$$\boldsymbol{x} = \boldsymbol{x}_r \implies V(\boldsymbol{x}) = 0$$

and there exists a $\delta > 0$ such that

$$\forall\ \boldsymbol{x} \in B_\delta(\boldsymbol{x}_r) \implies V(\boldsymbol{x}) \geq 0 \quad (V(\boldsymbol{x}) \leq 0)$$

excluding $\boldsymbol{x} = \boldsymbol{x}_r$. If the above property is true for any state vector $\boldsymbol{x}$, then $V(\boldsymbol{x})$ is said to be globally positive (negative) semi-definite.

If a function is only semi-definite, then this function may have extremas other than the desired target state. If the canyon had uneven walls with dips and rises, then it is possible for the ball to come to rest in such a "local valley" and not continue on to the center of the canyon. However, just because a function is only semi-definite near a reference state $\boldsymbol{x}_r$, does not guarantee that other local minima or maxima of $V(\boldsymbol{x})$ exists. A matrix $[K]$ is said to be positive or negative (semi-) definite if for every state vector $\boldsymbol{x}$

$$\boldsymbol{x}^T[K]\boldsymbol{x} \begin{cases} > 0 & \Rightarrow \text{positive definite} \\ \geq 0 & \Rightarrow \text{positive semi-definite} \\ < 0 & \Rightarrow \text{negative definite} \\ \leq 0 & \Rightarrow \text{negative semi-definite} \end{cases} \tag{7.14}$$

To prove stability of a dynamical system, special positive definite function called *Lyapunov functions* are sought.

Definitionn 7.11 (Lyapunov Function) *The scalar function $V(\boldsymbol{x})$ is a Lyapunov function for the dynamical system $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x})$ if it is continuous and there exists a $\delta > 0$ such that for any $\boldsymbol{x} \in B_\delta(\boldsymbol{x}_r)$*

1. *$V(\boldsymbol{x})$ is a positive definite function about $\boldsymbol{x}_r$*
2. *$V(\boldsymbol{x})$ has continuous partial derivatives*
3. *$\dot{V}(\boldsymbol{x})$ is negative semi-definite*

Even though $V(\boldsymbol{x})$ explicitly only depends on the state vector $\boldsymbol{x}$, since $\boldsymbol{x}(t)$ is time varying, the Lyapunov function V is time varying too. Using the chain rule, the derivative of V is found to be

$$\dot{V} = \frac{\partial V}{\partial \boldsymbol{x}}^T \dot{\boldsymbol{x}} = \frac{\partial V}{\partial \boldsymbol{x}}^T \boldsymbol{f}(\boldsymbol{x}) \tag{7.15}$$

where the last step holds since $\boldsymbol{x}(t)$ must satisfy the equations of motion $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x})$. Therefore the derivative $\dot{V}$ is often referred to as the directional derivative of V along the system trajectory. This idea is illustrated in Figure 7.2. The Lyapunov function is illustrated as a "bowl shaped" function over the state plane (x_1, x_2). If the projection of the motion $\boldsymbol{x}(t)$ onto the Lyapunov function $V(\boldsymbol{x})$ *always* has a non-positive slope, then V cannot grow larger and the corresponding dynamical system is stable about the origin. As will become evident in the sequel developments, the Lyapunov function, if one exists, is not unique – general stability in-the-large is often provable by any of a large family of Lyapunov functions. On the other hand, we will also see that the Lyapunov functions generalize the class of functions to which "total mechanical energy" belongs, and the simplest way to think qualitatively about Lyapunov functions is to simply view them as positive measures of displacement (in the state space) from a prescribed reference trajectory $\boldsymbol{x}_r(t)$.

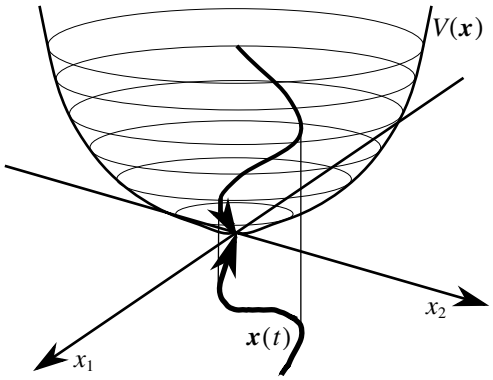


Figure 7.2: Illustration of a Lyapunov Function

If the following stability definitions are only valid in finite neighborhoods $B_\delta(\boldsymbol{x}_r)$, then they are only local stability theorems. However, if the Lyapunov function $V(\boldsymbol{x})$ is *radially unbounded* (i.e. $V(\boldsymbol{x}) \to \infty$ as $\|\boldsymbol{x}\| \to \infty$), then the stability claims are *globally valid*. To simplify the following theorems, it is assumed that the stability of $\boldsymbol{x}(t)$ is always examined relative the origin. If the origin is not the equilibrium state or nominal reference motion being examined, than a coordinate transformation can always be accomplished such that this is the case.

Theorem 7.2 (Lyapunov Stability) *If a Lyapunov function $V(\boldsymbol{x})$ exists for the dynamical system $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x})$, then this system is stable about the origin.*

Note that, as is the case with all Lyapunov stability theorems, if Theorem 7.2 is not fulfilled, then one *cannot conclude* that the system is unstable. In this case another Lyapunov function or stability theorem must be used to prove stability or instability. While using Lyapunov functions allows one to rigorously predict stability of nonlinear systems, finding an appropriate function to do so is not always a trivial matter. However, in many cases it is beneficial to first use the total energy expression as a first starting point for developing the Lyapunov function, as is done in the following example.

Example 7.4: The stability of the spring-mass system studied in Example 7.2 is verified here using Lyapunov stability theory. The dynamical system is given by

$$m\ddot{x} + kx = 0$$

The total kinetic and potential energy of this system provides a convenient Lyapunov function of the system motion about the system states $\dot{x} = 0$ and $x = 0$.

$$V(x, \dot{x}) = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}kx^2$$

By inspection $V(x, \dot{x})$ satisfies the three criteria of Definition 7.11. Note that $V(x, \dot{x})$ is radially unbounded and therefore any stability guarantee will be globally valid. Taking the derivative of V we find

$$\dot{V}(x, \dot{x}) = (m\ddot{x} + kx)\,\dot{x}$$

After substituting the dynamical system this is written as

$$\dot{V}(x, \dot{x}) = 0 \leq 0$$

Since $\dot{V}$ is negative semi-definite, the spring-mass system is only stable in the sense of Lyapunov, not asymptotically stable. Contrary to the more complicated stability proof in Example 7.2, this Lyapunov approach did not involve actually solving the equations of motion for $x(t)$, which could be a very challenging task to perform for many nonlinear systems. Since for this example, the Lyapunov function represents the total energy and its rate is shown to be zero for all time, we have also shown the well known truth that the total energy for the spring-mass system is conserved.

Theorem 7.3 (Asymptotic Stability) *Assume $V(\boldsymbol{x})$ is a Lyapunov function about $\boldsymbol{x}_r(t)$ for the dynamical system $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x})$, then the system is asymptotically stable if*

1. *the system is stable about $\boldsymbol{x}_r(t)$*
2. *$\dot{V}(\boldsymbol{x})$ is negative definite about $\boldsymbol{x}_r(t)$*

Theorem 7.4 (Exponential Stability) *Assume $V(\boldsymbol{x})$ is a Lyapunov function $V(\boldsymbol{x})$ of the dynamical system $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x})$ and the system is asymptotically stable, then the system is exponentially stable if there exists scalar constants $c_2 \geq c_1 > 0$ and $\lambda > 0$, $k > 0$ such that*

1. *$\dot{V} \leq -\lambda V$*
2. *$c_1 \|\boldsymbol{x}\|^k \leq V(\boldsymbol{x}) \leq c_2 \|\boldsymbol{x}\|^k$*

To guarantee convergence of $\boldsymbol{x}(t)$ to the reference motion $\boldsymbol{x}_r(t)$ (i.e. asymptotic stability), Theorem 7.3 states that a sufficient condition is $\dot{V} < 0$. However, this is only a sufficient, not a necessary condition. It is possible for a dynamical system to be asymptotically stable, while the Lyapunov function derivative along the system trajectory is only negative semi-definite. In essence, if $\dot{V}$ does vanish at some point other than $\boldsymbol{x}_r$, this point must not be an equilibrium state. The following very useful theorem allows one to prove asymptotic stability, if this indeed exists, when $\dot{V}(\boldsymbol{x}) \leq 0$ by investigating the higher order derivatives of the Lyapunov function.[6–8]

Theorem 7.5 *Assume there exists a Lyapunov function $V(\boldsymbol{x})$ of the dynamical system $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x})$. Let Ω be non-empty the set of state vectors such that*

$$\boldsymbol{x} \in \Omega \implies \dot{V}(\boldsymbol{x}) = 0$$

If the first $k-1$ derivatives of $V(\boldsymbol{x})$, evaluated on the set Ω, are zero

$$\frac{d^i V(\boldsymbol{x})}{d\boldsymbol{x}^i} = 0 \quad \forall\ \boldsymbol{x} \in \Omega \qquad i = 1, 2, \ldots, k-1$$

and the k-th derivative is negative definite on the set Ω

$$\frac{d^k V(\boldsymbol{x})}{d\boldsymbol{x}^k} < 0 \quad \forall\ \boldsymbol{x} \in \Omega$$

then the system $\boldsymbol{x}(t)$ is asymptotically stable if k is an odd number.

Example 7.5: To illustrate Theorems 7.3 through 7.5, the stability of the following linear spring-mass-damper system is studied.

$$m\ddot{x} + c\dot{x} + kx = 0$$

Again the total kinetic and potential energy is used as a radially unbounded Lyapunov function to measure the state errors from the equilibrium states $x = 0$ and $\dot{x} = 0$.

$$V(x, \dot{x}) = \frac{1}{2} m \dot{x}^2 + \frac{1}{2} k x^2$$

Taking the derivative of $V(x, \dot{x})$ and substituting the equations of motion, the following result is obtained:

$$\dot{V}(x, \dot{x}) = (m\ddot{x} + kx)\,\dot{x} = -c\dot{x}^2 \leq 0$$

Note that $\dot{V}$ is only negative semi-definite, and not negative definite. Even though we know from the easy analytical solution that this spring-mass-damper system is asymptotically stable, only Lyapunov stability can be concluded at this point in the analysis. Using Theorem 7.5, the higher order derivatives of V are investigated to prove asymptotic stability. The set of states Ω where $\dot{V} = 0$ is $\Omega = \{(x, \dot{x}) | \dot{x} = 0\}$. The second derivative of V is

$$\ddot{V} = -2c\ddot{x}\dot{x} = 2\frac{c}{m}\,(c\dot{x} + kx)\,\dot{x}$$

which is 0 when evaluated on the set Ω where $\dot{x} \equiv 0$. After taking another derivative and substituting the system equations of motion, the third derivative of V is expressed as

$$\dddot{V} = -2\frac{c}{m^2}\left((c\dot{x} + kx)^2 + c^2\dot{x}^2 + ckx\dot{x} - k\dot{x}^2\right)$$

Evaluated on the set Ω, this third Lyapunov derivative simplifies to

$$\dddot{V} = -2\frac{ck^2}{m^2}x^2$$

which is negative definite for for all x in Ω. Since the first non-zero higher order derivative of V is of odd order and negative definite on Ω, the system is globally asymptotically stable.

Note that the previous example does not prove that the given linear spring-mass-damper system is exponentially stable (which we know independently to be true from the analytical solution). Indeed with the chosen Lyapunov function it is impossible to satisfy the two primary conditions in Theorem 7.4. The following theorem allows Lyapunov theory to be conveniently used when proving various forms of stability of an autonomous linear system.

Theorem 7.6 (Lyapunov Stability Theorem for Linear Systems) *An autonomous linear system $\dot{\boldsymbol{x}} = [A]\boldsymbol{x}$ is stable if and only if for any symmetric, positive definite $[R]$ there exists a corresponding symmetric, positive definite $[P]$ such that*

$$[A]^T[P] + [P][A] = -[R] \tag{7.16}$$

is satisfied. Eq. (7.16) is called the algebraic Lyapunov equation.

The corollary to Theorem 7.6 is that if we know that $[A]$ is a stable matrix (i.e. all eigenvalues have negative real parts), then we know that for any choice of symmetric, positive definite $[P]$ we are guaranteed the existence of a corresponding symmetric, positive definite $[R]$. This property is very useful in the following example, and in fact, the proof follows from generalization of this example.

Example 7.6: Let us revisit the linear spring-mass-damper system from the previous example using an alternate Lyapunov function. First, we define the state vector $\boldsymbol{x}$ to be

$$\boldsymbol{x} = \begin{pmatrix} x \\ \dot{x} \end{pmatrix}$$

Now we are able to write the second order differential equation of motion in first order state space from.

$$\dot{\boldsymbol{x}} = \underbrace{\begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{c}{m} \end{bmatrix}}_{[A]} \boldsymbol{x}$$

To study the system stability about the fixed point $\boldsymbol{x} = 0$, we define the candidate Lyapunov function

$$V(\boldsymbol{x}) = \boldsymbol{x}^T[P]\boldsymbol{x}$$

where $[P]$ is some positive definite matrix. This causes $V(x)$ to be positive definite about the origin. Taking the derivative of $V(\boldsymbol{x})$ we find

$$\dot{V} = \boldsymbol{x}^T([A]^T[P] + [P][A])\boldsymbol{x}$$

Since $[A]$ is clearly a stable matrix for positive m, k and c, using Theorem 7.6 we are guaranteed that a symmetric, positive definite $[R]$ exists such that $[A]^T[P] + [P][A] = -[R]$. The Lyapunov rate is then rewritten as

$$\dot{V} = -\boldsymbol{x}^T[R]\boldsymbol{x}$$

which is negative definite in the state vector $\boldsymbol{x}$. By choosing this alternate Lyapunov function we are able to guarantee asymptotic stability in one step. This new V also lends itself to proof that all stable linear systems of the form $\dot{\boldsymbol{x}} = [A]\boldsymbol{x}$ are indeed exponentially stable. From linear algebra it is clear that the following Rayleigh-Ritz inequalities must hold for positive definite matrices:[1, 3]

$$\lambda_{P_{min}}||\boldsymbol{x}||^2 \leq \boldsymbol{x}[P]\boldsymbol{x} \leq \lambda_{P_{max}}||\boldsymbol{x}||^2$$
$$\lambda_{R_{min}}||\boldsymbol{x}||^2 \leq \boldsymbol{x}[R]\boldsymbol{x} \leq \lambda_{R_{max}}||\boldsymbol{x}||^2$$

where λ_{min} and λ_{max} are the respective smallest and largest eigenvalues of $[P]$ and $[R]$. If we chose λ such that

$$\lambda \leq \frac{\lambda_{R_{min}}}{\lambda_{P_{max}}}$$

then the first requirement of Theorem 7.6 is satisfied. To satisfy the second requirement, we chose $k = 2$ and $c_1 \leq \lambda_{P_{min}}$, $c_2 \geq \lambda_{R_{max}}$.

7.2 Generating Lyapunov Functions

Lyapunov's stability theory provides a very elegant method to guarantee stability characteristics of nonlinear dynamical systems without having to actually solve the corresponding equations of motion. Also, as will be evident, selection of Lyapunov functions can be approached simultaneously with control law designs. However, generating appropriate positive definite Lyapunov functions is not always a trivial matter. This section presents several Lyapunov functions that can be used to describe state errors of common aerospace systems. These functions are broken up into two categories, namely elemental Lyapunov functions that measure velocity or functions that measure position state errors. Separate elemental Lyapunov functions are linearly combined to provide the desired system Lyapunov function. The following motivational example illustrates how Lyapunov functions are used to generate control laws and make stability guarantees of the closed-loop system.

Example 7.7: We would like to have a particle m, whose position is given by the state vector $\boldsymbol{x}$, track a reference motion $\boldsymbol{x}_r(t)$. The dynamical system for this particle is given by Newton's second law.

$$m\ddot{\boldsymbol{x}} = \boldsymbol{u}$$

where the control vector $\boldsymbol{u}$ is the external force being applied to the particle. The tracking error $\delta\boldsymbol{x}$ is defined as

$$\delta\boldsymbol{x} = \boldsymbol{x} - \boldsymbol{x}_r$$

and the tracking error velocity as

$$\delta\dot{\boldsymbol{x}} = \dot{\boldsymbol{x}} - \dot{\boldsymbol{x}}_r$$

Having both $\delta\boldsymbol{x}$ and $\delta\dot{\boldsymbol{x}}$ go to zero implies that perfect tracking is being achieved. To develop an asymptotically stabilizing control law $\boldsymbol{u}$, the candidate Lyapunov function V is defined as the "error energy"

$$V(\delta\boldsymbol{x}, \delta\dot{\boldsymbol{x}}) = \frac{1}{2}m\delta\dot{\boldsymbol{x}}^T\delta\dot{\boldsymbol{x}} + \frac{1}{2}k\delta\boldsymbol{x}^T\delta\boldsymbol{x}$$

The first term of $V(\delta\boldsymbol{x}, \delta\dot{\boldsymbol{x}})$, which measures velocity errors, is a "kinetic-energy-error-like" term. If the reference velocity $\dot{\boldsymbol{x}}_r = 0$, then it would be the kinetic energy of the particle. The second term provides a positive definite measure of the position errors. It can be viewed as a "potential-energy-like" function. For example, the fictitious potential energy function used here is very similar in form to the real potential energy function of a linear spring. The parameter k can be thought of as a spring stiffness constant. Taking the first time derivative of V we find the "power" or work/energy equation

$$\dot{V} = \delta\dot{\boldsymbol{x}}^T\left(m\delta\ddot{\boldsymbol{x}} + k\delta\boldsymbol{x}\right)$$

For the closed loop system to be stable, Lyapunov stability theory requires that $\dot{V}$ be at least negative semi-definite. Therefore, we set $\dot{V}$ equal to

$$\dot{V} = -P\delta\dot{\boldsymbol{x}}^T\delta\dot{\boldsymbol{x}} \leq 0$$

where $P > 0$. Note that this $\dot{V}$ is not negative definite since it does not depend explicitly on the position error vector $\delta\boldsymbol{x}$. Setting these two Lyapunov function derivatives equal and using $\delta\ddot{\boldsymbol{x}} = \ddot{\boldsymbol{x}} - \ddot{\boldsymbol{x}}_r$ leads to the following stable closed-loop dynamical system.

$$m\ddot{\boldsymbol{x}} - m\ddot{\boldsymbol{x}}_r + k\delta\boldsymbol{x} + P\delta\dot{\boldsymbol{x}} = 0 \tag{7.17}$$

To find the control law $\boldsymbol{u}$ which will yield these dynamics, the system equations of motion are substituted into the closed-loop system, and we solve for the required control vector

$$\boldsymbol{u} = -k\delta\boldsymbol{x} - P\delta\dot{\boldsymbol{x}} + m\ddot{\boldsymbol{x}}_r$$

Note that the scalar parameters k and P are position and velocity feedback gains that provide stiffness and damping. To guarantee that this control law is indeed asymptotically stabilizing, the higher order time derivatives of V must be investigated. For this example $\dot{V}$ is zero whenever $\delta\dot{\boldsymbol{x}}$ is zero. The second (even) derivative of V is

$$\ddot{V} = -2P\delta\dot{\boldsymbol{x}}^T\delta\ddot{\boldsymbol{x}}$$

which is zero on the set $\Omega = \{(\delta\boldsymbol{x}, \delta\dot{\boldsymbol{x}}) | \delta\dot{\boldsymbol{x}} = 0\}$. The third (odd) derivative of V is

$$\dddot{V} = -2P\delta\ddot{\boldsymbol{x}}^T\delta\ddot{\boldsymbol{x}} - 2P\delta\dot{\boldsymbol{x}}^T\delta\dddot{\boldsymbol{x}}$$

Substituting the closed-loop dynamical system and setting $\delta\dot{\boldsymbol{x}}$ equal to zero yields

$$\dddot{V}(\delta\boldsymbol{x}, \delta\dot{\boldsymbol{x}} = 0) = -2P\frac{k^2}{m^2}\delta\boldsymbol{x}^T\delta\boldsymbol{x} < 0$$

which is negative definite. Since the first non-zero V derivative is of odd order, the control law $\boldsymbol{u}$ is actually asymptotically stabilizing.

This example illustrates the powerful fusion of control law design and system stability analysis with Lyapunov method. Finding a globally asymptotically stabilizing control law and proving the stability guarantees went hand-in-hand and were accomplished in a straight forward manner. A drawback to these Lyapunov methods is that the process of finding appropriate Lyapunov functions is not always obvious. The following two sections present several Lyapunov function prototypes that can be applied to many aerospace systems.

7.2.1 Elemental Velocity-Based Lyapunov Functions

We consider here a class of mechanical systems to provide physical motivation. We first consider the case that only the velocity of a dynamical system is to be controlled. Thus the state space of interest is simply $(\dot{\boldsymbol{q}})$ and not the classical $(\boldsymbol{q}, \dot{\boldsymbol{q}})$. The control vector here will be a force or torque type vector. We note the developments in this section seeks to rive $\dot{\boldsymbol{q}} \to 0$, but generally $\boldsymbol{q}$ will not be stabilized with input to a particular point. It is convenient to use the system kinetic energy expression T as the candidate Lyapunov function. For natural systems, the kinetic energy can always be written in the quadratic form

$$T = \frac{1}{2}\dot{\boldsymbol{q}}^T[M]\dot{\boldsymbol{q}} \tag{7.18}$$

where the vector $\boldsymbol{q}$ is a generalized position state vector. In general, the mass matrix $[M(\boldsymbol{q})]$ is positive definite and symmetric. The standard Lagrange equations of motion for a natural unconstrained system are

$$[M(\boldsymbol{q})]\ddot{\boldsymbol{q}} = -\dot{M}(\boldsymbol{q}, \dot{\boldsymbol{q}})]\dot{\boldsymbol{q}} + \frac{1}{2}\dot{\boldsymbol{q}}^T[M_{\boldsymbol{q}}(\boldsymbol{q})]\dot{\boldsymbol{q}} + \boldsymbol{Q} \tag{7.19}$$

where the vector $\boldsymbol{Q}$ is the generalized forcing term (includes both conservative and non-conservative forces). Note that the notation $[M_{\boldsymbol{q}}]$ indicates the partial derivative of the matrix $[M]$ with respect the vector $\boldsymbol{q}$ and the matrix product

$$\dot{\boldsymbol{q}}^T[M_{\boldsymbol{q}}(\boldsymbol{q})]\dot{\boldsymbol{q}} \equiv \begin{pmatrix} \dot{\boldsymbol{q}}^T\left[\frac{\partial M}{\partial q_1}\right]\dot{\boldsymbol{q}} \\ \vdots \\ \dot{\boldsymbol{q}}^T\left[\frac{\partial M}{\partial q_N}\right]\dot{\boldsymbol{q}} \end{pmatrix} \tag{7.20}$$

is an N-dimensional column vector.

If the reference velocity vector is the zero vector, then Eq. (7.18) itself provides a candidate Lyapunov function V.

$$V(\dot{\boldsymbol{q}}) = \frac{1}{2}\dot{\boldsymbol{q}}^T[M(\boldsymbol{q})]\dot{\boldsymbol{q}} \tag{7.21}$$

Since the mass matrix $[M]$ is symmetric, the derivative of V can be written as

$$\dot{V} = \dot{\boldsymbol{q}}^T[M]\ddot{\boldsymbol{q}} + \frac{1}{2}\dot{\boldsymbol{q}}^T[\dot{M}]\dot{\boldsymbol{q}} \tag{7.22}$$

Substituting for $[M]\ddot{\boldsymbol{q}}$ from the equations of motion in Eq. (7.19), $\dot{V}$ is reduced to

$$\dot{V} = \dot{\boldsymbol{q}}^T \left(-\frac{1}{2}[\dot{M}]\dot{\boldsymbol{q}} + \frac{1}{2}\dot{\boldsymbol{q}}^T[M_{\boldsymbol{q}}]\dot{\boldsymbol{q}} + \boldsymbol{Q} \right) \tag{7.23}$$

After noting the identity

$$\dot{\boldsymbol{q}}^T \left(\dot{\boldsymbol{q}}^T[M_{\boldsymbol{q}}]\dot{\boldsymbol{q}} \right) = \sum_{i=1}^{N} \dot{q}_i \left(\dot{\boldsymbol{q}}^T[M_{q_i}]\dot{\boldsymbol{q}} \right) = \dot{\boldsymbol{q}}^T[\dot{M}]\dot{\boldsymbol{q}} \tag{7.24}$$

the Lyapunov (kinetic energy) time derivative is written as the simple work-rate equation[9]

$$\dot{V} = \dot{\boldsymbol{q}}^T \boldsymbol{Q} \tag{7.25}$$

Suitable control vectors $\boldsymbol{Q}$ could now be developed to render $\dot{V}$ negative definite or negative semi-definite. For example, a simple control law would to set $\boldsymbol{Q} = -[P]\dot{\boldsymbol{q}}$ with $[P]$ being a positive definite matrix. This control law would asymptotically bring the system velocity $\dot{\boldsymbol{q}}$ of the system given in Eq. (7.19) to rest.

If the reference velocity vector $\dot{\boldsymbol{q}}_r$ is non-zero, then the Lyapunov function is defined in terms of the velocity state error vector $\delta\dot{\boldsymbol{q}} = \dot{\boldsymbol{q}} - \dot{\boldsymbol{q}}_r$ as the "kinetic-energy-like" function of departure velocity

$$V(\dot{\boldsymbol{q}}) = \frac{1}{2}\delta\dot{\boldsymbol{q}}^T[M(\boldsymbol{q})]\delta\dot{\boldsymbol{q}} \tag{7.26}$$

Taking the derivative of Eq. (7.26) we find

$$\dot{V} = \delta\dot{\boldsymbol{q}}^T \left([M]\delta\ddot{\boldsymbol{q}} + \frac{1}{2}[\dot{M}]\delta\dot{\boldsymbol{q}} \right) \tag{7.27}$$

After making use of the definition of $\delta\dot{\boldsymbol{q}}$ and substituting the equations of motion, $\dot{V}$ is written as

$$\dot{V} = \delta\dot{\boldsymbol{q}}^T \left(-\frac{1}{2}[\dot{M}] \left(\frac{1}{2}\dot{\boldsymbol{q}} + \dot{\boldsymbol{q}}_r \right) + \frac{1}{2}\dot{\boldsymbol{q}}^T[M_{\boldsymbol{q}}]\dot{\boldsymbol{q}} - [M]\ddot{\boldsymbol{q}}_r + \boldsymbol{Q} \right) \tag{7.28}$$

When tracking a time varying reference state, the elemental velocity-measure Lyapunov function rates no longer simplify to the classical power form of the work-energy equation in Eq. (7.25).

When controlling the angular velocities of rigid bodies, the elemental Lyapunov function in Eqs. (7.21) is specialized to kinetic energy expression of a rigid body given by

$$V(\boldsymbol{\omega}) = T = \frac{1}{2}\boldsymbol{\omega}^T[I]\boldsymbol{\omega} \tag{7.29}$$

where $[I]$ is the rigid body inertia matrix. Note that since the $\boldsymbol{\omega}$ and $[I]$ components are assumed to be taken in the body frame $\mathcal{B}$, taking the derivative of the scalar quantity V involves taking local derivatives of $\boldsymbol{\omega}$ and $[I]$ *as seen* by by $\mathcal{B}$ frame. Using Euler's rotational equations of motion in Eq. (4.32), the time derivative of the Lyapunov (kinetic energy) function is expressed as

$$\dot{V} = \boldsymbol{\omega}^T \left([I]\dot{\boldsymbol{\omega}}\right) = \boldsymbol{\omega}^T \left(-[\tilde{\boldsymbol{\omega}}][I]\boldsymbol{\omega} + \boldsymbol{Q}\right) = \boldsymbol{\omega}^T \boldsymbol{Q} \tag{7.30}$$

where $\boldsymbol{Q}$ is the total torque vector acting on the rigid body. The fact that $[I]$ is constant as seen by the $\mathcal{B}$ frame (i.e. rigid body) and that $\dot{\boldsymbol{\omega}} \equiv {}^{\mathcal{N}}d/dt(\boldsymbol{\omega}) = {}^{\mathcal{B}}d/dt(\boldsymbol{\omega})$ were used in deriving this expression.

To measure the angular velocity error relative to some reference rotation defined through $\boldsymbol{\omega}_r$, we define the angular velocity vector

$$\delta\boldsymbol{\omega} = \omega - \omega_r \tag{7.31}$$

Since both $\delta\boldsymbol{\omega}$ and $\boldsymbol{\omega}$ have components taken in the body frame $\mathcal{B}$, and the reference angular velocity vector $\boldsymbol{\omega}_r$ is typically given with the components taken in the reference frame $\mathcal{R}$, the angular velocity error ${}^{\mathcal{B}}\delta\boldsymbol{\omega}$ is computed as

$${}^{\mathcal{B}}\delta\boldsymbol{\omega} = {}^{\mathcal{B}}\boldsymbol{\omega} - [BR]{}^{\mathcal{R}}\boldsymbol{\omega}_r \tag{7.32}$$

The Lyapunov function is written as the kinetic energy like expression

$$V(\delta\boldsymbol{\omega}) = \frac{1}{2}\delta\boldsymbol{\omega}^T[I]\delta\boldsymbol{\omega} \tag{7.33}$$

The matrix components of $\delta\boldsymbol{\omega}$ and $[I]$ are implicitly taken in the $\mathcal{B}$ frame. Taking the derivative of the scalar quantity V involves taking derivatives of the scalar $\mathcal{B}$ frame components of Eq. (7.33).

$$\dot{V} = \delta\boldsymbol{\omega}^T[I]\frac{{}^{\mathcal{B}}d}{dt}\left(\delta\boldsymbol{\omega}\right) \tag{7.34}$$

Using the transport theorem in Eq. (1.21) and the identity in Eq. (4.30), the derivative of $\delta\boldsymbol{\omega}$ as seen by the $\mathcal{B}$ frame is given by

$$\frac{{}^{\mathcal{B}}d}{dt}\left(\delta\boldsymbol{\omega}\right) = \dot{\boldsymbol{\omega}} - \dot{\boldsymbol{\omega}}_r + \boldsymbol{\omega} \times \boldsymbol{\omega}_r \tag{7.35}$$

Substituting Eq. (7.35) and Euler's rotational equations of motion into the Lyapunov rate expression in Eq. (7.34), $\dot{V}$ is expressed as

$$\dot{V} = \delta\boldsymbol{\omega}^T \left(-[\tilde{\boldsymbol{\omega}}][I]\boldsymbol{\omega} + \boldsymbol{\omega} \times \boldsymbol{\omega}_r - [I]\dot{\boldsymbol{\omega}}_r + \boldsymbol{Q}\right) \tag{7.36}$$

Because most mechanical systems are natural systems, the Hamiltonian specializes for this case to the total system energy. This motivates the alternative use of the Hamiltonian as a more general Lyapunov function candidate.[10] Let

$\mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}})$ be the system Lagrangian. In the previous chapter the Hamiltonian H was defined in terms of the canonical coordinates $\boldsymbol{q}$ and $\boldsymbol{p}$ as

$$H(\boldsymbol{q}, \boldsymbol{p}) = \boldsymbol{p}^T \dot{\boldsymbol{q}} - \mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}}) \tag{7.37}$$

where $\boldsymbol{p}$ is the canonical (or conjugate) momentum defined

$$\boldsymbol{p} = \frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{q}}} \tag{7.38}$$

After some calculus and using Lagrange's equations, this leads to Hamilton's canonical equations of motion in terms of the gradient of H with respect to $(\boldsymbol{q}, \boldsymbol{p})$.

$$\dot{\boldsymbol{q}} = \frac{\partial H}{\partial \boldsymbol{p}} \tag{7.39}$$

$$\dot{\boldsymbol{p}} = \frac{\partial \mathcal{H}}{\partial \dot{\boldsymbol{q}}} + \boldsymbol{Q} \tag{7.40}$$

It is important in the partial derivatives of Eqs. (7.39) and (7.40) to consider $H = H(\boldsymbol{q}, \boldsymbol{p})$ rather than $H = H(\boldsymbol{q}, \dot{\boldsymbol{q}})$. The generalized coordinate rate vector $\dot{\boldsymbol{q}}$ has been eliminated by inverting Eq. (7.38) for $\boldsymbol{p} = \boldsymbol{g}(\boldsymbol{q}, \dot{\boldsymbol{q}})$. Taking the time derivative of the Hamiltonian H in Eq. (7.37) we find

$$\dot{H} = \dot{\boldsymbol{p}}^T \dot{\boldsymbol{q}} + \boldsymbol{p}\ddot{\boldsymbol{q}} - \frac{\partial \mathcal{L}}{\partial t} - \frac{\partial \mathcal{L}}{\partial \boldsymbol{q}}^T \dot{\boldsymbol{q}} - \frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{q}}}^T \ddot{\boldsymbol{q}} \tag{7.41}$$

After substituting Eqs. (7.38) and (7.40) and setting the Lyapunov function equal H, the Hamiltonian (Lyapunov) time rate is written as the modified power equation

$$\dot{V} = \dot{H} = \boldsymbol{Q}^T \dot{\boldsymbol{q}} - \frac{\partial \mathcal{L}}{\partial t} \tag{7.42}$$

Generally $\mathcal{L} = \mathcal{L}(t, \boldsymbol{q}, \dot{\boldsymbol{q}})$, but if $\mathcal{L} = \mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}})$ is not an explicit function of time, such as is the case with natural systems, the Hamiltonian rate reduces to the simple work-energy equation

$$\dot{V} = \dot{H}(\boldsymbol{q}, \boldsymbol{p}) = \boldsymbol{Q}^T \dot{\boldsymbol{q}} \tag{7.43}$$

Let the vector $\boldsymbol{F}$ be the external force being applied to a rigid body, $\boldsymbol{L}$ be the external torque vector, and $\boldsymbol{\omega}$ and $\boldsymbol{R}$ measure angular velocity and inertial position respectively, then $\dot{H}$ is written as

$$\dot{V} = \dot{H} = \boldsymbol{L}^T \boldsymbol{\omega} + \boldsymbol{F}^T \dot{\boldsymbol{R}} \tag{7.44}$$

Note that Eqs. (7.42) through (7.44) are *kinematic* results that were derived independent of the system dynamics! This has two important implications. First, when using the Hamiltonian (total energy) as the Lyapunov function of natural

system, it is not necessary to differentiate V explicitly and grind through the algebra to find an expression for $\dot{V}$. Instead, the work-energy rate expressions in Eqs. (7.43) and (7.44) can be written directly using some version of the work-energy equation. This can save a substantial amount of algebra and calculus. Second, any stabilizing control vector $\boldsymbol{Q}$ (that renders $\dot{V}$ negative semi-definite in some state space neighborhood) for the regulator control problem will remain stabilizing even in the presence of model errors. This is a direct consequence of the $\dot{V}$ expression being independent of the system dynamics (depends only upon forces, moments, and velocities of the points to which forces are applied. Therefore, if the inertia or mass matrix is modelled incorrectly, then the same control vector $\boldsymbol{Q}$ will still stabilize the system. The stability guarantees of such Lyapunov derived control laws, using total mechanical energy as the Lyapunov function, are thus *very robust* to the presence of modelling errors. Naturally the *performance* would differ in the system model was incorrect. It is implicitly necessary, however, that the actual system must be controllable and the actual Hamiltonian must be positive definite with respect to departures for the target state. Otherwise, it is necessary to establish sufficient insight to modify V and/or the number of control inputs. Whenever the actual system has additional degrees of freedom whose coordinates do not appear in the work/energy equation, this idealized analysis may break down, and caution showed be used to overstate stability guarantees.

Example 7.8: Assume the multi-link manipulator shown in Figure 7.3 is to be brought to rest. Choosing the inertial polar angles as generalized coordinates, the state vector is $\boldsymbol{q} = (\theta_1, \theta_2, \theta_3)^T$. The system mass matrix is then given by

$$M(\boldsymbol{q}) = \begin{bmatrix} (m_1+m_2+m_3)l_1^2 & (m_2+m_3)l_1 l_2 \cos(\theta_2-\theta_1) & m_3 l_1 l_3 \cos(\theta_3-\theta_1) \\ \cdots & (m_2+m_3)l_2^2 & m_3 2_1 l_3 \cos(\theta_3-\theta_2) \\ & m_3 l_2 l_3 \cos(\theta_3-\theta_2) & m_3 l_3^2 \end{bmatrix}$$

Assuming a torque Q_i is applied to each link, then the equations of motion for this 3-link manipulator system are given by

$$[M]\ddot{\boldsymbol{q}} + [\dot{M}]\dot{\boldsymbol{q}} - \frac{1}{2}\dot{\boldsymbol{q}}^T[M_{\boldsymbol{q}}]\dot{\boldsymbol{q}} = \boldsymbol{Q}$$

Since the final link orientation is not relevant in this velocity control situation, the Lyapunov function is simply chosen to be the total kinetic energy of the system given by

$$V(\dot{\boldsymbol{q}}) = \frac{1}{2}\dot{\boldsymbol{q}}^T[M(\boldsymbol{q})]\dot{\boldsymbol{q}}$$

Using Eq. (7.25), the Lyapunov rate is then given by

$$\dot{V} = \dot{\boldsymbol{q}}^T\boldsymbol{Q}$$

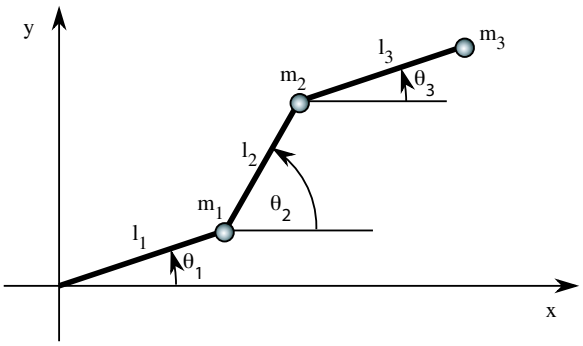


Figure 7.3: Three-Link Manipulator System Layout.

To guarantee that $\dot{V}$ is negative definite, standard control analysis leads to the velocity feedback control law $\boldsymbol{Q}_1$

$$\boldsymbol{Q}_1 = -P_1 \dot{\boldsymbol{q}}$$

where P_1 is a positive scalar velocity feedback gain. As shown in Refs. 11 and 12, since the mass matrix $[M]$ is symmetric positive-definite, using the velocity feedback control law $\boldsymbol{Q}_2$

$$\boldsymbol{Q}_2 = -P_2[M(\boldsymbol{q})]\dot{\boldsymbol{q}}$$

also leads to an asymptotically stable system with P_2 being a different positive scalar feedback gain. The configuration variable mass matrix $[M(\boldsymbol{q})]$ acts here as a variable feedback gains which produces some interesting and useful feedback performance enhancements. Note that $\boldsymbol{p} = [M(\boldsymbol{q})]\dot{\boldsymbol{q}} \equiv \frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{q}}}$ is the canonical conjugate momentum vector. One benefit of $\boldsymbol{Q}_2$ over $\boldsymbol{Q}_1$ is that $\boldsymbol{Q}_2$ can easily be shown to be exponentially stabilizing, providing the control designer with predictable exponential velocity error decay rate. Property 1 of Theorem 7.4 is trivially satisfied by setting $\lambda = 2P_2$.

$$\dot{V} = -P_2 \dot{\boldsymbol{q}}^T [M] \dot{\boldsymbol{q}} = -2P_2 V \leq -\lambda V$$

To verify property 2, we employ the Rayleigh-Ritz inequality[1, 3] which states

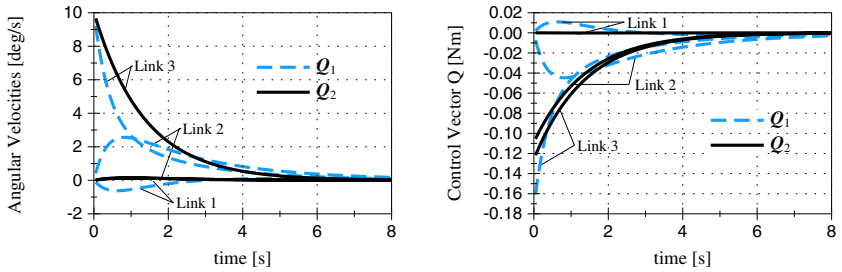
$$\lambda_{min}\left(\dot{\boldsymbol{q}}^T \dot{\boldsymbol{q}}\right) \leq \dot{\boldsymbol{q}}^T [M] \dot{\boldsymbol{q}} \leq \lambda_{max}\left(\dot{\boldsymbol{q}}^T \dot{\boldsymbol{q}}\right)$$

where λ_{min} and λ_{max} are respectively the smallest and largest eigenvalues of the system mass matrix $[M]$. Using the definition of the Lyapunov function, this inequality is rewritten as

$$2\lambda_{min}||\dot{\boldsymbol{q}}||^2 \leq V(\dot{\boldsymbol{q}}) \leq 2\lambda_{max}||\dot{\boldsymbol{q}}||^2$$

where the Euclidean norm is used. After setting $c_1 = 2\lambda_{min}$, $c_2 = 2\lambda_{max}$ and $k = 2$, the second property of Theorem 7.4 are verified and the control law $\boldsymbol{Q}_2$ is therefore exponentially stabilizing and $V(t) \leq V_0 e^{-\lambda t}$.

The following numerical simulation compares the performance of the two velocity feedback control laws $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$. The simulation parameters are



(i) $\dot{\boldsymbol{x}}$ Vector Components **(ii)** Control Vector $\boldsymbol{Q}$ Components

Figure 7.4: Isolated Initial Rotation Stabilization

Table 7.1: Parameters of Isolated Initial Motion Study

Parameter	Value	Units
l_i	1	m
m_i	1.0	kg
P_1	1.0	kg-m^2/sec
P_2	0.72	kg-m^2/sec
$\boldsymbol{x}(t_0)$	$[-90\ 30\ 0]$	deg
$\dot{\boldsymbol{x}}(t_0)$	[0.0 0.0 10]	deg/sec

given in Table 7.1. The feedback gains were chosen such that the maximum control torque encountered is the same for both control laws. The initial conditions are such that only the third link has some initial rotation. The other two links are at rest when the stabilizing control is turned on. The resulting motion for both control laws is shown in Figure 7.4. While $\boldsymbol{Q}_1$ is able to stabilize the system and bring all links to rest, the kinetic energy of the third link is partially transmitted to the other two links, thus exciting the entire system. However, the control law $\boldsymbol{Q}_2$ behaves quite differently as seen in Figure 7.4(i). The first two links remain essentially at rest while the the third link is brought to rest *separately*. This decoupling behavior was found will all chains of rigid links and is discussed in detail in References 11 and 12. The control torque components of each control law are shown in Figure 7.4(ii). While $\boldsymbol{Q}_1$ has all three torque motors active, $\boldsymbol{Q}_2$ only drives the second and third torque motors. For the same maximum allowable control torque, the $\boldsymbol{Q}_2$ was found to have much better state error convergence to zero in the end game.

7.2.2 Elemental Position-Based Lyapunov Functions

This section provides elemental position-based Lyapunov functions that allow us to control the position of a body. Analogous to the elemental velocity-based Lyapunov functions, the state space of interest here is simply $(\boldsymbol{q})$ and not the

more general $(\boldsymbol{q}, \dot{\boldsymbol{q}})$. Note that $\dot{\boldsymbol{q}}$ is treated as as the control variable. The control laws that are developed using purely position-based Lyapunov functions are often referred to as steering laws. The control law will determine a desired $\dot{\boldsymbol{q}}(t)$ trajectory that must be followed to stabilize the system about a desired position. To achieve this $\dot{\boldsymbol{q}}(t)$ time history, separate lower level servo loops are assumed to be present that will maintain the desired coordinate rate. For example, consider the multi-link system shown in Example 7.8. If a position based steering law were applied, then it would be assumed that servo loops were present on each joint link to maintain the $\dot{\boldsymbol{q}}(t)$ required by the steering law.

However, besides leading to system steering laws, the elemental position-based Lyapunov functions can also be combined with the elemental velocity-based Lyapunov functions to develop control laws that stabilize both velocity and position errors. This type of development is shown in the next section.

To provide a scalar measure of position displacements relative to a target state, potential energy-like functions are created which are zero at the target state. In many instances it is possible to use an actual mechanical potential energy function as the Lyapunov function. Consider a linear spring-mass system with the coordinate x measuring the displacement of the spring and the scalar parameter K being the spring constant. For this system, the spring potential function provides the positive definite measure of the displacement x.

$$V(x) = \frac{1}{2}kx^2 \tag{7.45}$$

The derivative of V is then

$$\dot{V} = \dot{x}\,(kx) \tag{7.46}$$

To combine the position-based Lyapunov functions later on with the velocity-based Lyapunov functions, it is usually be necessary to write the velocity expressions in the position-based Lyapunov functions in terms of the same velocity-coordinates used in the velocity-based Lyapunov function. With many dynamical systems it is not possible to express the position errors relative to some target state in terms of an actual potential energy function, because there may be no inherent "stiffness" that attracts the system to the desired state. Instead, a fictitious potential energy function is created which is zero at the target state and positive elsewhere (i.e. positive definite about the target state). A standard approach to create such a function is to express the position error as the weighted sum square of all position coordinates. This is written in matrix notation as

$$V(\boldsymbol{q}) = \frac{1}{2}\boldsymbol{q}^T[K]\boldsymbol{q} \tag{7.47}$$

where the position vector $\boldsymbol{q}$ is assumed to be measured relative to the target state. The symmetric matrix $[K]$ must be positive definite to guarantee that the Lyapunov function is a positive definite function of $\boldsymbol{q}$. When using these Lyapunov functions to create feedback control laws, the matrix $[K]$ assumes

the role of a position feedback gain matrix and also has the perfect analog to a system of linear springs. Since $[K] = [K]^T$, the derivative of Eq. (7.47) is

$$\dot{V} = \dot{\boldsymbol{q}}^T ([K]\boldsymbol{q}) \tag{7.48}$$

To create a steering law for $\boldsymbol{q}$, the Lyapunov function in Eq. (7.47) is written without the gain matrix $[K]$ as

$$V(\boldsymbol{q}) = \frac{1}{2}\boldsymbol{q}^T\boldsymbol{q} \tag{7.49}$$

Defining the steering $\dot{\boldsymbol{q}}$ to be

$$\dot{\boldsymbol{q}} = -[K]\boldsymbol{q} \tag{7.50}$$

the resulting $\dot{V} = -\boldsymbol{q}^T[K]\boldsymbol{q}$ is negative definite in $\boldsymbol{q}$. Thus this $\dot{\boldsymbol{q}}$ steering law would bring $\boldsymbol{q}$ asymptotically to zero. Note that in all steering laws (controlling $\boldsymbol{q}$ and treating $\dot{\boldsymbol{q}}$ as a control variable), the system dynamics do not appear. The internal servo control loops, which maintain the desired $\dot{\boldsymbol{q}}(t)$ coordinate rates, effectively hide the system level dynamics from the steering law. However, every system will exhibit certain limits as to how fast it is able to move and accelerate. Steering law gains must be carefully chosen such that the required $\dot{\boldsymbol{q}}(t)$ time histories to not exceed these limits. Otherwise the servo loops will not be able to track the required coordinate rates and the steering law stability guarantees are no longer valid without further analysis.

With the remaining position-based Lyapunov functions presented in this section, analogous steering laws could be constructed for each system. All these control laws demand a specific coordinate rate and assume a lower level system servo loop will achieve this desired rate. The advantage of using steering laws is that the control designer can focus the control on having the system states avoid singularities or other constraints. However, a drawback is that the internal servo control loops must run at a much higher digital sampling frequency to be able to track the desired coordinate rate time histories. This type of steering control is often used in robotics applications where a desired joint time history is prescribed, and each joint degree of freedom has a separate control servo loop which attempt to track the prescribed coordinate rates. The steering law can then be designed such that joint limits and singularities are not approached, while leaving the system level dynamics to be compensated for by the rate servo loop.

In rigid body dynamics, the kinetic energy is typically not expressed in terms of position coordinate derivatives (i.e. $\dot{q}_i$'s), but rather in terms of the body angular velocity vector $\boldsymbol{\omega}$. Therefore, velocity expressions in $\dot{V}$ for rigid bodies will need to be written in terms of $\boldsymbol{\omega}$ too. As was discussed in the chapter on rigid body kinematics, there is a multitude of attitude coordinates available to describe rigid body orientations. Convenient Lyapunov functions for a selected subset of the attitude coordinates discussed are presented below. A popular set of attitude coordinates is the Euler angle vector $\boldsymbol{\theta}$, where θ_i could

be either the 3-2-1 yaw, pitch and roll angles, the 3-1-3 Euler angles or any other set of sequential rotational coordinates. Assume $\boldsymbol{\theta}$ measures the current rigid body attitude relative to some target orientation, then the candidate Lyapunov function

$$V(\boldsymbol{\theta}) = \frac{1}{2}\boldsymbol{\theta}^T[K]\boldsymbol{\theta} \tag{7.51}$$

provides a positive definite measure of the attitude error. Using either Eq. (3.56) or (3.58) to express $\dot{\boldsymbol{\theta}} = [B(\boldsymbol{\theta})]\boldsymbol{\omega}$, the derivative of Eq. (7.51) is expressed as

$$\dot{V} = \boldsymbol{\omega}^T\left([B(\boldsymbol{\theta})][K]\boldsymbol{\theta}\right) \tag{7.52}$$

Contrary to the $\dot{V}$ in Eq. (7.48), the $\dot{V}$ in Eq. (7.52) will not lead to a linear feedback law in terms of the position/orientation coordinates due to the nonlinear nature of the $[B(\boldsymbol{\theta})]$ matrix. Also, If the target attitude is non-stationary, but defined through the reference body angular velocity vector $\boldsymbol{\omega}_r$, then the relative attitude error rate $\dot{\boldsymbol{\theta}}$ is given by

$$\dot{\boldsymbol{\theta}} = [B(\boldsymbol{\theta})]\delta\boldsymbol{\omega} \tag{7.53}$$

where $\boldsymbol{\theta}$ is the attitude vector from the reference frame to the body frame. The corresponding Lyapunov function time derivative is

$$\dot{V} = \delta\boldsymbol{\omega}^T\left([B(\boldsymbol{\theta})][K]\boldsymbol{\theta}\right) \tag{7.54}$$

For the remainder of this section, unless noted otherwise, it will always be assumed that the attitude vector is measured relative to the target state and not relative to some inertial frame. Therefore no distinction will be made if this reference state is stationary or not since the corresponding angular velocities $\boldsymbol{\omega}$ and $\delta\boldsymbol{\omega}$ can be interchanged trivially as shown in Eqs. (7.52) and (7.54).

The Gibbs or classical Rodrigues parameter vector $\boldsymbol{q}$ is another popular attitude coordinate vector used to describe large rotations. To establish a feedback control law with a fully populated positive definite feedback gain matrix $[K]$, a corresponding candidate Lyapunov function is expressed as

$$V(\boldsymbol{q}) = \boldsymbol{q}^T[K]\boldsymbol{q} \tag{7.55}$$

with the time derivative

$$\dot{V} = \boldsymbol{\omega}^T\left(\left(I - [\tilde{\boldsymbol{q}}] + \boldsymbol{q}\boldsymbol{q}^T\right)[K]\boldsymbol{q}\right) \tag{7.56}$$

If the feedback gain is permitted to be a scalar value K, then $\dot{V}$ in Eq. (7.56) is simplified to

$$\dot{V} = \boldsymbol{\omega}^T\left(K\left(1 + q^2\right)\boldsymbol{q}\right) \tag{7.57}$$

where the notation $q^2 = \boldsymbol{q}^T\boldsymbol{q}$ is used again. Due to the $(1+q^2)$ term, Eq. (7.57) leads to a nonlinear feedback control law in $\boldsymbol{q}$. This nonlinear scaling term can

be avoided by choosing a different Lyapunov function as was shown in Ref. 13. Instead of using the standard weighted sum square approach to generating an attitude Lyapunov function, a logarithmic sum squared approach is used.

$$V(\boldsymbol{q}) = K \ln\left(1 + \boldsymbol{q}^T\boldsymbol{q}\right) \tag{7.58}$$

Taking the derivative of Eq. (7.58) we find

$$\dot{V} = \frac{2K}{1+q^2}\dot{\boldsymbol{q}}^T\boldsymbol{q} \tag{7.59}$$

which, after substituting the Gibbs vector differential kinematic equation in Eq. (3.128), is reduced to the remarkably simply form

$$\dot{V} = \boldsymbol{\omega}^T\left(K\boldsymbol{q}\right) \tag{7.60}$$

Using Eq. (7.58), an attitude steering law Lyapunov function can be defined as

$$V(\boldsymbol{q}) = \ln\left(1 + \boldsymbol{q}^T\boldsymbol{q}\right) \tag{7.61}$$

Since the Lyapunov rate equation is given by the compact form

$$\dot{V} = \boldsymbol{\omega}^T\boldsymbol{q} \tag{7.62}$$

the steering law

$$\boldsymbol{\omega} = -[K]\boldsymbol{q} \tag{7.63}$$

would asymptotically drive the Gibbs attitude error $\boldsymbol{q}$ to zero if the gain matrix $[K]$ is positive definite. Once again, to implement this steering law, it is assumed that a lower level servo loop is present which would maintain the desired $\boldsymbol{\omega}(t)$ time histories.

Attitude Lyapunov functions in terms of MRPs are generated in a very similar way. For a fully populated feedback gain matrix $[K]$, the Lyapunov function

$$V(\boldsymbol{\sigma}) = 2\boldsymbol{\sigma}^T[K]\boldsymbol{\sigma} \tag{7.64}$$

can be used to measure the attitude error. Its derivative is of the complex form

$$\dot{V} = \boldsymbol{\omega}^T\left(\left(\left(1-\sigma^2\right)I - 2[\tilde{\boldsymbol{\sigma}}] + 2\boldsymbol{\sigma}\boldsymbol{\sigma}^T\right)[K]\boldsymbol{\sigma}\right) = \boldsymbol{\omega}^T[B(\boldsymbol{\sigma})][K]\boldsymbol{\sigma} \tag{7.65}$$

If the feedback gain is a scalar K, then $\dot{V}$ is simplified to

$$\dot{V} = \boldsymbol{\omega}^T\left(K(1+\sigma^2)\boldsymbol{\sigma}\right) \tag{7.66}$$

which will also lead to a nonlinear feedback control law in terms of $\boldsymbol{\sigma}$. To obtain a simpler $\dot{V}$ with a scalar K, the Lyapunov function is written as

$$V(\boldsymbol{\sigma}) = 2K\ln\left(1 + \boldsymbol{\sigma}^T\boldsymbol{\sigma}\right) \tag{7.67}$$

Note that by switching the MRP's to their alternate shadow set whenever $\sigma^2 > 1$, this Lyapunov function can describe *any* orientation error without encountering a singularity. Further, it provides a bounded measure of the attitude error. This is convenient, since two orientations can only differ by a *finite* rotation. Therefore, having a bounded Lyapunov function describing the attitude error inherently reflects this fact and will have some important consequences when designing attitude feedback control laws in terms of the MRPs. After substituting the MRP differential kinematic equations in Eq. (3.150), has the simple first time derivative

$$\dot{V} = \boldsymbol{\omega}^T (K\boldsymbol{\sigma}) \tag{7.68}$$

As was shown with the Gibbs vector steering law, the Lyapunov function $V(\boldsymbol{\sigma}) = 2\ln\left(1 + \boldsymbol{\sigma}^T\boldsymbol{\sigma}\right)$ allows us to show that the MRP steering law

$$\boldsymbol{\omega} = -[K]\boldsymbol{\sigma} \tag{7.69}$$

is asymptotically stabilizing. Note however that such steering laws are not typically applied to the rigid body attitude problem. Rather, the control law is formulated to provide a torque level input and control both the attitude and rotational rate. This development will be shown in the following section.

The most popular redundant, non-singular attitude coordinates are the Euler parameters β_i. The zero orientation vector is defined in terms of Euler parameters as

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{7.70}$$

With $\boldsymbol{\beta}$ being the orientation vector relative to the desired orientation, we define the candidate Lyapunov function V as

$$V(\boldsymbol{\beta}) = K\left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\right)^T \left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\right) \tag{7.71}$$

Using the Euler parameter differential equation in Eq. (3.105), and since $\hat{\boldsymbol{\beta}}$ is constant, the Lyapunov derivative is given by

$$\dot{V} = K\boldsymbol{\omega}^T[B(\boldsymbol{\beta})]^T \left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\right) \tag{7.72}$$

Making use of the identity $[B(\boldsymbol{\beta})]^T\boldsymbol{\beta} = 0$ in Eq. (3.107), the Lyapunov rate expression is reduced to the simple form

$$\dot{V} = K\boldsymbol{\omega}^T \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \boldsymbol{\omega}^T (K\boldsymbol{\epsilon}) \tag{7.73}$$

where the $\boldsymbol{\epsilon}$ definition in Eq. (3.109) is used.

Defining $V(\boldsymbol{\beta}) = \left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\right)^T \left(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}\right)$, the steering law

$$\boldsymbol{\omega} = -K\boldsymbol{\epsilon} \tag{7.74}$$

leads to the negative semi-definite Lyapunov rate $\dot{V} = -K\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}$. Note that $\dot{V}$ goes to zero whenever $\boldsymbol{\epsilon} \to 0$. However, having $\boldsymbol{\epsilon} = \mathbf{0}$ implies that β_0 is either +1 or -1. Thus at first glance it might appear as if the steering law would not always reorient the body to the desired orientation. Recalling that having $\beta_0 = \pm 1$ represents the same attitude (due to the duality of the Euler parameters $\boldsymbol{\beta}$ and $-\boldsymbol{\beta}$ for the same direction cosine matrix), the steering law in Eq. (7.74) will orient the body to the desired attitude. However, it is not guaranteed that the steering law will guide the body along the shortest rotational path to the desired orientation.

7.3 Nonlinear Feedback Control Laws

The elemental Lyapunov functions can be linearly combined to develop velocity and/or position feedback control laws for numerous mechanical aerospace systems. This section will develop and analyze in detail a reference trajectory tracking attitude and angular velocity feedback control law that will stabilize the rotation of a rigid body. Any set of attitude coordinates could be used describe the rigid body orientation. However, since large and arbitrary rotations must be considered, certain coordinates such as the Euler angles are less suited for large motion cases. A very popular set of coordinates used when performing large rotations are the Euler parameter β_i. Since they are nonsingular, a globally stable feedback control law in terms of β_i will be able to stabilize a body from any attitude error. Instead of using these well known redundant coordinates, the feedback control law can be developed in terms of the more recently developed modified Rodrigues parameters σ_i. We will adopt the σ_i's to illustrate the process. With only the minimal number of three coordinates, they achieve many similar properties as is accomplished with the Euler parameters. While the expression of the final feedback control law will depend on which attitude coordinates were chosen, the steps taken in the development of this control law holds for any choice of attitude coordinates.

7.3.1 Unconstrained Control Law

The modified Rodrigues parameter vector $\boldsymbol{\sigma}$ is very well suited for describing attitude errors in a feedback control law setting. Particularly when very large attitude errors are present, the MRPs are extremely attractive. By switching between the original and shadow MRP set they are able to describe any arbitrary orientation without encountering singularities by only using three parameters instead of four as do the Euler parameters. Adopting the switching surface $\sigma^2 = 1$ bounds the attitude error vector norm within the unit sphere $|\boldsymbol{\sigma}| \leq 1$ where the $\boldsymbol{\sigma}$-motion is approximately linear with respect to $\boldsymbol{\omega}$. This bounded attitude

error property is very useful since it will make designing the attitude feedback gain much easier. Choosing the $\sigma^2 = 1$ switching surface also has a big benefit when trying to bring a tumbling rigid body to rest. Conventional attitude parameters such as the Euler angles have no explicit means of determining the shortest rotational distance back to the reference attitude. Consider this one dimensional example. If a rigid body has tumbled past 180° from the reference attitude, it would be much simpler for the control law to just assist the body in completing the tumble and then bring it to rest as it approaches the reference attitude from the opposite direction, as opposed to "unwinding" the motion through a rotation of greater than 180°. Using MRPs with the $\sigma^2 = 1$ switching surface (in a feedback control law setting) provides a set of attitude coordinates that will naturally do just that. As is shown in the cases in Eq. (3.140), bounding the MRP vector to unit magnitude or less limits corresponding principal rotation angle Φ to be 180 degrees or less. In other words, these MRPs will always measure the shortest rotational error to the reference attitude, and control laws seeking to null $\boldsymbol{\sigma}$ will implicitly seek the shortest angular path to the target state. Of course the discontinuous switch of $\sigma^2 = 1$ is a cause for some concern, but these concerns typically turn out to have negligible practical consequences. Difficulties may arise if an unusual large external disturbance causes cyclic motion through this 180° condition.

Let $[I]$ be the rigid body inertia matrix, $\boldsymbol{\omega}(t)$ be the body angular velocity vector and $\boldsymbol{u}(t)$ be some *unconstrained* external torque vector. The vector $\boldsymbol{L}$ is some *known* external torque acting on the body. Euler's rotational equations of motion for a rigid body are given by[14]

$$[I]\dot{\boldsymbol{\omega}} = -[\tilde{\boldsymbol{\omega}}][I]\boldsymbol{\omega} + \boldsymbol{u} + \boldsymbol{L} \tag{7.75}$$

The vector $\boldsymbol{\sigma}(t)$ measures the attitude error of this rigid body to some reference trajectory which itself is defined through the reference angular velocity vector $\boldsymbol{\omega}_r(t)$. The error $\delta\boldsymbol{\omega}(t)$ in angular velocities is defined as

$$\delta\boldsymbol{\omega} = \boldsymbol{\omega} - \boldsymbol{\omega}_r \tag{7.76}$$

The MRP rate vector $\dot{\boldsymbol{\sigma}}$ and the body angular velocity error vector $\delta\boldsymbol{\omega}$ are then related through

$$\dot{\boldsymbol{\sigma}} = \frac{1}{4}\left[(1-\sigma^2)I + 2[\tilde{\boldsymbol{\sigma}}] + 2\boldsymbol{\sigma}\boldsymbol{\sigma}^T\right]\delta\boldsymbol{\omega} \tag{7.77}$$

Combining Eqs. (7.36) and (7.67) leads to the Lyapunov function[15–17]

$$V(\boldsymbol{\omega}, \boldsymbol{\sigma}) = \frac{1}{2}\delta\boldsymbol{\omega}^T[I]\delta\boldsymbol{\omega} + 2K\log\left(1 + \boldsymbol{\sigma}^T\boldsymbol{\sigma}\right) \tag{7.78}$$

which provides a positive definite, radially unbounded measure of the rigid body state error relative to the reference trajectory. The parameter K is a positive *scalar* attitude feedback gain. Making use of the Lyapunov rates in Eq. (7.34) and (7.68), the derivative of the Lyapunov function V is expressed as

$$\dot{V} = \delta\boldsymbol{\omega}^T\left([I]\frac{\mathcal{B}d}{dt}(\delta\boldsymbol{\omega}) + K\boldsymbol{\sigma}\right) \tag{7.79}$$

We write the derivative of $\delta\boldsymbol{\omega}$ as ${}^{\mathcal{B}}d/dt(\delta\boldsymbol{\omega})$ to clarify that $\delta\boldsymbol{\omega}$ is not treated here as a regular vector, but rather as a 3×1 matrix of scalar $\mathcal{B}$ frame components. To guarantee stability we force $\dot{V}$ to be negative semi-definite by setting it equal to

$$\dot{V} = -\delta\boldsymbol{\omega}^T[P]\delta\boldsymbol{\omega} \tag{7.80}$$

where $[P]$ is the positive definite angular velocity feedback gain matrix. This leads to the following stability constraint.

$$[I]\frac{{}^{\mathcal{B}}d}{dt}(\delta\boldsymbol{\omega}) + [P]\delta\boldsymbol{\omega} + K\boldsymbol{\sigma} = 0 \tag{7.81}$$

After making use of the local derivative expression of $\delta\boldsymbol{\omega}$ in the $\mathcal{B}$ frame given in Eq. (7.35) and substituting the rigid body dynamics in Eq. (7.75) into Eq. (7.81), the feedback control $\boldsymbol{u}$ is given by

$$\boldsymbol{u} = -K\boldsymbol{\sigma} - [P]\delta\boldsymbol{\omega} + [I](\dot{\boldsymbol{\omega}}_r - [\tilde{\boldsymbol{\omega}}]\boldsymbol{\omega}_r) + [\tilde{\boldsymbol{\omega}}][I]\boldsymbol{\omega} - \boldsymbol{L} \tag{7.82}$$

Since the Lyapunov function V in Eq. (7.78) is radially unbounded and the MRPs with the $\sigma^2 = 1$ switching surface are non-singular, the feedback control law $\boldsymbol{u}$ is guaranteed to be globally stabilizing. The cross-coupling term $\boldsymbol{\omega} \times \boldsymbol{\omega}_r$ is sometimes neglected in this feedback control law. For typical spacecraft maneuvers, both $\boldsymbol{\omega}$ and $\boldsymbol{\omega}_r$ are relatively small and this cross product is not scaled by any inertia components or feedback gains. Therefore this product of $\boldsymbol{\omega}$ and $\boldsymbol{\omega}_r$ usually has a negligible impact on the control law performance. However, to rigorously guarantee stability or to include the possibility of large $\boldsymbol{\omega}$ and $\boldsymbol{\omega}_r$ vectors, this cross-coupling term must be included.

Note that if the reference trajectory is a stationary attitude (i.e. $\boldsymbol{\omega}_r(t) = 0$), then the globally stabilizing feedback control law of Eqs. (7.82) simplifies to the elegant (linear in $\boldsymbol{\sigma}, \boldsymbol{\omega}$ law:[15, 16]

$$\boldsymbol{u} = -K\boldsymbol{\sigma} - [P]\boldsymbol{\omega} - \boldsymbol{L} \tag{7.83}$$

The last term in Eq. (7.82) is not needed here since for this case this gyroscopic term is non working ($\boldsymbol{\omega}^T[\tilde{\boldsymbol{\omega}}][I]\boldsymbol{\omega}$ is zero). Note that in either control law in Eq. (7.82) or (7.83), whenever the magnitude of the MRP vector $\boldsymbol{\sigma}$ grows larger than one and it is mapped to its shadow set inside the unit sphere, the sign of the MRP attitude feedback term will change. This corresponds to attitude error between the rigid body and the reference trajectory having having grown larger than 180^o. By switching the MRP vectors, the control law now no longer tries to force the rigid body to "unwind" back to the previous path, but rather it allows it to complete the revolution and seeks to bring the body to rest as it approaches the desired reference trajectory from the opposite direction. For a given choice on K and P, we can see in Eq. (7.83) that a large $\boldsymbol{L}$ may result in an un-achievable control if $\boldsymbol{u}$ is bounded. As we show below, this difficulty is compounded if $\boldsymbol{L}$ is unknown. Also note, at the precise instant of the switch near $\sigma^2 = 1$, the step from Eq. (7.81) to Eq. (7.82) involves

substituting for derivatives along assumed continuous trajectories. Unless a large $\boldsymbol{L}(t)$ causes the $\sigma^2 = 1$ surface to be frequently encountered, no practical difficulty is expected due to "chatter", but pathological problems can obviously be invented. Typically switches at $\sigma^2 = 1$ are rare events, since $\boldsymbol{\sigma}$ is an attitude error expected to be small.

If we had chosen to use the Euler parameters as our attitude coordinates and used elemental Lyapunov function in Eq. (7.71) instead of Eq. (7.67), then only the attitude feedback term in our control law would change as follows:

$$\boldsymbol{u} = -K\boldsymbol{\epsilon} - [P]\delta\boldsymbol{\omega} + [I]\dot{\boldsymbol{\omega}}_r - \boldsymbol{\omega} \times \boldsymbol{\omega}_r + [\tilde{\boldsymbol{\omega}}][I]\boldsymbol{\omega} - \boldsymbol{L} \tag{7.84}$$

The vector $\boldsymbol{\epsilon} = (\beta_1, \beta_2, \beta_3)^T$ is defined in Eq. (3.109). While both control laws in Eq. (7.82) and (7.84) are essentially singularity free, the Euler parameter control law control law will not automatically drive the body back to the reference trajectory through the shortest rotational path as was the case with the combined MRP and shadow-MRP control law. If a body nearly completes a full rotation, then the above Euler parameter control law will try to reverse this rotation, even though the body is already near the correct attitude. This problem can be avoided with a minor modification performed in a similar manner as was done with the MRPs. Since the Euler parameters too are non-unique, one can always switch the attitude description to the alternate set; in this case the transformation would be simply $\boldsymbol{\beta}' = -\boldsymbol{\beta}$ (or $\boldsymbol{\epsilon} = -\boldsymbol{\epsilon}$ for the shown Euler parameter feedback control law). By ensuring that the current β_i parameters have $\beta_0 \geq 0$ one is guaranteed that the vectors describes the shortest rotational distance back to the reference trajectory. Thus in both cases, a switch at the 180^o error condition is required to obtain this desired attitude control law property.

7.3.2 Asymptotic Stability Analysis

Since $\dot{V}$ in Eq. (7.80) is only negative semi-definite, it can only be concluded at this point that the control law $\boldsymbol{u}$ in Eq. (7.82) is globally stabilizing. To prove that it is indeed globally asymptotically stabilizing, the higher time derivatives of the Lyapunov function V can be investigated as indicated in Theorem 7.5. A sufficient condition to guarantee asymptotic stability is that the first nonzero higher-order derivative of V, evaluated on the set of states such that $\dot{V}$ is zero, must be of odd order and be negative definite.[6–8] To simplify the notation from here on in this chapter, it understood that the derivative expression $\delta\dot{\boldsymbol{\omega}}$ is actually ${}^{\mathcal{B}}d/dt(\delta\boldsymbol{\omega}_r)$. The same holds true for higher derivatives of $\delta\boldsymbol{\omega}$. For this dynamical system $\dot{V}$ is zero if $\delta\dot{\boldsymbol{\omega}}$ is zero. Differentiating Eq. (7.80) yields

$$\ddot{V} = -2\delta\boldsymbol{\omega}^T[P]\delta\dot{\boldsymbol{\omega}} \tag{7.85}$$

which is zero for the set where $\delta\boldsymbol{\omega}$ is zero. Differentiating again the third derivative of the Lyapunov function V is

$$\dddot{V} = -2\delta\boldsymbol{\omega}^T[P]\delta\ddot{\boldsymbol{\omega}} - 2\delta\dot{\boldsymbol{\omega}}^T[P]\delta\dot{\boldsymbol{\omega}} \tag{7.86}$$

Substituting Eq. (7.81) into Eq. (7.86) and setting $\delta\boldsymbol{\omega} = 0$, the third derivative of the Lyapunov function is expressed as

$$\dddot{V}(\boldsymbol{\sigma}, \delta\boldsymbol{\omega} = 0) = -K^2\boldsymbol{\sigma}^T \left([I]^{-1}\right)[P][I]\boldsymbol{\sigma} \tag{7.87}$$

which is a negative definite quantity since both $[I]$ and $[P]$ are positive definite matrices. Therefore the control law $\boldsymbol{u}$ in Eq. (7.82) is globally asymptotically stabilizing.

If some *unmodeled* external torque $\Delta\boldsymbol{L}$ is present, then Euler's rotational equations of motion are written as

$$[I]\dot{\boldsymbol{\omega}} = -[\tilde{\boldsymbol{\omega}}][I]\boldsymbol{\omega} + \boldsymbol{u} + \boldsymbol{L} + \Delta\boldsymbol{L} \tag{7.88}$$

Substituting these equations of motion into the Lyapunov rate expression in Eq. (7.79) and keeping the same feedback control law $\boldsymbol{u}$ given in Eq. (7.82) leads to the new Lyapunov rate expression

$$\dot{V} = -\delta\boldsymbol{\omega}^T[P]\delta\boldsymbol{\omega} + \delta\boldsymbol{\omega}^T\Delta\boldsymbol{L} \tag{7.89}$$

For a nonzero $\Delta\boldsymbol{L}$ vector, this $\dot{V}$ is not negative semi-definite and the control law $\boldsymbol{u}$ is no longer said to be globally stabilizing in the sense of Lyapunov. However, for a constant, bounded external torque vector $\Delta\boldsymbol{L}$, Eq. (7.89) shows that $\delta\boldsymbol{\omega}$ cannot become unstable and grow unbounded in magnitude. The reason for this is that since the expression $-\delta\boldsymbol{\omega}^T[P]\delta\boldsymbol{\omega}$ is negative semi-definite and $\Delta\boldsymbol{L}$ is constant, the first term in the $\dot{V}$ expression is guaranteed to become negative and dominant as $\delta\omega$ grows in magnitude. As soon as this happens the Lyapunov function, which is a measure of the state errors, will decay again and the angular velocities will not grow unbounded. The new closed-loop equations of motion are written as

$$[I]\delta\dot{\boldsymbol{\omega}} + [P]\delta\boldsymbol{\omega} + K\boldsymbol{\sigma} = \Delta\boldsymbol{L} \tag{7.90}$$

Taking the derivative of Eq. (7.90) and making use of $\dot{\boldsymbol{\sigma}} = \frac{1}{4}[B(\boldsymbol{\sigma})]\delta\boldsymbol{\omega}$ given in Eq. (3.150), we obtain a second order differential equation in terms of $\delta\boldsymbol{\omega}$.

$$[I]\delta\ddot{\boldsymbol{\omega}} + [P]\delta\dot{\boldsymbol{\omega}} + \frac{K}{4}[B(\boldsymbol{\sigma})]\delta\boldsymbol{\omega} = \Delta\dot{\boldsymbol{L}} \approx 0 \tag{7.91}$$

The last step holds if the unmodeled torque vector is assumed to change very slowly with time. Note that this differential equation is of the standard form of a damped, spring mass system with a nonlinear spring stiffness matrix $K[B(\boldsymbol{\sigma})]$ if the matrix $[B(\boldsymbol{\sigma})]$ can be shown to be positive definite. To do so we verify that $\boldsymbol{\omega}^T[B(\boldsymbol{\sigma})]\boldsymbol{\omega} > 0$ for any nonzero angular velocity vector $\boldsymbol{\omega}$. Using Eq. (3.150) we find

$$\boldsymbol{\omega}^T[B(\boldsymbol{\sigma})]\boldsymbol{\omega} = \boldsymbol{\omega}^T\left[(1-\sigma^2)I + 2[\tilde{\boldsymbol{\sigma}}] + 2\boldsymbol{\sigma}\boldsymbol{\sigma}^T\right]\boldsymbol{\omega} \tag{7.92}$$

$$= (1-\sigma^2)\boldsymbol{\omega}^T\boldsymbol{\omega} + 2\left(\boldsymbol{\omega}^T\boldsymbol{\sigma}\right)^2 > 0 \tag{7.93}$$

where the last step holds since the MRP attitude vector is maintained such that $\sigma^2 \leq 1$. Therefore the spring-damper-mass system in Eq. (7.91) is stable the angular velocity error vector $\delta\boldsymbol{\omega}$ will approach a constant steady-state value $\delta\boldsymbol{\omega}_{ss}$ as time grows large. Taking the limit of Eq. (7.91) we find the steady state condition

$$K[B(\boldsymbol{\sigma}_{ss})]\delta\boldsymbol{\omega}_{ss} = 0 \tag{7.94}$$

Since $[B(\boldsymbol{\sigma})]$ was shown to be near-orthogonal in Eq. (3.152), it is always of full rank. Therefore the steady-state angular velocity tracking error is

$$\delta\boldsymbol{\omega}_{ss} = 0 \tag{7.95}$$

Thus, even in the presence of an unmodeled external torque vector $\Delta\boldsymbol{L}$, the angular velocity tracking errors will decay to zero asymptotically. However, the attitude tracking errors will not decay to zero. Taking the limit of Eq. (7.90) we find the steady state attitude error to be

$$\boldsymbol{\sigma}_{ss} = \lim_{t\to\infty} \boldsymbol{\sigma} = \frac{1}{K}\Delta\boldsymbol{L} \tag{7.96}$$

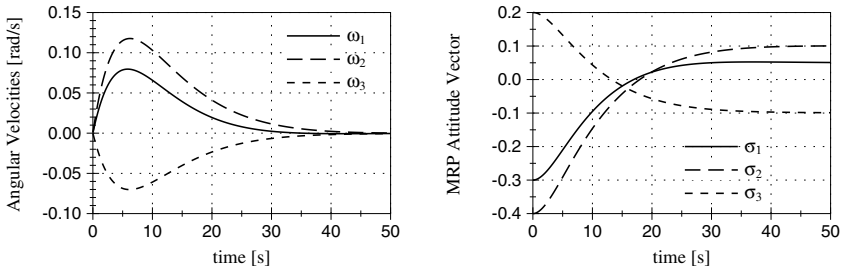
Without further modification to the control law in Eq. (7.82), the attitude tracking errors will settle on a finite offset $\boldsymbol{\sigma}_{ss}$ in the presence of a constant, unmodeled external torque. The control law $\boldsymbol{u}$ is therefore stabilizing in the sense of Lagrange, since the state tracking errors are only guaranteed to remain bounded. The magnitude of the attitude offset $\boldsymbol{\sigma}_{ss}$ can be controlled with the attitude feedback gain K. This steady-state attitude offset is common to PD type control laws (Proportional-Derivative) and is not related to the choice of attitude coordinates. Had Euler parameters or other attitude coordinates been chosen, a similar behavior would have been observed. By using the MRPs though it was possible to analytically predict what the $\boldsymbol{\sigma}_{ss}$ will be. This behavior can be visualized by considering a mass being suspended by the ceiling by a spring with stiffness K. Since a constant gravity force of magnitude mg is acting on the mass, the spring must deflect a certain amount before it can cancel the gravity force. Including drag effects, the mass will then come to rest with the spring stretched a certain distance past its natural, undeformed length. By increasing the spring stiffness, this offset is reduced. Analogous behavior is evident with the PD control laws derived in Eqs. (7.82) through (7.84).

Example 7.9: A rigid body with a large initial attitude error is to be brought to rest at a zero reference attitude. All three principal inertias are 10 kg-m^2. The rigid body is initially at rest with an MRP attitude vector $\boldsymbol{\sigma}(t_0) = (-0.3, -0.4, 0.2)^T$. The control law used to stabilize the body is of the simple PD form

$$\boldsymbol{u} = -K\boldsymbol{\sigma} - P\boldsymbol{\omega}$$

which was shown to be globally asymptotically stabilizing if no unmodeled torques are present. The scalar feedback gains K and P are chosen to be

1 kg-m^2/sec^2 and 3 kg-m^2/sec respectively. To illustrate the steady-state attitude tracking error produced by this control law, an unmodeled, constant external torque vector $\Delta\boldsymbol{L} = (0.05, 0, 10, -0.10)$ Nm is added. This torque is chosen to be much larger than what a spacecraft would normally experience in orbit due to solar or atmospheric drag, to more clearly illustrate its effect and verify the validity of the $\boldsymbol{\sigma}_{ss}$ estimation of Eq. (7.96), even for large disturbances.



(i) Angular Vel. Vector $\boldsymbol{\omega}$ **(ii)** Attitude Vector $\boldsymbol{\sigma}$

Figure 7.5: Steady-State Attitude Offset Due to Unmodeled External Torque Vector

The resulting maneuver is illustrated in Figure 7.5. As predicted in Eq. (7.95), Figure 7.5(i) shows the angular velocity errors decay to zero despite the presence of the external torque vector $\Delta\boldsymbol{L}$. The initial attitude error is reduced by the feedback control law $\boldsymbol{u}$. However, instead of asymptotically approaching zero, they settle down at the offset $\boldsymbol{\sigma}_{ss}$ predicted in Eq. (7.96) given by

$$\boldsymbol{\sigma}_{ss} = \frac{1}{K}\Delta\boldsymbol{L} = \begin{pmatrix} 0.05 \\ 0.10 \\ -0.10 \end{pmatrix}$$

To reduce this attitude offset $\boldsymbol{\sigma}_{ss}$ in the presence of this large $\Delta\boldsymbol{L}$ vector, the attitude feedback gain K would need to be enlarged. This would stiffen up the feedback control and may cause the control devices to be more quickly saturated; other methods to address steady state offset are available.

To achieve asymptotic tracking with unmodeled external torques present, the control law in Eq. (7.82) is modified by adding an integral feedback term. To accomplish this, a new state vector $\boldsymbol{z}$ is introduced.[18]

$$\boldsymbol{z}(t) = \int_0^t (K\boldsymbol{\sigma} + [I]\delta\dot{\boldsymbol{\omega}})\, dt \tag{7.97}$$

Note that $\delta\dot{\boldsymbol{\omega}}$ is given by Eq. (7.35). If the steady state attitude vector $\boldsymbol{\sigma}_{ss}$ is non-zero, then the corresponding state vector $\boldsymbol{z}$ would grow without bounds. Designing a control law that forces $\boldsymbol{z}$ to remain bounded will implicitly force $\boldsymbol{\sigma}$

to go to zero. To design this feedback control law $\boldsymbol{u}$, the Lyapunov function in Eq. (7.78) is augmented by an additional positive definite quantity in terms of the state vector $\boldsymbol{z}$.

$$V(\boldsymbol{\omega}, \boldsymbol{\sigma}, \boldsymbol{z}) = \frac{1}{2}\delta\boldsymbol{\omega}^T[I]\delta\boldsymbol{\omega} + 2K\log\left(1+\boldsymbol{\sigma}^T\boldsymbol{\sigma}\right) + \frac{1}{2}\boldsymbol{z}^T[K_I]\boldsymbol{z} \tag{7.98}$$

where the positive definite matrix $[K_I]$ is the integral feedback gain matrix. Taking the derivative of Eq. (7.98) leads to Lyapunov rate expression

$$\dot{V} = \left(\delta\boldsymbol{\omega} + [K_I]\boldsymbol{z}\right)^T\left([I]\delta\dot{\boldsymbol{\omega}} + K\boldsymbol{\sigma}\right) \tag{7.99}$$

To ensure stability, $\dot{V}$ is set equal to the following negative semi-definite expression

$$\dot{V} = -\left(\delta\boldsymbol{\omega} + [K_I]\boldsymbol{z}\right)^T[P]\left(\delta\boldsymbol{\omega} + [K_I]\boldsymbol{z}\right) \tag{7.100}$$

where $[P]$ is again the positive definite angular velocity feedback gain matrix. Assume at first that no unmodeled external torques are present, then equating Eqs. (7.99) and (7.100) the closed-loop error dynamics are give by

$$[I]\delta\dot{\boldsymbol{\omega}} + [P]\delta\boldsymbol{\omega} + K\boldsymbol{\sigma} + [P][K_I]\boldsymbol{z} = 0 \tag{7.101}$$

Substituting Eq. (7.35) and the rotational equations of motion in Eq. (7.75) into Eq. (7.101) leads to the following feedback control law $\boldsymbol{u}$.

$$\boldsymbol{u} = -K\boldsymbol{\sigma} - [P]\delta\boldsymbol{\omega} - [P][K_I]\boldsymbol{z} + [I]\left(\dot{\boldsymbol{\omega}}_r - [\tilde{\boldsymbol{\omega}}]\boldsymbol{\omega}_r\right) + [\tilde{\boldsymbol{\omega}}][I]\boldsymbol{\omega} - \boldsymbol{L} \tag{7.102}$$

While the definition of the internal error state vector $\boldsymbol{z}$ in Eq. (7.97) is convenient for control analysis purposes, it is not very convenient to implement. In particular the term $[I]\delta\dot{\boldsymbol{\omega}}$ could cause problems since it requires angular acceleration information. Since the inertia matrix of a rigid body $[I]$ is constant, the $\boldsymbol{z}$ vector can also be written in the useful form

$$\boldsymbol{z}(t) = K\int_0^t \boldsymbol{\sigma}dt + [I]\left(\delta\boldsymbol{\omega} - \delta\boldsymbol{\omega}_0\right) \tag{7.103}$$

where $\delta\boldsymbol{\omega}_0$ is the initial body angular velocity error vector. Using this $\boldsymbol{z}$ vector expression, the control law $\boldsymbol{u}$ is expressed as

$$\begin{aligned}\boldsymbol{u} = -K\boldsymbol{\sigma} - \left([P] + [P][K_I][I]\right)\delta\boldsymbol{\omega} - K[P][K_I]\int_0^t \boldsymbol{\sigma}dt + \\ [P][K_I][I]\delta\boldsymbol{\omega}_0 + [I]\left(\dot{\boldsymbol{\omega}}_r - [\tilde{\boldsymbol{\omega}}]\boldsymbol{\omega}_r\right) + [\tilde{\boldsymbol{\omega}}][I]\boldsymbol{\omega} - \boldsymbol{L}\end{aligned} \tag{7.104}$$

The integral feedback gain matrix $[K_I]$ is typically kept small in size relative to the angular velocity feedback gain matrix $[P]$. The integral feedback term is only added to rid the closed-loop dynamics of any non-zero steady-state attitude vectors. It is not desirable for this integral feedback term to drastically

change the closed loop response behavior. As is evident in Eq. (7.104), a relatively large $[K_I]$ would change the effective closed loop frequency and damping characteristics in a substantial manner.

Since $\dot{V}$ in Eq. (7.100) is negative semi-definite, all the states $\boldsymbol{\omega}$, $\boldsymbol{\sigma}$ and $\boldsymbol{z}$ are Lyapunov stable. Further, Eq. (7.100) shows that the quantity $\delta\boldsymbol{\omega}+[K_I]\boldsymbol{z}$ will go to zero. To investigate if the states are asymptotically stable without unmodeled external torques present, the higher order derivatives of V are investigated on the set where $\delta\boldsymbol{\omega} + [K_I]\boldsymbol{z} = 0$. The second derivative is clearly zero on this set. The third derivative, evaluated on this set, can be shown to be

$$\dddot{V}(\boldsymbol{\sigma}, \delta\boldsymbol{\omega} + [K_I]\boldsymbol{z} = 0) = -K^2\boldsymbol{\sigma}^T \left([I]^{-1}\right)[P][I]\boldsymbol{\sigma} \tag{7.105}$$

which is a negative definite quantity. Therefore the attitude error vector $\boldsymbol{\sigma}$ is guaranteed to decay to zero. Due to the kinematic relationship between $\boldsymbol{\sigma}$ and $\delta\boldsymbol{\omega}$, if $\boldsymbol{\sigma}$ goes to zero over time, then so must $\delta\boldsymbol{\omega}$. Since Eq. (7.100) shows that the quantity $\delta\boldsymbol{\omega} + [K_I]\boldsymbol{z}$ goes to zero, then so must the state vector $\boldsymbol{z}$. Thus, in the absence of unmodeled external torques, adding the integral term to the feedback control law did not change the asymptotic stability property.

If an unmodeled external torque $\Delta\boldsymbol{L}$ is included in the dynamical system by substituting Eq. (7.88) into Eq. (7.99), using the feedback control law $\boldsymbol{u}$ in Eq. (7.102) then leads to the following Lyapunov rate expression.

$$\dot{V} = -\left(\delta\boldsymbol{\omega} + [K_I]\boldsymbol{z}\right)^T \left([P]\left(\delta\boldsymbol{\omega} + [K_I]\boldsymbol{z}\right) - \Delta\boldsymbol{L}\right) \tag{7.106}$$

For a non-zero $\Delta\boldsymbol{L}$, this $\dot{V}$ is not negative semi-definite and stability of the states $\delta\boldsymbol{\omega}$, $\boldsymbol{\sigma}$ and $\boldsymbol{z}$ is not guaranteed. However, for a bounded $\Delta\boldsymbol{L}$ Eq. (7.106) does show that $\delta\boldsymbol{\omega}$ and $\boldsymbol{z}$ cannot grow unbounded, since for sufficiently large $\delta\boldsymbol{\omega}$ and $\boldsymbol{z}$ the term

$$-\left(\delta\boldsymbol{\omega} + [K_I]\boldsymbol{z}\right)^T [P]\left(\delta\boldsymbol{\omega} + [K_I]\boldsymbol{z}\right)$$

will dominate and the Lyapunov rate $\dot{V}$ is guaranteed to become negative. For the state $\boldsymbol{z}$ to converge to a finite value as time grows to infinite, then both $\boldsymbol{\sigma}$ and $\dot{\boldsymbol{\omega}}$ must decay to zero. If $\boldsymbol{\sigma}$ would converge to a non-zero vector, then the integral expression in either Eq. (7.97) or (7.103) would grow to infinity. Since $\boldsymbol{\sigma} \to 0$, then due to their kinematic relationship so must $\delta\boldsymbol{\omega} \to 0$. Therefore both and $\delta\boldsymbol{\omega}$ are asymptotically stable. The question that remains is what happens to the state vector $\boldsymbol{z}$ if $\Delta\boldsymbol{L}$ is non-zero. Just because $\boldsymbol{z}$ must remain bounded, we cannot conclude that it also must go to zero. For the $\dot{V}$ expression in Eq. (7.106) to approach $\dot{V} = 0$, the following limit must be true.

$$\lim_{t\to\infty} \left([P]\left(\delta\boldsymbol{\omega} + [K_I]\boldsymbol{z}\right) - \Delta\boldsymbol{L}\right) = 0 \tag{7.107}$$

Since $\delta\boldsymbol{\omega} \to 0$, then

$$\lim_{t\to\infty} \boldsymbol{z} = [K_I]^{-1}[I]^{-1}\Delta\boldsymbol{L} \tag{7.108}$$

This agrees with our earlier stability analysis which stated if $\Delta\boldsymbol{L}$ is zero, then $\boldsymbol{z}$ would go to zero. The closed-loop equations of motion for this system are written as

$$[I]\delta\dot{\boldsymbol{\omega}} + [P]\delta\boldsymbol{\omega} + K\boldsymbol{\sigma} = \Delta\boldsymbol{L} - [P][K_I]\boldsymbol{z} \tag{7.109}$$

Therefore, as $[P][K_I]\boldsymbol{z} \to \Delta\boldsymbol{L}$ as shown in Eq. (7.108), then the integral feedback term in time will effectively cancel the external torque disturbance. Once this happens the closed-loop dynamics are the same as Eq. (7.101) where no unmodeled external torques were present.

Example 7.10: The simulation in Example (7.9) is repeated here with the integral feedback term added to the control law. The initial attitude error is the same as before, but the initial angular velocity error is $\boldsymbol{\omega}(t_0) = (0.2, 0.2, 0.2)$ rad/sec. Since all three principal inertias are set to be equal, then we can use the short-hand notation $I = 10$ kg-m^2. The feedback control law used is

$$\boldsymbol{u} = -K\boldsymbol{\sigma} - [\tilde{\boldsymbol{\omega}}][I]\boldsymbol{\omega} - P(1 + K_I I)\boldsymbol{\omega} - PK_I K\int_0^t \boldsymbol{\sigma}dt + PK_I I\boldsymbol{\omega}(t_0)$$

The scalar integral feedback gain K_I was set to 0.01 sec^{-1}. Having the additional integral feedback term should make $\boldsymbol{\sigma}_{ss}$ go to zero. Further, since $\Delta\boldsymbol{L}$ is a non-zero vector for this simulation, the state vector $\boldsymbol{z}$ is expected to approach the finite limit

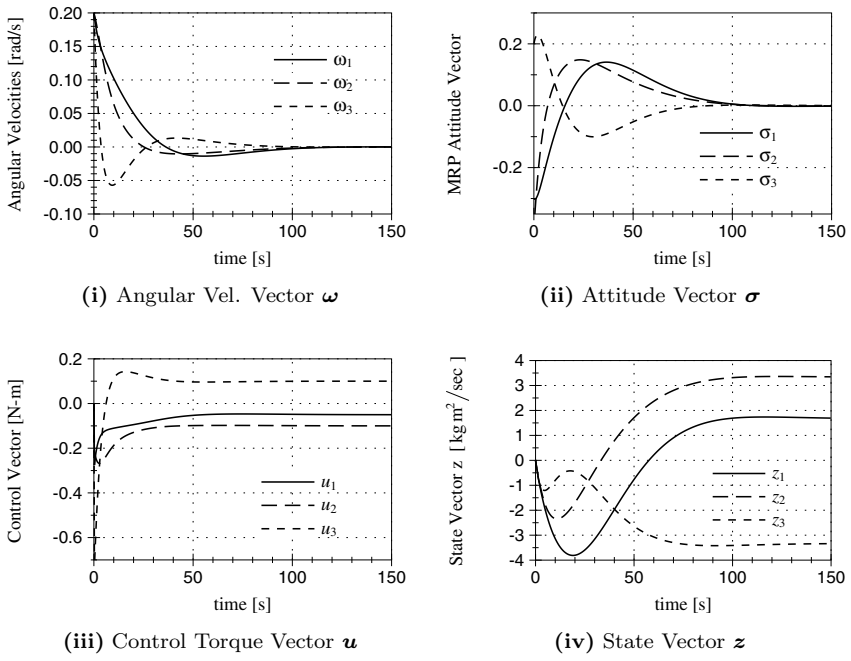
$$\lim_{t\to\infty} \boldsymbol{z} = \frac{1}{K_I I}\Delta\boldsymbol{L} = \begin{pmatrix} 1.66 \\ 3.33 \\ -3.33 \end{pmatrix} \text{ kg-m}^2/\text{sec}$$

The resulting maneuver is illustrated in Figure 7.6. Figures 7.6(i) and 7.6(ii) clearly show that the state errors $\boldsymbol{\omega}$ and $\boldsymbol{\sigma}$ indeed both decay to zero. The control torque vector components are shown Figure 7.6(iii). Since the feedback control law $\boldsymbol{u}$ has to compensate for the external disturbance, its components remain non-zero at the maneuver end. The state vector $\boldsymbol{z}$ is shown in Figure 7.6(iv). As indicated in Eqs. (7.108) and (7.109), the $\boldsymbol{z}$ vector does not decay to zero. Rather, it asymptotically approaches the prescribed values to cancel the influence the external torque disturbance.

Another method to reduce the tracking error $\boldsymbol{\sigma}_{ss}$ would be to adaptively learn the external torque vector by comparing the predicted Lyapunov decay rate to the actual decay rate as shown in Ref. 19. This method has been found to yield a very simple adaptive law that is also able to compensate for inertia matrix modeling errors as well as external torque vectors.

7.3.3 Feedback Gain Selection

To determine appropriate feedback gains, the closed loop dynamics are studied. Substituting the feedback control law $\boldsymbol{u}$ given in Eq. (7.82) into Eq. (7.75)



(i) Angular Vel. Vector $\boldsymbol{\omega}$ **(ii)** Attitude Vector $\boldsymbol{\sigma}$

(iii) Control Torque Vector $\boldsymbol{u}$ **(iv)** State Vector $\boldsymbol{z}$

Figure 7.6: Integral Feedback Control Law Simulation with Unmodeled External Torque Vector

the closed loop dynamics for the controlled rigid body in the absence of any unmodeled external torques are[17]

$$[I]\delta\dot{\boldsymbol{\omega}} = -[P]\delta\boldsymbol{\omega} - K\boldsymbol{\sigma} \tag{7.110}$$

Note that due to the use of the MRPs and the logarithm function in Eq. (7.78), the closed loop tracking dynamics are *rigorously linear* in both the body angular velocity and attitude error vectors, even though large general motions are considered. Eq. (7.110) is solved along with the differential kinematic equation expressed in Eq. (3.150) to determine the exact closed-loop performance. The only nonlinearities that appear in these two differential equations are the quadratic nonlinearities present in the MRP kinematic differential equation. Linearizations are attractive in order to impose terminal motion control law design criteria. Linearizing Eq. (3.150) about $\boldsymbol{\sigma}= 0$ yields

$$\dot{\boldsymbol{\sigma}} \simeq \frac{1}{4}\boldsymbol{\omega} \tag{7.111}$$

Since the MRPs "behave like angles over four" for small angles, this linearization will be applicable for a relatively large range of rotations. As a comparison, the classical Rodrigues parameters only linearize as angles over two, whereas the standard Euler angles simply linearize as angle type quantities.

The linearized set of closed loop equations are given by

$$\begin{pmatrix} \dot{\boldsymbol{\sigma}} \\ \delta\dot{\boldsymbol{\omega}} \end{pmatrix} = \begin{bmatrix} 0 & \frac{1}{4}I \\ -K[I]^{-1} & -[I]^{-1}[P] \end{bmatrix} \begin{pmatrix} \boldsymbol{\sigma} \\ \delta\boldsymbol{\omega} \end{pmatrix} \tag{7.112}$$

Given the rigid body inertia matrix $[I]$, any standard linear control design method such as a pole placement method can be used to determine the desired response of the linearized closed loop dynamics. If both the inertia matrix $[I]$ and the angular velocity feedback gain matrix $[P]$ are diagonal matrices with entries I_i and P_i respectively, then Eq. (7.112) can be conveniently decoupled into three sets of differential equations[12, 17]

$$\begin{pmatrix} \dot{\boldsymbol{\sigma}}_i \\ \delta\dot{\boldsymbol{\omega}}_i \end{pmatrix} = \begin{bmatrix} 0 & \frac{1}{4} \\ -\frac{K}{I_i} & -\frac{P_i}{I_i} \end{bmatrix} \begin{pmatrix} \boldsymbol{\sigma}_i \\ \delta\boldsymbol{\omega}_i \end{pmatrix} \qquad i = 1, 2, 3 \tag{7.113}$$

whose roots are explicitly given by

$$\lambda_i = -\frac{1}{2I_i}\left(P_i \pm \sqrt{-KI_i + {P_i}^2}\right) \qquad i = 1, 2, 3 \tag{7.114}$$

For an underdamped system, the corresponding closed loop natural frequencies ω_{n_i} and damping ratios ξ_i are

$$\omega_{n_i} = \frac{1}{2I_i}\sqrt{KI_i - 2P_i^2} \tag{7.115}$$

$$\xi_i = \frac{P_i}{\sqrt{KI_i - 2P_i^2}} \tag{7.116}$$

The decay time constants T_i which indicates how long it would take for the state errors to decay to $\frac{1}{e}$ of their respective initial values are given by

$$T_i = \frac{2I_i}{P_i} \tag{7.117}$$

It is interesting to note that only the angular velocity feedback gain constants P_i dictate how fast the state errors will decay. The attitude feedback gain constant K contributes to both the natural frequency and damping ratios of the closed loop response. The damped natural frequency ω_{d_i} is given by

$$\omega_{d_i} = \frac{1}{2I_i}\sqrt{KI_i - P_i^2} \tag{7.118}$$

This enables an explicit "pole placement" design process in which (K, P_i) can be chosen to achieve specific desired (ω_{d_i}, T_i) or (ω_{d_i}, ξ_i) characteristics for the closed loop system. The use of the feedback control law in Eq. (7.82), along with the design of the feedback gains, is illustrated in the following numerical simulation.

Example 7.11: Assume a rigid body with body axes aligned along the principal axes is initially in a tumbling situation. The reference trajectory is set to be the zero attitude at rest. The parameter values for the numerical simulation are shown in Table 7.11. The relative orientation of the rigid body to the zero attitude is expressed through the MRP vector $\boldsymbol{\sigma}$. Note that the rigid body has a large initial angular velocity about the first body axis which will cause the body to tumble through the $\Phi = 180$ degree orientation. Other three-parameter sets of attitude coordinates cannot describe arbitrary rotations without encountering singularities. For example, had the classical Rodrigues parameters been used in the simulation they would early on encounter a singularity at $\Phi = 180$ degrees.

Table 7.2: Parameter of MRP Control Law Numerical Simulation

Parameter	Value	Units
I_1	140.0	kg-m^2
I_2	100.0	kg-m^2
I_3	80.0	kg-m^2
$\boldsymbol{\sigma}(t_0)$	$[0.60 \ -0.40 \ 0.20]$	
$\boldsymbol{\omega}(t_0)$	$[0.70 \ 0.20 \ -0.15]$	rad/sec
$[P]$	$[18.67 \ 2.67 \ 10.67]$	kg-m^2/sec
K	7.11	kg-m^2/sec^2

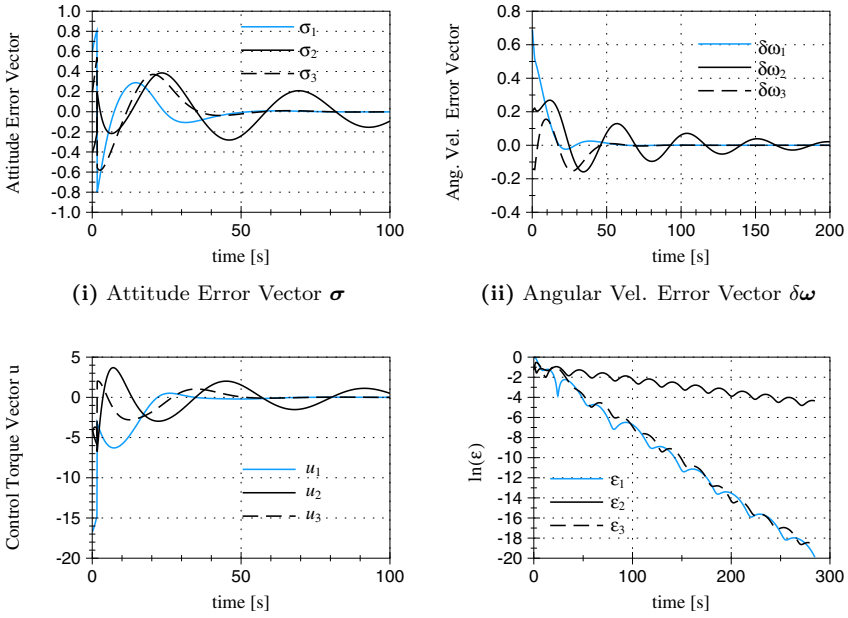
The feedback gains for this simulation were chosen such that the closed loop dynamics will be very underdamped. Clearly the resulting performance would not be what is needed to control a real system. However, having visible state oscillations present will allow for the predicted damped natural frequency in Eq. (7.118) and decay time constants in Eq. (7.117) to be verified.

The results of the numerical simulation are shown in Figure 7.7. The control vector $\boldsymbol{u}$ stabilizes the tumbling rigid body and brings it to rest at the zero attitude. The decay time constant T_2, which controls how fast the states σ_s and ω_s are reduced, was chosen purposely to be much larger than the other two time constants. This results in the second body axis state errors being reduced much slower than the other two, simulating a situation where less control authority is present about this axis. As is seen in Figures 7.7(i) and 7.7(ii) the nonlinear response corresponds very well with the linearized prediction. As the body tumbles through the "upside down" orientation at $\Phi = 180°$, the MRP vector switches automatically near $\sigma_2 = 1$ to the alternate set. At this point the corresponding control law ceases to fight the tumble and lets the body complete the revolution before bringing it to rest at the origin as seen in Figure 7.7(iii).

Let vector $\boldsymbol{\epsilon}$ be the state error vector whose components are given by

$$\epsilon_i = \sqrt{\sigma_i^2 + \omega_i^2} \qquad i = 1, 2, 3 \tag{7.119}$$

To study the damped natural frequencies ω_{d_i} and the decay times T_i the natural logarithm of ϵ_i is plotted in Figure 7.7(iv). This Figure clearly shows the decay rate and the natural oscillations of the underdamped response.



(i) Attitude Error Vector $\boldsymbol{\sigma}$ **(ii)** Angular Vel. Error Vector $\delta\boldsymbol{\omega}$

(iii) Feedback Control Vector $\boldsymbol{u}$ **(iv)** State Error Vector $\boldsymbol{\epsilon}$

Figure 7.7: MRP Feedback Control Law Simulation Without Control Constraints

Note that the simulated maneuver performs a very large rotation which includes a complete tumble. Typically, when studying the closed loop response of a control law, only small attitude errors in the order of 10s of degrees are used. Table 7.3 compares the actual averaged decay rates and damped natural frequencies of the nonlinear system to the ones predicted by the linearized feedback gain design. As expected, the linearization used in Eq. (7.111) yields accurate closed loop performance predictions because of the extremely large domain in which the exact $\boldsymbol{\sigma}$-motion is near linear. The percent differences between the actual nonlinear T_i and ω_{d_i} and the ones obtained from the linearized model are only in the 1 to 2 percent range. Thus the MRP feedback law in Eq. (7.82) achieves predictable, global, asymptotic stability by only using three attitude coordinates as compared to four coordinates required by Euler parameter feedback laws. Some control laws using other three parameter sets of attitude coordinates such as the standard yaw, pitch and roll angles also claim to have global stability. However, they all come with a disclaimer warning against rotating the rigid body to certain attitudes because of the inherent singularities of the chosen attitude coordinates. Such control laws can therefore hardly be considered globally stabilizing. The MRP attitude description allows for arbitrary rotations and has the added benefit of always indicating the shortest rotational distance back to the origin when the switching surface $\sigma^2 = 1$ is chosen. There is one caveat for bounded

Table 7.3: Comparison of Actual Averaged Closed-Loop Response Parameters vs. Predicted Linearized Values

Parameter	Actual Average	Predicted Value	Percent Difference
T_1	14.71 s	15.00 s	1.97%
T_2	76.92 s	75.00 s	-2.50%
T_3	14.71 s	15.00 s	1.97%
ω_{d_1}	0.0938 rad/s	0.0909 rad/s	-3.12%
ω_{d_2}	0.1326 rad/s	0.1326 rad/s	0.08%
ω_{d_3}	0.1343 rad/s	0.1333 rad/s	-0.74%

controls: large unknown disturbances which cause cyclic passage through the 180° error condition could cause a problem with control chatter.

7.4 Lyapunov Optimal Control Laws

The feedback control laws in Eqs. (7.82) and (7.104) were developed assuming that no control magnitude constraints are present. However, most control devices such as reaction wheels, CMGs or thrusters have an upper bound on how much control authority they can exert onto a system. If a control device is operating at such a bound, it is said to be saturated. This section investigates the stability of dynamical systems with saturated control present.

There are essentially two possibilities for dealing with saturated controls. One solution is to reduce the feedback gains such that the anticipated required control effort never saturates *any* control devices. This is typically done when designing open-loop reference trajectories. However, this method has the drawback that the overall performance of the feedback control law is greatly reduced, perhaps to an un-acceptable degree. When trying to stabilize a system about a reference state, a more efficient method of dealing with saturated controls is to allow individual control devices to become saturated. This leads to a saturated control law which is said to be *Lyapunov optimal*. Being Lyapunov optimal means that the time derivative of the given Lyapunov function V is made as negative as possible during intervals where one or more of the control devices are saturated.[9, 20, 21] However, certain difficulties, including possible loss of controllability, are potentially implicit in this approach if the most negative $\dot{V}$ is still positive!

The goal of the controller design process is to choose a control law from an admissible set that will stabilize the system in an optimal fashion (i.e. make $\dot{V}$ as negative as possible). For saturated controls of a natural system, the classical stabilizing controller takes the form

$$Q_i = -Q_{i_{max}} sgn(\dot{q}_i) \tag{7.120}$$

where Q_i are the generalized control forces, $\dot{q}_i$ are the derivatives of the generalized coordinates and $Q_{i_{max}}$ are the control bounds. This control law is Lyapunov

optimal for minimizing the performance index J

$$J = \dot{V} = \sum_{i=1}^{n} \dot{q}_i Q_i \tag{7.121}$$

The control law is optimal in a sense analogous to Pontryagin's Principle for optimal control because the controls are selected from an admissible set $|Q_i| \leq Q_{i_{max}}$ such that the instantaneous work rate (in the common event that V has an energy interpretation) is minimized at every point in time. Note that mathematical difficulties and practical system performance issues arise if this controller is implemented directly for most systems.[22] The discontinuity at the origin must typically be replaced with a region of unsaturated control to avoid chattering near $\dot{q}_i = 0$. This unsaturated controller can either approximate the discontinuity or be some other stable/optimal feedback controller that transitions from the saturated controller on the saturation boundary. We restrict attention here to control laws that transition continuously at the saturation boundary. The obvious choice is to augment Eq. (7.120) with a linear controller of the type

$$Q_i = \begin{cases} -K_i \dot{q}_i & \text{for } |K_i \dot{q}_i| \leq Q_{i_{max}} \\ -Q_{i_{max}} sgn(\dot{q}_i) & \text{for } |K_i \dot{q}_i| > Q_{i_{max}} \end{cases} \tag{7.122}$$

where $K_i > 0$ is a chosen feedback gain. This control continuously transitions across the saturation boundary and eliminates chattering. Note that Eq. (7.122) allows some elements of the control vector to become saturated, while others are still in the unsaturated range. This differs from conventional gain scheduling and deadband methods which typically reduce the feedback gains to keep all controls in the unsaturated range.

Example 7.12: Let us design a saturated control law for a single degree of freedom nonlinear oscillator. Assuming $m, c, k, k_N > 0$, the equation of motion of a Duffing oscillator are given by

$$m\ddot{x} + c\dot{x} + kx + k_N x^3 = u$$

The Lyapunov function (the system Hamiltonian of the unforced and undamped system) is

$$V = \frac{1}{2} m\dot{x}^2 + \frac{1}{2} kx^2 + \frac{1}{4} k_N x^4$$

V is positive definite and vanishes only at the origin, which is the only real equilibrium point of the un-forced system. The performance index is the time derivative of V and can be written immediately from Eq. (7.43) as

$$J = \dot{V} = \dot{x}Q = \dot{x}(-c\dot{x} + u)$$

For bounded control $|u| \leq u_{max}$, the performance index J is minimized by the feedback controller

$$u = -u_{max} sgn(\dot{x})$$

Using this control law, $\dot{V}$ is reduced to the energy dissipation rate

$$\dot{V}(x,\dot{x}) = -c\dot{x}^2 - u_{max}\dot{x}\cdot sgn(\dot{x})$$

It is of interest to note that an arbitrary, unknown, positive definite potential energy function $\Delta V(x)$ could be added to V — and *exactly the same result is obtained* for $\dot{V}$ and u. Thus the structure of the control law and the stability guarantee is invariant (and therefore inherently robust) with respect to a large family of modeling assumptions.

Since $\dot{V}(x,\dot{x})$ is negative semi-definite, it can only be concluded at this point that the system has globally stable motion near the origin; thus x and $\dot{x}$ will remain bounded. Since the control u is bounded by definition, the duffing oscillator equations of motion show that $\ddot{x}$ will also be bounded. To prove asymptotic stability, the higher derivatives of V must be investigated. The only point where $\dot{V}$ vanishes is $\dot{x}=0$. The second derivative of V is

$$\frac{d^2V}{dt^2} = -2c\dot{x}\ddot{x} - u_{max}\ddot{x}sgn(\dot{x})$$

which is also zero for all x when $\dot{x}=0$. The third derivative of V is

$$\frac{d^3V}{dt^3} = -2c\ddot{x}^2 - 2c\dot{x}\frac{d^3x}{dt^3} - u_{max}\frac{d^3x}{dt^3}sgn(\dot{x})$$

Using the duffing oscillator equations of motion, we find on the set where $\dot{x}=0$ that

$$\left.\frac{d^3V}{dt^3}\right|_{\dot{x}=0} = -2\frac{c}{m}\left(kx + k_N x^3\right)^2$$

which is a negative definite function of x. Therefore, according to Theorem 7.5, the saturated control law u is globally asymptotically stabilizing.

If a tracking control law is subjected to control constraints, then Lyapunov optimality is difficult to define because tracking stability cannot be guaranteed during saturated control intervals. Nevertheless, globally asymptotically stable tracking controllers can often be achieved by generalizing the method developed in this section. A generalized work/energy equation that is equivalent to Eq. (7.43) is not possible because the position and/or attitude error tracking coordinates are measured in a non-inertial reference frame (thus a more tedious process is required to establish the $\dot{V}$ equation for each system. Also, consideration must be given to whether the prescribed trajectory is a feasible exact trajectory of the system.

Consider the case of having a rigid body track a given reference trajectory $\boldsymbol{\omega}_r(t)$. The unsaturated control law $\boldsymbol{u}_{us}$ given by

$$\boldsymbol{u}_{us} = -K\boldsymbol{\sigma} - [P]\delta\boldsymbol{\omega} + [I]\left(\dot{\boldsymbol{\omega}}_r - [\tilde{\boldsymbol{\omega}}]\boldsymbol{\omega}_r\right) + [\tilde{\boldsymbol{\omega}}][I]\boldsymbol{\omega} - \boldsymbol{L} \tag{7.123}$$

has been shown to be globally, asymptotically stabilizing. The corresponding Lyapunov rate function $\dot{V}$ can be expressed as

$$\dot{V} = \delta\boldsymbol{\omega}^T\left(-[\tilde{\boldsymbol{\omega}}][I]\boldsymbol{\omega} + \boldsymbol{u} - [I]\left(\dot{\boldsymbol{\omega}}_r - [\tilde{\boldsymbol{\omega}}]\boldsymbol{\omega}_r\right) + K\boldsymbol{\sigma}\right) \tag{7.124}$$

Assume that the available control torque about the i-th body axis is limited by u_{max_i}. Then following earlier analysis, we augment the unsaturated control law $\boldsymbol{u}_{us}$ with a Lyapunov optimal saturated term to yield a modified control law $\boldsymbol{u}$.

$$u_i = \begin{cases} u_{us_i} & \text{for } |u_{us_i}| \leq u_{max_i} \\ u_{us_i} \cdot sgn(u_{us_i}) & \text{for } |u_{us_i}| > u_{max_i} \end{cases} \tag{7.125}$$

A conservative stability boundary (a sufficient condition for stability) for this modified control torque $\boldsymbol{u}$ is found to be

$$|\left([I]\left(\dot{\boldsymbol{\omega}}_r - [\tilde{\boldsymbol{\omega}}]\boldsymbol{\omega}_r\right) + [\tilde{\boldsymbol{\omega}}][I]\boldsymbol{\omega} - K\boldsymbol{\sigma}\right)_i| \leq u_{max_i} \tag{7.126}$$

Note that, for this higher dimensional system, this stability constraint may be overly conservative. The condition in Eq. (7.126) is clearly violated if the inequality fails about any one body axis.

Let us now consider the problem of a tumbling rigid body where the controls are saturated. In such a case tracking a reference trajectory is no longer a primary concern, rather stabilizing the motion is. Therefore $\boldsymbol{\omega}_r(t)$ is set to zero. This allows $\dot{V}$ of Eq. (7.124) to be simplified, using $\boldsymbol{\omega}^T[\tilde{\boldsymbol{\omega}}]$, to

$$\dot{V} = \boldsymbol{\omega}^T\left(\boldsymbol{u} + K\boldsymbol{\sigma}\right) \tag{7.127}$$

The control torque $\boldsymbol{u}_{us}$ for unsaturated conditions is then reduced to

$$\boldsymbol{u}_{us} = -K\boldsymbol{\sigma} - [P]\boldsymbol{\omega} \tag{7.128}$$

A conservative stability condition is found by studying $\dot{V}$ in Eq. (7.127):

$$K|\sigma_i| \leq u_{max_i} \tag{7.129}$$

Since the magnitude of the MRP attitude error vector $\boldsymbol{\sigma}$ is bounded by 1, this stability condition can also be written as

$$K \leq u_{max_i} \tag{7.130}$$

As shown in Ref. 9, while this condition in Eq. (7.129) guarantees stability, it it not a necessary condition for stability. If one simply wanted to stop the tumbling motion without regard to the final attitude, then one could set $K = 0$. Assume that the velocity feedback gain matrix $[P]$ is diagonal. Then the saturated control law $\boldsymbol{u}$ is

$$u_i = \begin{cases} -P_{ii}\omega_i & \text{for } |P_{ii}\omega_i| \leq u_{max_i} \\ -u_{max_i} \cdot sgn(\omega_i) & \text{for } |P_{ii}\omega_i| > u_{max_i} \end{cases} \tag{7.131}$$

which leads to the Lyapunov rate function

$$\dot{V}(\boldsymbol{\omega}) = -\sum_{i=1}^{M} P_{ii}\omega_i^2 - \sum_{i=M+1}^{N} \omega_i \cdot sgn(\omega_i) \tag{7.132}$$

where M is the number of unsaturated control inputs currently present. Since this $\dot{V}$ is *negative definite* (we only care in this case about $\boldsymbol{\omega}$, not about attitude), the combined saturated, unsaturated control law in Eq. (7.131) is globally asymptotically stabilizing.

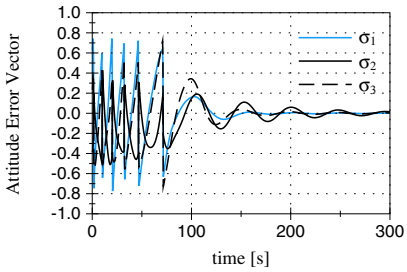
Limiting K through the stability condition in Eq. (7.130) is usually overly conservative. As the numerical simulation in Example 7.13 will illustrate, having a $K > u_{max_i}$ still typically leads to an asymptotically stable closed-loop dynamics. The reason for this is the bounded nature of the attitude error vector. The stability condition in Eq. (7.129), and therefore the requirement of $\dot{V}$ being negative, may indeed be locally violated for finite periods of time. These violations are likely to occur whenever the rigid body tumbles towards the $\Phi = \pm 180$ degrees condition. After the body tumbles past $\Phi = \pm 180$ degrees, the sign of attitude vector components are switched through $\sigma^S = -\sigma$. As is seen in Figure 7.7(iii), the required unsaturated control torque drops drastically in magnitude during this switching. Before the switching, where the body is still rotating away from the origin, both the angular velocity and the attitude feedback are demanding a control torque in the same direction and their effects are added up to produce the large control torque before the switching. After the switching at the $\sigma^2 = 1$ surface, the body now starts to rotate back towards the origin and the sign of the attitude feedback control is switched. This results in the angular velocity and attitude feedback control partially cancelling each other and therefore producing a much smaller control torque. Therefore the required control torques are larger and more likely to be saturated approaching $\Phi = \pm 180$ degrees then they are leaving the "upside-down orientation." Since the body is tumbling, the $\boldsymbol{\sigma}$ vector magnitude will always periodically come close to zero where the stability condition in Eq. (7.129) is satisfied and kinetic energy is guaranteed to be pumped out of the system because we guarantee in Eq. (7.132) that $\dot{V} \leq 0$. Eventually, perhaps after *several* revolutions or tumbles, the body will come to rest.

Example 7.13: The rigid body detumbling maneuver in Example 7.11 is repeated here in the presence of control constraints $u_{max_i} = 1$ Nm. The unsaturated control law

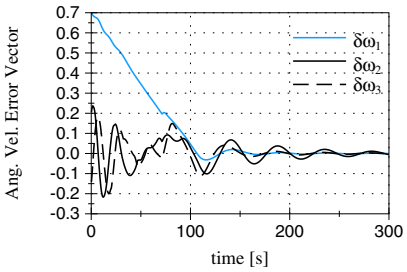
$$\boldsymbol{u} = -K\boldsymbol{\sigma} - [P]\boldsymbol{\omega} \tag{7.133}$$

is augmented with a saturated control law as shown in Eq. (7.125). The numerical simulation results are shown in Figure 7.8.

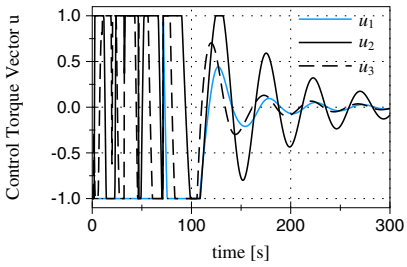
As Figure 7.8(i) shows, with the limited control effort present the rigid body now performs about five tumbles before coming to rest at the origin. The large initial body angular velocity about the first body axis is gradually reduced until the control torque $\boldsymbol{u}$ remains in the unsaturated regime. As shown in Figure 7.8(ii), from there on ω_1 starts to exhibit the anticipated underdamped oscillations as were present with the unsaturated control law. For the first 100 seconds into the simulation, the control torque components u_i remain mostly saturated as shown in Figure 7.8(iii). Once the angular velocity errors are sufficiently reduced, the required control effort remains in the unsaturated



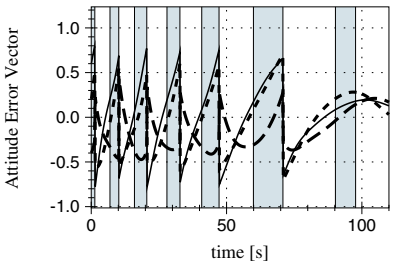
(i) Attitude Error Vector $\boldsymbol{\sigma}$



(ii) Angular Vel. Error Vector $\delta\boldsymbol{\omega}$



(iii) Feedback Control Vector $\boldsymbol{u}$



(iv) Illustration of Positive $\dot{V}$ Regions With the MRP Vector Components Superimposed

Figure 7.8: Saturated MRP Feedback Control Law Simulation

regime. Figure 7.8(iv) shows the MRP attitude vector components for up to 120 seconds into the simulation. In the background the time regions are grayed out where the Lyapunov function time derivative $\dot{V}$ is actually positive for this dynamical system. As predicted, $\dot{V}$ becomes temporarily positive when the rigid body is rotating towards the "upside-down" orientation. As soon as the body rotates past this orientation $\dot{V}$ becomes negative again. This happens even though the control torque vector components u_i are mostly still saturated.

Therefore, even though K was chosen to be much larger than $u_{max_i} = 1$ for this simulation, the saturated control law in Eq. (7.125) still asymptotically stabilized the rigid body. Thus, one beautiful property of this MRP feedback control law is that not only does it perform well for small orientation errors, it also scales well to handle the much tougher problem of controlling *arbitrary* large tumbling motions in the presence of control saturation.

7.5 Linear Closed-Loop Dynamics

Whereas the previous attitude feedback control laws were found by first defining a candidate Lyapunov function and then extracting the corresponding stabilizing control, it it also possible to start out instead with a desired (or prescribed) set of stable closed-loop dynamics and then extract the corresponding feedback control law using a variation of the "inverse dynamics" approach common in robotics open-loop path planning problems. This technique is very general and can be applied to a multitude of systems. Paielli and Bach present such a control law derived in terms of the Euler parameter components in Ref. 23. Let the $\boldsymbol{\epsilon} = (\beta_1, \beta_2, \beta_3)^T$ be the vector portion of the Euler parameters as defined in Eq. (3.109). Note that $\boldsymbol{\epsilon}$ contains information about both the principal rotation axis and principal rotation angle. Therefore, if $\boldsymbol{\epsilon} \to 0$, then the body has rotated back to the origin. Let's assume that we desire the closed loop dynamics to have the following prescribed *linear* form

$$\ddot{\boldsymbol{\epsilon}} + P\dot{\boldsymbol{\epsilon}} + K\boldsymbol{\epsilon} = 0 \tag{7.134}$$

where P and K are the positive scalar velocity and position feedback gains. From linear control theory it is evident that for any initial $\boldsymbol{\epsilon}$ and $\dot{\boldsymbol{\epsilon}}$ vectors, the resulting motion would obviously be asymptotically stable. If desired, one could also easily add an integral feedback term to the desired closed loop equations.

$$\ddot{\boldsymbol{\epsilon}} + P\dot{\boldsymbol{\epsilon}} + K\boldsymbol{\epsilon} + K_i \int_0^t \boldsymbol{\epsilon} dt = 0 \tag{7.135}$$

With judicious choices for P, K and K_i, stable $\boldsymbol{\epsilon}(t)$ motions can be specified by Eq. (7.135). Note that instead of the Euler parameter vector component $\boldsymbol{\epsilon}$, any attitude or position vector could have been used. Next we will impose kinematic and dynamical differential constraints to find the control law that will render the closed loop dynamics of a rigid body equal to Eq. (7.134). Through Euler's rotational equation of motion

$$[I]\dot{\boldsymbol{\omega}} + [\tilde{\boldsymbol{\omega}}][I]\boldsymbol{\omega} = \boldsymbol{u} \tag{7.136}$$

once the body angular acceleration vector $\dot{\boldsymbol{\omega}}$ is found consistent with Eq. (7.134) or (7.135), then the control law vector $\boldsymbol{u}$ is also given. To find an expression for $\dot{\boldsymbol{\omega}}$, we need to differentiate the Euler parameter kinematic differential equation. Assuming the target state has zero angular velocity, from Eq. (3.104) the vector $\dot{\boldsymbol{\epsilon}}$ is expressed as

$$\dot{\boldsymbol{\epsilon}} = \frac{1}{2}[T]\boldsymbol{\omega} \tag{7.137}$$

where the matrix $[T] = [T(\beta_0, \boldsymbol{\epsilon})]$ is given by

$$[T] = \begin{bmatrix} \beta_0 & -\beta_3 & \beta_2 \\ \beta_3 & \beta_0 & -\beta_1 \\ -\beta_2 & \beta_1 & \beta_0 \end{bmatrix} = \beta_0 I_{3\times 3} + [\tilde{\boldsymbol{\epsilon}}] \tag{7.138}$$

The term $\dot{\beta}_0$ is also expressed from Eq. (3.104) as

$$\dot{\beta}_0 = -\frac{1}{2}\boldsymbol{\epsilon}^T\boldsymbol{\omega} \tag{7.139}$$

Differentiating the differential kinematic equation in Eq. (7.137) we find

$$\ddot{\boldsymbol{\epsilon}} = \frac{1}{2}[T]\dot{\boldsymbol{\omega}} + \frac{1}{2}[\dot{T}]\boldsymbol{\omega} \tag{7.140}$$

Using Eq. (7.138) the term $[\dot{T}]\omega$ is expressed as

$$[\dot{T}]\omega = \dot{\beta}_0\boldsymbol{\omega} - [\tilde{\boldsymbol{\omega}}]\dot{\boldsymbol{\epsilon}} \tag{7.141}$$

Substituting this $[\dot{T}]\omega$ and making use of Eqs. (7.137) and (7.139), the vector $\ddot{\boldsymbol{\epsilon}}$ is written as

$$\ddot{\boldsymbol{\epsilon}} = \frac{1}{2}[T]\dot{\boldsymbol{\omega}} - \frac{1}{4}\left(\boldsymbol{\epsilon}^T\boldsymbol{\omega}\boldsymbol{\omega} + [\tilde{\boldsymbol{\omega}}][T]\boldsymbol{\omega}\right) \tag{7.142}$$

This expression can be further simplified by substituting Eq. (7.138) and using the identities $[\tilde{\boldsymbol{a}}]\boldsymbol{a} = 0$ and

$$[\tilde{\boldsymbol{a}}][\tilde{\boldsymbol{a}}] = \boldsymbol{a}\boldsymbol{a}^T - \boldsymbol{a}^T\boldsymbol{a}I_{3\times 3} \tag{7.143}$$

Using these Eq. (7.142) is written in its most compact form

$$\ddot{\boldsymbol{\epsilon}} = \frac{1}{2}[T]\dot{\boldsymbol{\omega}} - \frac{1}{4}\omega^2\boldsymbol{\epsilon} \tag{7.144}$$

where the shorthand notation $\omega^2 = \boldsymbol{\omega}^T\boldsymbol{\omega}$ is used. Eq. (7.144) introduces the necessary $\dot{\boldsymbol{\omega}}$ term which leads to the control vector $\boldsymbol{u}$. After substituting Eqs. (7.137) and (7.144) into our desired linear closed loop dynamics of Eq. (7.134), the following constraint equation can be found.

$$[T]\left(\dot{\boldsymbol{\omega}} + P\boldsymbol{\omega} + [T]^{-1}\left(-\frac{1}{2}\omega^2\boldsymbol{\epsilon} + 2K\boldsymbol{\epsilon}\right)\right) = 0 \tag{7.145}$$

The matrix inverse of $[T]$ can be written explicitly as

$$[T]^{-1} = [T]^T + \frac{1}{\beta_0}\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T \tag{7.146}$$

This expression can be readily verified by using it to confirm $[T]^{-1}[T] = I_{3\times 3}$. From Eq. (7.146) it is evident that the matrix inverse of $[T]$ is always possible except when $\beta_0 \to 0$. This corresponds to the rigid body being rotated $\pm$ 180 degrees relative to the target state at $\boldsymbol{\epsilon} = 0$. Since $[T]$ is of full rank everywhere with the exception when $\beta_0 = 0$, other than at this particular orientation, from Eq. (7.145) the following acceleration constraint must hold to achieve the desired tracking dynamics of Eq. (7.134):

$$\dot{\boldsymbol{\omega}} + P\boldsymbol{\omega} + [T]^{-1}\left(-\frac{1}{2}\omega^2\boldsymbol{\epsilon} + 2K\boldsymbol{\epsilon}\right) = 0 \tag{7.147}$$

Substituting Eq. (7.138) and (7.146), the body angular acceleration vector $\dot{\boldsymbol{\omega}}$ is expressed as

$$\dot{\boldsymbol{\omega}} = -P\boldsymbol{\omega} - 2\left(K - \frac{\omega^2}{4}\right)\frac{\boldsymbol{\epsilon}}{\beta_0} \tag{7.148}$$

After substituting this desired (or required) acceleration $\dot{\boldsymbol{\omega}}$ into Euler's rotational equations of motion, we can solve for the required nonlinear feedback control law vector $\boldsymbol{u}$ as

$$\boldsymbol{u} = [\tilde{\boldsymbol{\omega}}][I]\boldsymbol{\omega} + [I]\left(-P\boldsymbol{\omega} - 2\left(K - \frac{\omega^2}{4}\right)\frac{\boldsymbol{\epsilon}}{\beta_0}\right) \tag{7.149}$$

Note that since the vector $\boldsymbol{q} = \boldsymbol{\epsilon}/\beta_0$ is the Gibbs vector, this control law is alternatively written as a function of $(\boldsymbol{q}, \boldsymbol{\omega})$ and is singular near principal rotations of $\pm$ 180 degrees. Therefore this control law is not globally stabilizing despite the linear closed loop dynamics (motions which tumble through $\pm$180 degrees are excluded). It is interesting to compare this control law with the Gibbs vector control law previously derived from a Lyapunov function. Using the Lyapunov functions in Eq. (7.29) and (7.58) leads to the asymptotically stabilizing control law

$$\boldsymbol{u} = [\tilde{\boldsymbol{\omega}}][I]\boldsymbol{\omega} - P\boldsymbol{\omega} - K\boldsymbol{q} \tag{7.150}$$

The reason the control law in Eq. (7.150) does not lead to linear closed loop dynamical equations is because of the quadratic nonlinearity present in the Gibbs vector differential kinematic equations. However, this nonlinearity is very weak for a large range of rotations and $\omega^2\boldsymbol{\epsilon}$ and $[\tilde{\boldsymbol{\omega}}][I]\boldsymbol{\omega}$ can typically be ignored when designing the feedback gains. Only if it is necessary to precisely predict the closed loop behavior with linear control theory is it advantageous to add the more complex $\boldsymbol{\omega}$ and $\boldsymbol{q}$ feedback in Eq. (7.149). We remark that it is possible to parallel the above developments using the MRP vector $\boldsymbol{\sigma}$ instead of $\boldsymbol{\epsilon}$ (or $\boldsymbol{q}$), and eliminate the singularity at $\pm$180 degrees, and still have a linear tracking error dynamics. This is evident in Example 7.14 below.

An attractive part of this methodology is that the structure of the closed loop equations can easily be modified using standard linear control theory techniques. It is also of paramount importance that this methodology has been generalized with adaptive control methods to obtain even more robust version of these feedback control laws.[24, 25] If it is necessary that the feedback control reject external disturbances, an integral measure of the attitude error is added to the closed loop equations as shown in Eq. (7.135). Following similar steps as were done previously in this section, the desired body angular acceleration vector $\dot{\boldsymbol{\omega}}$ which results in an integral feedback control law is then written as

$$\dot{\boldsymbol{\omega}} = -P\boldsymbol{\omega} - 2\left(K - \frac{\omega^2}{4}\right)\frac{\boldsymbol{\epsilon}}{\beta_0} - 2K_i\left([T]^T + \frac{1}{\beta_0}\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T\right)\int_0^t \boldsymbol{\epsilon} dt \tag{7.151}$$

Where the integral feedback term in the Lyapunov function derived control law in Eq. (7.104) has a constant feedback gain, the integral feedback of $\boldsymbol{\epsilon}$ is scaled by a nonlinear term.

If the target state angular velocity vector $\boldsymbol{\omega}_r$ is non-zero, then the $\boldsymbol{\epsilon}$ differential kinematic equation is written as

$$\dot{\boldsymbol{\epsilon}} = \frac{1}{2}[T]\delta\boldsymbol{\omega} \tag{7.152}$$

where the vector $\delta\boldsymbol{\omega}$ is the error vector in body angular velocities defined as

$$\delta\boldsymbol{\omega} = \boldsymbol{\omega} - \boldsymbol{\omega}_r \tag{7.153}$$

The error angular acceleration vector is found be taking the inertial derivative of Eq. (7.153).

$$\delta\dot{\boldsymbol{\omega}} = \dot{\boldsymbol{\omega}} - \dot{\boldsymbol{\omega}}_r \tag{7.154}$$

Note that here $\delta\boldsymbol{\omega}$ is treated as a vector, not as a 3×1 matrix. Therefore this $\delta\dot{\boldsymbol{\omega}}$ expression is different then the local derivative expression in Eq. (7.35) used in deriving the Lyapunov feedback control laws. After differentiating Eq. (7.152) the vector $\ddot{\boldsymbol{\epsilon}}$ is found.

$$\ddot{\boldsymbol{\epsilon}} = \frac{1}{2}[T]\delta\dot{\boldsymbol{\omega}} + \frac{1}{2}[\dot{T}]\delta\boldsymbol{\omega} \tag{7.155}$$

Substituting these $\dot{\boldsymbol{\epsilon}}$ and $\ddot{\boldsymbol{\epsilon}}$ expressions into the linear closed loop dynamics in Eq. (7.134), and making use of Eq. (7.154), the body angular acceleration vector $\dot{\boldsymbol{\omega}}$ is found to be

$$\dot{\boldsymbol{\omega}} = \dot{\boldsymbol{\omega}}_r - P\delta\boldsymbol{\omega} - 2\left(K - \frac{\delta\omega^2}{4}\right)\frac{\boldsymbol{\epsilon}}{\beta_0} \tag{7.156}$$

The $\dot{\boldsymbol{\omega}}$ in Eq. (7.151) which leads to an integral feedback control law can be modified in a similar manner to track a reference rotation $\boldsymbol{\omega}_r(t)$.

Example 7.14: Instead of using the Euler parameter vector $\boldsymbol{\epsilon}$ as the attitude measure, this example will use the MRP vector $\boldsymbol{\sigma}$. Assume that the closed loop attitude error dynamics are desired to be of the stable second order form

$$\ddot{\boldsymbol{\sigma}} + P\dot{\boldsymbol{\sigma}} + K\boldsymbol{\sigma} = 0$$

where P and K are positive scalar feedback gains. The kinematic differential equation of the modified Rodrigues parameters is expressed as

$$\dot{\boldsymbol{\sigma}} = \frac{1}{4}[B(\boldsymbol{\sigma})]\boldsymbol{\omega}$$

where the matrix $[B]$ is defined in Eq. (3.150). To introduce the $\dot{\boldsymbol{\omega}}$ term, the time derivative of this kinematic equation is taken to produce the exact relation

$$\ddot{\boldsymbol{\sigma}} = \frac{1}{4}[B]\dot{\boldsymbol{\omega}} + \frac{1}{4}[\dot{B}]\boldsymbol{\omega}$$

Substituting the expressions for $\dot{\boldsymbol{\sigma}}$ and $\ddot{\boldsymbol{\sigma}}$ into the desired linear closed loop dynamical equations, the following constraint condition is found.

$$\ddot{\boldsymbol{\sigma}} + P\dot{\boldsymbol{\sigma}} + K\boldsymbol{\sigma} = \frac{1}{4}[B]\left(\dot{\boldsymbol{\omega}} + P\boldsymbol{\omega} + [B]^{-1}\left([\dot{B}]\boldsymbol{\omega} + 4K\boldsymbol{\sigma}\right)\right) = 0$$

Since for $|\boldsymbol{\sigma}| \leq 1$ the matrix $[B]$ is always invertible, then the constraint on $\dot{\boldsymbol{\omega}}$ is

$$\dot{\boldsymbol{\omega}} + P\boldsymbol{\omega} + [B]^{-1}\left([\dot{B}]\boldsymbol{\omega} + 4K\boldsymbol{\sigma}\right) = 0$$

Using the vector product definition of the $[B]$ matrix in Eq. (3.150), the matrix product $[\dot{B}]\boldsymbol{\omega}$ is expressed as

$$[\dot{B}]\boldsymbol{\omega} = \frac{1}{2}\Big(2\boldsymbol{\sigma}^T\boldsymbol{\omega}\left(1-\sigma^2\right)\boldsymbol{\omega} - \left(1+\sigma^2\right)\omega^2\boldsymbol{\sigma} - 4\boldsymbol{\sigma}^T\boldsymbol{\omega}[\tilde{\boldsymbol{\omega}}]\boldsymbol{\sigma} + 4\left(\boldsymbol{\sigma}^T\boldsymbol{\omega}\right)^2\boldsymbol{\sigma}\Big)$$

Using this expression along with the analytic inverse of matrix $[B]$ given in Eq. (3.152) as

$$[B]^{-1} = \frac{1}{\left(1+\sigma^2\right)^2}[B]^T$$

allows the body angular acceleration vector $\dot{\boldsymbol{\omega}}$ constraint to be reduced to the remarkably simple form

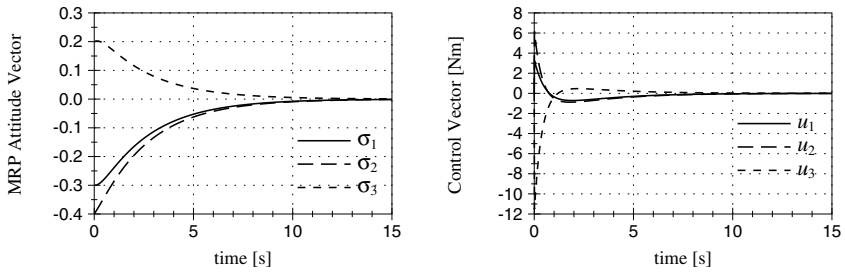
$$\dot{\boldsymbol{\omega}} = -P\boldsymbol{\omega} - \left(\boldsymbol{\omega}\boldsymbol{\omega}^T + \left(\frac{4K}{1+\sigma^2} - \frac{\omega^2}{2}\right)I_{3\times 3}\right)\boldsymbol{\sigma} \tag{7.157}$$

This MRP feedback control is only slightly more complicated than the angular acceleration associated with the Gibbs vector control in Eq. (7.148). However, since whenever $\sigma^2 > 1$ the MRPs are switched to their corresponding shadow set, this $\dot{\boldsymbol{\omega}}$ yields a globally, asymptotically stable feedback control law, whereas Eq. (7.148) is singular at ± 180 degree rotations about any axis.

Table 7.4: Linear MRP Closed Loop Dynamics Numerical Simulation Parameters

Parameter	Value	Units
I_1	30.0	kg-m^2
I_2	20.0	kg-m^2
I_3	10.0	kg-m^2
$\boldsymbol{\sigma}(t_0)$	$[-0.30 \ -0.40 \ 0.20]$	
$\boldsymbol{\omega}(t_0)$	[0.20 0.20 0.20]	rad/sec
$[P]$	3.0	kg-m^2/sec
K	1.0	kg-m^2/sec^2

The following numerical example illustrates the linear closed loop dynamics of this nonlinear feedback control law. The simulation parameters are shown in Table 7.4 and the resulting reorientation is shown in Figure 7.9. No external disturbances were included in this simulation. Figure 7.9(i) shows the MRP attitude vector $\boldsymbol{\sigma}$ components. Their behavior can easily be verified by



(i) Attitude Vector $\boldsymbol{\sigma}$ **(ii)** Control Vector $\boldsymbol{u}$

Figure 7.9: Linear Closed-Loop Dynamics using the MRP Vector

solving the linear differential equation $\ddot{\boldsymbol{\sigma}} + P\dot{\boldsymbol{\sigma}} + K\boldsymbol{\sigma} = 0$ for the given initial conditions. Figure 7.9(ii) shows the corresponding nonlinear control $\boldsymbol{u}$.

7.6 Reaction Wheel Control Devices

Instead of using propellant expelling thrusters, often spacecraft rotational maneuvers are performed using some type of momentum exchange devices. The two most common such devices are the Reaction Wheels (RWs) and the Control Moment Gyroscopes (CMGs). Both are electrically powered and are thus well suited for long-duration missions. Reaction wheels are body fixed disks which are spun up or down to exert a torque onto the spacecraft. They have a relatively simple construction and are cheaper to produce than CMGs. However, their torque output is rather small compared to the torque output of a single-axis CMG.

This section develops feedback control laws that control a spacecraft containing reaction wheel control devices. The following section will deal with CMG and VSCMG control and steering laws. Assume a rigid spacecraft has N reaction wheels attached. Each RW spin axis is denoted through the body fixed vector $\hat{\boldsymbol{g}}_{s_i}$. The equations of motion for this system were developed in Chapter 4 within Example 4.5 and are repeated here for the reader's convenience.

$$[I_{RW}]\dot{\boldsymbol{\omega}} = -[\tilde{\boldsymbol{\omega}}]\left([I_{RW}]\boldsymbol{\omega} + [G_s]\boldsymbol{h}_s\right) - [G_s]\boldsymbol{u}_s + \boldsymbol{L} \tag{7.158}$$

Note that the inertia matrix $[I_{RW}]$ is fixed as seen by the body frame $\mathcal{B}$ and is defined as

$$[I_{RW}] = [I_s] + \sum_{i=1}^{N}\left(J_{t_i}\hat{\boldsymbol{g}}_{t_i}\hat{\boldsymbol{g}}_{t_i}^T + J_{g_i}\hat{\boldsymbol{g}}_{g_i}\hat{\boldsymbol{g}}_{g_i}^T\right) \tag{7.159}$$

The RW motor torques u_{s_i} are given by

$$u_{s_i} = J_{s_i}\left(\dot{\Omega}_i + \hat{\boldsymbol{g}}_{s_i}^T\boldsymbol{\omega}\right) \tag{7.160}$$

The angular vector components of $\boldsymbol{h}_s$ are given by

$$h_{s_i} = J_{s_i}\left(\omega_{s_i} + \Omega_i\right) \tag{7.161}$$

Let the angular velocity error vector $\delta\boldsymbol{\omega}$ be defined as

$$\delta\boldsymbol{\omega} = \boldsymbol{\omega} - \boldsymbol{\omega}_r \tag{7.162}$$

where $\boldsymbol{\omega}_r$ is the reference angular velocity vector. The attitude error between the current body and reference frames is chosen to be expressed through the MRP vector $\boldsymbol{\sigma}$. Other attitude parameterizations could have been used here instead too. To provide a positive definite scalar measure of both the attitude and angular velocity tracking error, we use Eqs. (7.33) and (7.67) to construct our Lyapunov function V.

$$V(\boldsymbol{\sigma}, \delta\boldsymbol{\omega}) = \frac{1}{2}\delta\boldsymbol{\omega}^T[I_{RW}]\delta\boldsymbol{\omega} + 2K\ln\left(1 + \boldsymbol{\sigma}^T\boldsymbol{\sigma}\right) \tag{7.163}$$

Note that the components of $\delta\boldsymbol{\omega}$ and $[I_{RW}]$ are taken in the $\mathcal{B}$ frame. After setting the derivative of V equal to the negative semi-definite function

$$\dot{V} = -\delta\boldsymbol{\omega}[P]\delta\boldsymbol{\omega} \tag{7.164}$$

the following closed loop dynamics are obtained.

$$[I_{RW}]\frac{{}^{\mathcal{B}}d}{dt}\left(\delta\boldsymbol{\omega}\right) = -K\boldsymbol{\sigma} - [P]\delta\boldsymbol{\omega} \tag{7.165}$$

After substituting the equations of motion in Eq. (7.158) and making use of Eqs. (7.35) and (7.162), the RW motor torque vector is defined through the constraint

$$\begin{aligned}[G_s]\boldsymbol{u}_s = K\boldsymbol{\sigma} + [P]\delta\boldsymbol{\omega} - [\tilde{\boldsymbol{\omega}}]\left([I_{RW}]\boldsymbol{\omega} + [G_s]\boldsymbol{h}_s - \boldsymbol{\omega}_r\right) \\ - [I_{RW}]\left(\dot{\boldsymbol{\omega}}_r - \boldsymbol{\omega}\times\boldsymbol{\omega}_r\right) + \boldsymbol{L}\end{aligned} \tag{7.166}$$

Let us combine the terms of the right hand side of Eq. (7.166) to form the required torque vector $\boldsymbol{L}_r$ that the RW cluster must produce. The control constraint is then expressed compactly as

$$[G_s]\boldsymbol{u}_s = \boldsymbol{L}_r \tag{7.167}$$

If at least three or more RWs are present and their spin axes $\hat{\boldsymbol{g}}_{s_i}$ span the entire three-dimensional space, then the RW cluster will be able to produce in principle any required torque vector $\boldsymbol{L}_r$. For the special case where the spacecraft only

contains three RWs, and each is aligned with one of the principal body axes, then $[G_s] = I_{3\times 3}$ and the RW motor torque vector is simply given by

$$\boldsymbol{u}_s = \boldsymbol{L}_r \tag{7.168}$$

For the more general case where a redundant set of RW are present, the vectors $\boldsymbol{u}_s$ and $\boldsymbol{L}_r$ will not have a one-to-one correspondence. Here there actually exists an infinite number of u_{s_i} combinations that will produce the required torque vector $\boldsymbol{L}_r$. One common method used to solve for the actual RW motor torques is to use a minimum norm inverse.

$$\boldsymbol{u}_s = [G_s]^T \left([G_s][G_s]^T\right)^{-1} \boldsymbol{L}_r \tag{7.169}$$

This method provides at any instance of time the smallest set of RW motor torques that combined produce $\boldsymbol{L}_r$. Even though $\dot{V}$ is only negative semi-definite, it was shown in Eq. (7.87) that this control law is indeed asymptotic stabilizing. A major advantage of this RW control law when compared to those of other moment exchange devices is that it is relatively simple in nature and easy to implement. The limitation of RWs include the relatively small torque produced by the device and the problems of having RW saturate. Each RW has an upper limit on how fast its rotor can be safely spun up. This rotor rate range constrains the torque that can be produced. Further, the faster a RW is spinning, the more power is consumed to produce a required torque. The kinetic work rate for a spacecraft with a system of RW can be deduced from the more general work rate in Eq. (4.119).

$$\dot{T} = \boldsymbol{\omega}^T \boldsymbol{L} + \sum_{i=1}^{N} \Omega_i u_{s_i} \tag{7.170}$$

This work rate expression shows clearly that the larger Ω_i is, the larger the work rate of the RW motor torque u_{s_i} will be.

7.7 Variable Speed Control Moment Gyroscopes

Control moment gyroscopes contain a rotor whose spin rate Ω is held constant. By rotating (also referred to as gimbaling) this rotor about some axis other than the spin axis, a gyroscopic torque is produced onto the spacecraft (see Figure 7.10. Single-gimbal CMGs only contain one body fixed gimbal axis. Their major advantage is that for a small torque input about the gimbal axis, a large torque output is produced about the transverse axis. This phenomena is called the *torque amplification effect*. Their drawback is that the corresponding steering law is more complicated than that of the RWs and that singular gimbal angle configurations exist where the required torque is only partially produced, if at all. Dual-gimbal CMGs have a second gimbal axis. These devices don't have as large of a torque amplification effect, but their steering law is less prone to encounter singularities.

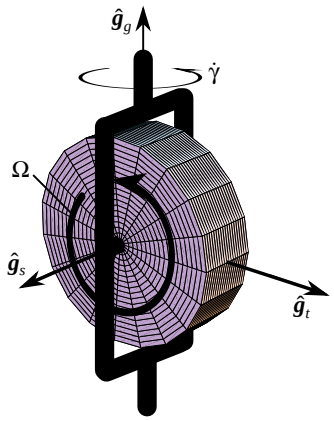


Figure 7.10: Schematic Illustration of a Momentum Exchange Device

The concept of Variable Speed Control Moment Gyroscopes (VSCMGs) are a more recent addition to the family of momentum exchange devices. VSCMGs are essentially single-gimbal CMGs where the rotor spin speed Ω is allowed to be time-varying. Each device is capable of producing two torque outputs if necessary giving a spacecraft more redundancy in case of hardware failure. Nominally these devices are operated as regular CMGs unless the gimbal angles approach a singular configuration, where the RW mode is utilized to ensure that the required torque is produced at all times.

The control and steering laws of the single-gimbal CMGs and the VSCMGs are developed here as one. By holding the RW rotor rate Ω_i constant, the VSCMG steering law is falls back to the classical CMG steering law.

7.7.1 Control Law

The following VSCMG feedback control law is derived using Lyapunov control theory. Given some initial angular velocity and attitude measure, the goal of this control law is to track some reference trajectory defined through the reference angular velocity vector $\boldsymbol{\omega}_r(t)$. By holding the RW spin speeds Ω_i constant this control law, along with the associated steering laws, is easily converted to the conventional CMG feedback control and steering laws. The MRP attitude vector $\boldsymbol{\sigma}$ is chosen again to measure the relative attitude error between the body frame $\mathcal{B}$ and the reference frame $\mathcal{R}$. The body angular velocity error vector $\delta\boldsymbol{\omega}$ is defined in Eq. (7.31) along with its local $\mathcal{B}$ frame derivative defined in Eq. (7.35). Note that the vector components of $\boldsymbol{\omega}_r$ and $\dot{\boldsymbol{\omega}}_r$ are typically given in the $\mathcal{R}$ frame and are assumed to be known.

In designing the feedback control law, it is assumed that estimates of the state vectors $\boldsymbol{\omega}$, $\boldsymbol{\sigma}$, Ω_i and γ_i are available. The following Lyapunov function V is a positive definite, radially unbounded measure of the total system state error relative to the target state $\delta\boldsymbol{\omega} = \boldsymbol{\sigma} = 0$, where K is a scalar attitude feedback

gain.[17]

$$V(\delta\boldsymbol{\omega},\boldsymbol{\sigma}) = \frac{1}{2}\delta\boldsymbol{\omega}^T[I]\delta\boldsymbol{\omega} + 2K\log\left(1+\boldsymbol{\sigma}^T\boldsymbol{\sigma}\right) \tag{7.171}$$

The total spacecraft inertia matrix for this dynamical system was developed in Eq. (4.115) and is expressed as

$$[I] = [I_s] + \sum_{i=1}^{N}\left[J_{s_i}\hat{\boldsymbol{g}}_{s_i}\hat{\boldsymbol{g}}_{s_i}^T + J_{t_i}\hat{\boldsymbol{g}}_{t_i}\hat{\boldsymbol{g}}_{t_i}^T + J_{g_i}\hat{\boldsymbol{g}}_{g_i}\hat{\boldsymbol{g}}_{g_i}^T\right] \tag{7.172}$$

Note that all the body angular velocity vectors and inertia matrices have components taken in the $\mathcal{B}$ frame in Eq. (7.171). Using Eq. (7.68), the Lyapunov function rate $\dot{V}$ is then given by

$$\dot{V} = \delta\boldsymbol{\omega}^T\left([I]\frac{{}^{\mathcal{B}}d}{dt}(\delta\boldsymbol{\omega}) + \frac{1}{2}\frac{{}^{\mathcal{B}}d}{dt}[I]\delta\boldsymbol{\omega} + K\boldsymbol{\sigma}\right) \tag{7.173}$$

Note that the spacecraft inertia matrix $[I]$ must now be treated as time varying as seen by the $\mathcal{B}$ frame because of the CMG gimbaling. Since the Lyapunov function V is a scalar quantity, taking its derivative simply involves taking the derivatives of its scalar components. Since the inertia matrix $[I]$ and $\delta\boldsymbol{\omega}$ have components taken in the $\mathcal{B}$ frame, their derivatives are taken as seen by the $\mathcal{B}$ frame. Using the inertia matrix definition in Eq. (7.172) and the $\mathcal{B}$ frame derivatives of the gimbal frame unit vectors in Eq. (4.98), the $\mathcal{B}$ frame derivative of $[I]$ is

$$\frac{{}^{\mathcal{B}}d}{dt}[I] = \sum_{i=1}^{N}\gamma_i\left(J_{s_i} - J_{t_i}\right)\left(\hat{\boldsymbol{g}}_{s_i}\hat{\boldsymbol{g}}_{t_i}^T + \hat{\boldsymbol{g}}_{t_i}^T\hat{\boldsymbol{g}}_{s_i}^T\right) \tag{7.174}$$

To guarantee stability of the closed-loop system, the Lyapunov rate function is set equal to the negative semi-definite function $\dot{V} = -\delta\boldsymbol{\omega}^T[P]\delta\boldsymbol{\omega}$, which, when combined with Eq. (7.173), leads to the stability constraint:

$$[I]\delta\dot{\boldsymbol{\omega}} = -K\boldsymbol{\sigma} - [P]\delta\boldsymbol{\omega} - \frac{1}{2}\frac{{}^{\mathcal{B}}d}{dt}[I]\delta\boldsymbol{\omega} \tag{7.175}$$

After substituting Eqs. (4.117) and (7.174) into Eq. (7.175), the following stability constraint is obtained.

$$\begin{aligned}\sum_{i=1}^{N}J_{s_i}\dot{\Omega}_i\hat{\boldsymbol{g}}_{s_i} + \sum_{i=1}^{N}J_{g_i}\ddot{\gamma}_i\hat{\boldsymbol{g}}_{g_i} + \sum_{i=1}^{N}\dot{\gamma}_i\Big(&J_{s_i}\Omega_i\hat{\boldsymbol{g}}_{t_i} + \frac{1}{2}\left(J_{s_i}-J_{t_i}\right)\left(\omega_{t_i}\hat{\boldsymbol{g}}_{s_i} + \omega_{s_i}\hat{\boldsymbol{g}}_{t_i}\right)\\ &+ J_{g_i}\left(\omega_{t_i}\hat{\boldsymbol{g}}_{s_i} - \omega_{s_i}\hat{\boldsymbol{g}}_{t_i}\right) + \frac{1}{2}\left(J_{s_i}-J_{t_i}\right)\left(\hat{\boldsymbol{g}}_{s_i}\hat{\boldsymbol{g}}_{t_i}^T\boldsymbol{\omega}_r + \hat{\boldsymbol{g}}_{t_i}\hat{\boldsymbol{g}}_{s_i}^T\boldsymbol{\omega}_r\right)\Big)\\ &= K\boldsymbol{\sigma} + [P]\delta\boldsymbol{\omega} + \boldsymbol{L} - [\tilde{\boldsymbol{\omega}}][I]\boldsymbol{\omega} - [I]\left(\dot{\boldsymbol{\omega}}_r - [\tilde{\boldsymbol{\omega}}]\boldsymbol{\omega}_r\right)\\ &\quad - \sum_{i=1}^{N}J_{s_i}\left(\Omega_i\omega_{g_i}\hat{\boldsymbol{g}}_{t_i} - \Omega_i\omega_{t_i}\hat{\boldsymbol{g}}_{g_i}\right)\end{aligned} \tag{7.176}$$

To express this condition in a more compact and useable form, let us define the following 3xN matrices, where all components are taken in the $\mathcal{B}$ frame:

$$[D_0] = [\cdots \hat{\boldsymbol{g}}_{s_i} J_{s_i} \cdots] \tag{7.177a}$$

$$[D_1] = [\cdots J_{s_i} \left(\left(\Omega_i + \frac{1}{2}\omega_{s_i} \right) \hat{\boldsymbol{g}}_{t_i} + \frac{1}{2}\omega_{t_i} \hat{\boldsymbol{g}}_{s_i} \right) \cdots] \tag{7.177b}$$

$$[D_2] = [\cdots \frac{1}{2} J_{t_i} \left(\omega_{t_i} \hat{\boldsymbol{g}}_{s_i} + \omega_{s_i} \hat{\boldsymbol{g}}_{t_i} \right) \cdots] \tag{7.177c}$$

$$[D_3] = [\cdots J_{g_i} \left(\omega_{t_i} \hat{\boldsymbol{g}}_{s_i} - \omega_{s_i} \hat{\boldsymbol{g}}_{t_i} \right) \cdots] \tag{7.177d}$$

$$[D_4] = [\cdots \frac{1}{2} \left(J_{s_i} - J_{t_i} \right) \left(\hat{\boldsymbol{g}}_{s_i} \hat{\boldsymbol{g}}_{t_i}^T \boldsymbol{\omega}_r + \hat{\boldsymbol{g}}_{t_i} \hat{\boldsymbol{g}}_{s_i}^T \boldsymbol{\omega}_r \right) \cdots] \tag{7.177e}$$

$$[B] = [\cdots \hat{\boldsymbol{g}}_{g_i} J_{g_i} \cdots] \tag{7.177f}$$

Let $\dot{\boldsymbol{\Omega}}$, $\ddot{\boldsymbol{\gamma}}$ and $\dot{\boldsymbol{\gamma}}$ be Nx*1* vectors whose i-th element contains the respective VSCMG angular velocity or acceleration or RW spin rate. The stability constraint in Eq. (7.176) then is expressed as[12]

$$[D_0]\dot{\boldsymbol{\Omega}} + [B]\ddot{\boldsymbol{\gamma}} + [D]\dot{\boldsymbol{\gamma}} = \boldsymbol{L}_r \tag{7.178}$$

where $[D] = ([D_1] - [D_2] + [D_3] + [D_4])$ and the required torque vector $\boldsymbol{L}_r$ is defined to be

$$\boldsymbol{L}_r = K\boldsymbol{\sigma} + [P]\delta\boldsymbol{\omega} + \boldsymbol{L} - [\tilde{\boldsymbol{\omega}}][I]\boldsymbol{\omega} - [I]\left(\dot{\boldsymbol{\omega}}_r - [\tilde{\boldsymbol{\omega}}]\boldsymbol{\omega}_r \right) - \sum_{i=1}^{N} J_{s_i} \left(\Omega_i \omega_{g_i} \hat{\boldsymbol{g}}_{t_i} - \Omega_i \omega_{t_i} \hat{\boldsymbol{g}}_{g_i} \right) \tag{7.179}$$

Dropping the $[D_0]\dot{\boldsymbol{\Omega}}$ term in Eq. (7.178), the standard single-gimbal CMG stability constraint is retrieved as it is developed in Ref. 26. Note that the formulation presented here does not require any matrix multiplications of sparse matrices and the effects of the individual VSCMG inertia terms are immediately evident. The condition in Eq. (7.178) only guarantees global stability in the sense of Lyapunov for the states $\delta\boldsymbol{\omega}$ and $\boldsymbol{\sigma}$, since $\dot{V}$ was only set to be negative semi-definite, not negative definite. However, the negative semi-definite Lyapunov rate expression does show that $\delta\boldsymbol{\omega} \to 0$ as time goes to infinity. To prove that the stability constraint in Eq. (7.178) guarantees asymptotic stability of all states including $\boldsymbol{\sigma}$, once again the higher time derivatives of V must be investigated. For this dynamical system $\dot{V}$ is zero whenever $\delta\boldsymbol{\omega}$ is zero. Using Eq. (7.175) and setting $\delta\boldsymbol{\omega} = 0$, the third derivative of the Lyapunov function V is found to be the first, non-zero higher order derivative and is expressed as

$$\frac{d^3}{d^3t}V = -K^2\boldsymbol{\sigma}^T \left([I]^{-1} \right)^T [P][I]^{-1}\boldsymbol{\sigma} \tag{7.180}$$

which is a negative definite quantity since both $[I]$ and $[P]$ are positive definite matrices. Therefore the stability constraint in Eq. (7.178) does guarantee global asymptotic stability.

If the desired spacecraft trajectory is a stationary attitude, then the reference body angular velocity vector $\boldsymbol{\omega}_r$ is zero. For these rest-to-rest or motion-to-rest type maneuvers, the feedback control law in Eq. (7.178) can be greatly simplified. After substituting Eqs. (4.117) and (7.174) into Eq. (7.173) and rearranging some terms, the Lyapunov rate function $\dot{V}$ is expressed as

$$\dot{V} = -\boldsymbol{\omega}^T \Bigg([\tilde{\boldsymbol{\omega}}][I]\boldsymbol{\omega} + \sum_{i=1}^{N} J_{s_i}\dot{\Omega}_i \hat{\boldsymbol{g}}_{s_i} + \sum_{i=1}^{N} J_{g_i}\ddot{\gamma}_i \hat{\boldsymbol{g}}_{g_i} + \sum_{i=1}^{N} J_{s_i}\dot{\gamma}_i \left(\Omega_i + \frac{1}{2}\left(\omega_{t_i}\hat{\boldsymbol{g}}_{s_i} + \omega_{s_i}\hat{\boldsymbol{g}}_{t_i}\right)\right) - K\boldsymbol{\sigma} - \boldsymbol{L} + \sum_{i=1}^{N} J_{s_i}\Omega\left(\omega_{g_i}\hat{\boldsymbol{g}}_{t_i} - \omega_{t_i}\hat{\boldsymbol{g}}_{g_i}\right) - \sum_{i=1}^{N} J_{t_i}\dot{\gamma}_i \frac{1}{2}\left(\omega_{s_i}\hat{\boldsymbol{g}}_{t_i} + \omega_{t_i}\hat{\boldsymbol{g}}_{s_i}\right) + \sum_{i=1}^{N} J_{g_i}\dot{\gamma}_i\left(\omega_{t_i}\hat{\boldsymbol{g}}_{g_i} - \omega_{g_i}\hat{\boldsymbol{g}}_{t_i}\right)\Bigg) \quad (7.181)$$

For this regulator problem, several terms in Eq. (7.181) can be shown to be nonworking and are neglected in the resulting feedback control law. Setting $\dot{V} = -\boldsymbol{\omega}^T[P]\boldsymbol{\omega}$ and performing further algebraic manipulations, the simplified stability constraint for the regulator problem is found to be

$$\sum_{i=1}^{N} J_{s_i}\dot{\Omega}_i\hat{\boldsymbol{g}}_{s_i} + \sum_{i=1}^{N} J_{g_i}\ddot{\gamma}_i\hat{\boldsymbol{g}}_{g_i} + \sum_{i=1}^{N} J_{s_i}\dot{\gamma}_i\left(\Omega_i + \omega_{s_i}\right)\hat{\boldsymbol{g}}_{t_i} - \sum_{i=1}^{N} J_{t_i}\omega_{s_i}\dot{\gamma}_i\hat{\boldsymbol{g}}_{t_i} = K\boldsymbol{\sigma} + [P]\boldsymbol{\omega} + L = \boldsymbol{L}_r \quad (7.182)$$

Note that $\boldsymbol{L}_r$ defined in Eq. (7.182) is a simplified version of the one defined in Eq. (7.179). Making use of the $3 \times N$ matrices

$$[D_0] = [\cdots \hat{\boldsymbol{g}}_{s_i} J_{s_i} \cdots] \quad (7.183a)$$

$$[D_1] = [\cdots \hat{\boldsymbol{g}}_{t_i} J_{s_i}\left(\Omega_i + \omega_{s_i}\right) \cdots] \quad (7.183b)$$

$$[D_2] = [\cdots \hat{\boldsymbol{g}}_{t_i} J_{t_i}\omega_{s_i} \cdots] \quad (7.183c)$$

$$[B] = [\cdots \hat{\boldsymbol{g}}_{g_i} J_{g_i} \cdots] \quad (7.183d)$$

the stability constraint is written in the following compact form[12, 27]

$$[D_0]\dot{\boldsymbol{\Omega}} + [B]\ddot{\boldsymbol{\gamma}} + [D]\dot{\boldsymbol{\gamma}} = \boldsymbol{L}_r \quad (7.184)$$

where $[D] = ([D_1] - [D_2])$. Note that the matrices $[D_0]$ and $[B]$ are the same as with the general feedback law. The matrices $[D_1]$ and $[D_2]$ are simplified and have columns which solely depend on the $\hat{\boldsymbol{g}}_{t_i}$ directions. The matrices $[D_3]$ and $[D_4]$ do not appear at all in this control law. Since this regulator control law is a specialization of the more general trajectory tracking control law, it too is globally asymptotically stabilizing.

7.7.2 Velocity Based Steering Law

Note that the stability constraints in Eqs. (7.178) and (7.184) do not contain the physical control torques u_{s_i} and u_{g_i} explicitly. Instead, only gimbal rates

and accelerations and RW accelerations appear. This will lead to a steering law that determines the required time history of $\boldsymbol{\gamma}$ and $\boldsymbol{\Omega}$ such that Eq. (7.178) is satisfied. The reason for this is two fold. First, currently available CMGs typically require the gimbal rate vector $\dot{\boldsymbol{\gamma}}$ as the input, not the actual physical torque vector $\boldsymbol{u}_g$. Secondly, writing Eqs. (7.178) and (7.184) in terms of the torque vectors $\boldsymbol{u}_s$ and $\boldsymbol{u}_g$ and then solving for these would lead to a control law that is equivalent to solving Eq. (7.178) directly for the gimbal acceleration vector $\ddot{\boldsymbol{\gamma}}$. As has been pointed out in Ref. 26, this has been found to give a very undesirable control law with excessive gimbal rates. A physical reason for this is that such control laws provide the required control torque mainly through the $[B]\ddot{\boldsymbol{\gamma}}$ term. In this setup the CMGs are essentially being used as RWs and the potential torque amplification effect in not being exploited. Because CMG gimbal inertias J_g are typically small compared to their spin inertia J_s, the corresponding $[B]$ will also be very small which leads to very large $\ddot{\boldsymbol{\gamma}}$ vectors.

To take advantage of the potential torque amplification effect, most of the required control torque vector $\boldsymbol{L}_r$ should be produced by the larger gyroscopic coupling $[D]\dot{\boldsymbol{\gamma}}$ term. This is why classical CMG steering laws control primarily the $\dot{\boldsymbol{\gamma}}$ vector and not $\ddot{\boldsymbol{\gamma}}$. For the VSCMGs it is desirable to have the required torque $\boldsymbol{L}_r$ be produced by a combination of the $\dot{\boldsymbol{\Omega}}$ and $\dot{\boldsymbol{\gamma}}$ terms in Eqs. (7.178) and (7.184). To simplify the further development, we assume that the final angular velocity is zero Therefore the stability constraint in Eq. (7.184) is used. However, the results are equally valid for the trajectory tracking control law. To force the required torque to be produced by the gimbal rates, the terms containing the transverse and gimbal VSCMG inertias are ignored at this level. Eq. (7.184) then becomes

$$[D_0]\dot{\boldsymbol{\Omega}} + [D_1]\dot{\boldsymbol{\gamma}} = \boldsymbol{L}_r \tag{7.185}$$

Comparing the $[D_1]$ matrix to that of conventional CMG steering laws of the form

$$[D_1]\dot{\boldsymbol{\gamma}} = \boldsymbol{L}_r \tag{7.186}$$

such as ones found in Ref. 26, it is evident that an extra $\hat{\boldsymbol{g}}_t J_s \omega_s$ term is present in the definition of the $[D_1]$ matrix in the VSCMG formulation. This term is neglected in the standard CMG formulation since it can be assumed that ω_s will typically be much smaller than Ω. However, since for a VSCMG the RW spin speed Ω is variable, this assumption is no longer justified for the more general VSCMG case and this term is retained in this formulation. To solve the conventional CMG feedback control constraint in Eq. (7.186) for the gimbal rates, a minimum norm inverse is typically used. However, if the rank of the $[D_1]$ matrix drops below 3, then a steering law singularity where the minimum norm inverse may not be possible mathematically. To operate mathematically in the neighborhood of these singular gimbal configurations, Nakamura and Hanafusa introduce in Ref. 28 a modified minimum norm inverse

$$\dot{\boldsymbol{\gamma}} = [D_1]^T \left([D_1][D_1]^T + \alpha I_{3\times 3}\right)^{-1} \boldsymbol{L}_r \tag{7.187}$$

The scalar parameter is only non-zero in the neighborhood of a singularity and is typically very small in magnitude. This allows the CMG steering law to produce gimbal rates even in the mathematically ill-conditioned singular neighborhoods. The draw-back of this method is that the resulting torque produced by the CMG cluster is not precisely equal to the required torque vector $\boldsymbol{L}_r$. This will cause path deviations of the spacecraft from the prescribed trajectories. Also, the modified minimum norm inverse does not avoid the problem of having a gimbal lock. If the $\boldsymbol{L}_r$ is perpendicular to the range of $[D_1]$, then a zero gimbal rate vector is produced. The gimbal effectively remain locked in this configuration, producing no effective torque, until the required torque vector $\boldsymbol{L}_r$ is changed somehow.

By finding a steering law for the VSCMG case, we avoid many of the conventional CMG singularities by taking into account that the RW rotor speeds are allowed to be time varying. For notational convenience, we introduce the *2N*x*1* state vector $\boldsymbol{\eta}$

$$\boldsymbol{\eta} = \begin{bmatrix} \boldsymbol{\Omega} \\ \boldsymbol{\gamma} \end{bmatrix} \tag{7.188}$$

and the *3*x*2N* matrix $[Q]$

$$[Q] = \begin{bmatrix} D_0 \vdots D_1 \end{bmatrix} \tag{7.189}$$

Eq. (7.185) is then written compactly as

$$[Q]\dot{\boldsymbol{\eta}} = \boldsymbol{L}_r \tag{7.190}$$

Note that each column of the $[D_0]$ matrix is a scalar multiple of the $\hat{\boldsymbol{g}}_{s_i}$ vectors, while each column of $[D_1]$ is a scalar multiple of the $\hat{\boldsymbol{g}}_{t_i}$ vectors. In the classical 4 single-gimbal CMG cluster, singular gimbal configurations are encountered whenever the rank of $[D_1]$ is less than 3. This occurs whenever the $\hat{\boldsymbol{g}}_{t_i}$ axes no longer span the three-dimensional space, but form a plane or a line. Any required torque which does not lie perfectly in this plane or line cannot be generated exactly by the CMG cluster and the spacecraft would deviate from the desired trajectory. If the required control torque is perpendicular to this plane, then the CMG cluster produces no effective torque on the spacecraft. This singular behavior is illustrated in Figure 7.11 with two CMGs gimbaling to produce a constant torque vector $\boldsymbol{L}_r$. Since each CMG produces a torque about its transverse axis, the two wheels must be gimbaled symmetrically and at the same rate to produce the indicated required torque vector $\boldsymbol{L}_r$. As both transverse axes rotate toward perpendicular orientations relative to $\boldsymbol{L}_r$, the associated gimbal rates become exceedingly large to produce the required torque. This is referred to as operating in the neighborhood of a singular configuration. If both transverse axes are perpendicular to $\boldsymbol{L}_r$, then no torque is produced (referred to as gimbal lock).

These singular configurations *can never occur* with a VSCMG steering since the rank of the $[Q]$ matrix will never be less than 3. Since the $\hat{\boldsymbol{g}}_{s_i}$ vectors are

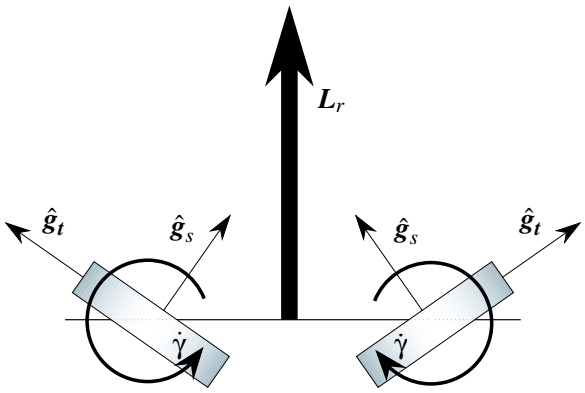


Figure 7.11: Dual CMG System Encountering a Singularity

perpendicular to the $\hat{\boldsymbol{g}}_{t_i}$ vectors, even when all the transverse axes are coplanar, there will always be at least one spin axis that is not in this plane. Therefore the columns of $[Q]$ will *always* span the entire three-dimensional space as long as at least 2 or more VSCMGs are used with distinct $\hat{\boldsymbol{g}}_{g_i}$ vectors. We mention, however, that while a singularity does not occur, this does not imply that other difficulties, such as wheel saturation, will not be encountered occasionally.

Since the $[Q]$ matrix will never be rank deficient, a minimum norm solution for $\dot{\boldsymbol{\eta}}$ can be obtained using the standard Moore-Penrose inverse of Eq. (7.190). However, since ideally the VSCMGs are to act like classical CMGs away from single-gimbal CMG singular configurations, a weighted pseudo inverse is recommended instead.[29] Let $[W]$ be a $2N$x$2N$ diagonal matrix

$$[W] = diag\{W_{s_1}, \ldots, W_{s_N}, W_{g_1}, \ldots, W_{g_N}\} \tag{7.191}$$

where W_{s_i} and W_{g_i} are the weights associated with how active each of the RW and CMG modes are. Setting a weight to zero effectively turns that particular mode off. The larger the weight, the more important that mode is in the VSCMG steering law. The desired $\dot{\boldsymbol{\eta}}$ is then found through[27]

$$\dot{\boldsymbol{\eta}} = \begin{bmatrix} \dot{\boldsymbol{\Omega}} \\ \dot{\boldsymbol{\gamma}} \end{bmatrix} = [W][Q]^T \left([Q][W][Q]^T\right)^{-1} \boldsymbol{L}_r \tag{7.192}$$

Note that there is no need here to introduce a modified pseudo-inverse as Nakamura and Hanafusa did in developing the singularity robustness steering law in Ref. 28. To achieve the desired VSCMG behavior, the weights are made dependent on the proximity to a single-gimbal CMG singularity. One method to measure the proximity to a singularity is to compute the non-dimensional scalar factor δ defined as

$$\delta = \det \frac{1}{h^2} \left([D_1][D_1]^T\right) \tag{7.193}$$

where $\bar{h}$ is a nominal RW angular momentum. With classical CMG configurations, typically each CMG has the same spin axis angular momentum magnitude h. This quantity can easily be factored out of $[D_1]$ to render δ non-dimensional. However, since VSCMGs have potentially time varying spin axis angular momentum magnitudes, one has to divide $[D_1]$ by a *nominal* spin axis angular momentum magnitude $\bar{h}$ to render δ non-dimensional. As the gimbals approach a singular CMG configuration, this parameter δ goes to zero. The weights W_{s_i} can be defined to as the functions

$$W_{s_i} = W_{s_i}^0 e^{(-\mu\delta)} \tag{7.194}$$

where $W_{s_i}^0$ and μ are positive scalars to be chosen by the control designer. The gains W_{g_i} are simply held constant. Away from CMG singularities, this steering law will have very small weights on the RW mode and essentially perform like a classical single-gimbal CMG. As a singularity is approached, the steering law will start to use the RW mode to ensure that the gimbal rates do not become excessive and that the required control torque $\boldsymbol{L}_r$ is actually produced by the VSCMG cluster.

Two types of CMG singularities are commonly discussed. The simpler type of singularity is when the rank of the $[D_1]$ matrix drops below 3 which is indicated by δ, defined in Eq. (7.193), approaching or becoming zero. The VSCMG velocity steering law in Eq. (7.192) handles temporary rank deficiencies very well. The required control torque is always produced correctly by making use of the addition control authority provided by the RW modes. Another type of singularity is when the required control torque is exactly perpendicular to the span of the transverse VSCMG axis (i.e. $\boldsymbol{L}_r$ is in the nullspace of $[D_1]$). Naturally, this is only possible whenever δ is zero. To measure how close the required torque $\boldsymbol{L}_r$ is to lying in the nullspace of $[D_1]$, the scalar orthogonality index $\mathcal{O}$ is used.[26]

$$\mathcal{O} = \frac{1}{\bar{h}^2} \frac{\boldsymbol{L}_r^T [D_1]^T [D_1] \boldsymbol{L}_r}{||\boldsymbol{L}_r||^2} \tag{7.195}$$

Whenever $\boldsymbol{L}_r$ becomes part of the nullspace of $[D_1]$, then $\mathcal{O}$ will tend towards zero. A classical single-gimbal CMG steering law demands a zero $\dot{\boldsymbol{\gamma}}$ vector with this type of singularity which "locks up" the gimbals produces no effective torque on the spacecraft. The VSCMG steering does not prevent the gimbals from being locked up in these singular orientations; however, the $\boldsymbol{L}_r$ vector is still being produced thanks to the RW mode of the VSCMGs. If a gimbal lock is actually achieved, then without any further changes, such as a change in the required $\boldsymbol{L}_r$, the VSCMG will simply continue the maneuvers acting like pure RWs. Running numerical simulations it was found that unless one starts the simulation in a pure gimbal lock situation, it was very unlikely for the VSCMG steering law to lock up the gimbals. Once a singularity is approached, the RWs are automatically spun up or down which also in return affects the gimbal orientation and lowers the likelihood of having the orthogonality index $\mathcal{O}$ go to zero. However, the the current form this VSCMG steering law makes

no explicit effort to avoid these singular configurations during a maneuver. In essence, once the momentum symmetry (associated with CMG geometry and constant wheel speeds) is destroyed; by virtue of variable wheel speeds, the corresponding singular geometries are also eliminated.

7.7.3 VSCMG Null Motion

To perform a given spacecraft maneuver, there are an infinity of possible CMG configurations that would produce the required torque vector $\boldsymbol{L}_r$. Depending on the torque direction and a given CMG momentum, some of these initial gimbal configurations will encounter CMG singularities during the resulting maneuver while others will not. Vadali et al. show in Ref. 30 a method to compute a preferred set of initial gimbal angles $\boldsymbol{\gamma}(t_0)$ with which the resulting maneuver will not encounter any CMG singularities. The method computes $\boldsymbol{\gamma}(t_0)$ off-line before the maneuver is performed. To reorient the CMG cluster to these preferred gimbal angles, the null motion of the steering law $[D_1]\dot{\boldsymbol{\gamma}} = \boldsymbol{L}_r$ is used. This null motion allows for the gimbals to be reconfigured without applying any torque on the spacecraft. However, the set of gimbal angles between which one can reorient the classical CMGs is very limited, since the internal CMG cluster momentum vector must remain constant during this maneuver. Also, the null motion involves the inverse of the $[D_1][D_1]^T$ matrix which has to be approximated with the singularity robustness inverse whenever the determinant goes to zero. This approximation results in a small torque being applied to the spacecraft itself.

Rearranging VSCMGs instead of CMGs however, there are now twice as many degrees of control available. A common four-CMG pyramid configuration would have eight degrees of control instead of four, and more importantly, instead of having a one dimensional nullspace, we have a five dimensional nullspace. In particular, the CMG angles can be rearranged in a more general manner by also varying the RW spin speed vector $\boldsymbol{\Omega}$. The null motion of the control law in Eq. (7.190) is given by

$$\dot{\boldsymbol{\eta}} = \left[[\hat{W}][Q]^T \left([Q][\hat{W}][Q]^T \right)^{-1} [Q] - [I_{2N\times 2N}] \right] \boldsymbol{d} = [\tau]\boldsymbol{d} \tag{7.196}$$

where $[\hat{W}]$ is another diagonal weight matrix which controls how heavily either the CMG or RW mode is used in this null motion. The vector $\boldsymbol{d}$ provides a direction to which to drive the state vector $\boldsymbol{\eta}$.

Note that the matrix $[\tau]$ is a projection matrix and has the useful property that $[\tau][\tau] = [\tau]$. Let the constant vector $\boldsymbol{\eta}_f$ be a preferred set of $\boldsymbol{\Omega}_f$ and $\boldsymbol{\gamma}_f$. The difference between the current and the preferred VSCMG states is expressed as

$$\Delta\boldsymbol{\eta} = \boldsymbol{\eta} - \boldsymbol{\eta}_f = \begin{pmatrix} \boldsymbol{\Omega} - \boldsymbol{\Omega}_f \\ \boldsymbol{\gamma} - \boldsymbol{\gamma}_f \end{pmatrix} = \begin{pmatrix} \Delta\boldsymbol{\Omega} \\ \Delta\boldsymbol{\gamma} \end{pmatrix} \tag{7.197}$$

The state error vector $\boldsymbol{e}$ is defined as

$$\boldsymbol{e} = [A]\Delta\boldsymbol{\eta} \tag{7.198}$$

where $[A]$ is the diagonal matrix

$$[A] = \begin{bmatrix} a_{RW}[I_{N\times N}] & [0_{N\times N}] \\ [0_{N\times N}] & a_{CMG}[I_{N\times N}] \end{bmatrix} \tag{7.199}$$

The parameters a_{RW} and a_{CMG} are either 1 or 0. If one is set to zero, this means that the resulting null motion will be performed with no preferred set of either $\boldsymbol{\Omega}_f$ or $\boldsymbol{\gamma}_f$. The derivative of $\boldsymbol{e}$ is

$$\dot{\boldsymbol{e}} = [A]\dot{\boldsymbol{\eta}} \tag{7.200}$$

The total error between preferred and actual CMG angular speed states is given through the Lyapunov function

$$V_e(\boldsymbol{e}) = \frac{1}{2}\boldsymbol{e}^T\boldsymbol{e} \tag{7.201}$$

Using Eqs. (7.196) and (7.200), the derivative of the Lyapunov function is

$$\dot{V}_e = \boldsymbol{e}^T\dot{\boldsymbol{e}} = \boldsymbol{e}^T[A][\tau]\boldsymbol{d} \tag{7.202}$$

After setting $\boldsymbol{d} = -k_e\boldsymbol{e}$, where the scalar k_e is a positive gain, and making use of the properties $[A]\boldsymbol{e} = \boldsymbol{e}$ and $[\tau][\tau] = [\tau]$, $\dot{V}_e$ is rewritten as

$$\dot{V}_e = -k_e\boldsymbol{e}^T[\tau][\tau]\boldsymbol{e} \leq 0 \tag{7.203}$$

If the weight matrix $[\hat{W}]$ is written as

$$[\hat{W}] = W[I_{2N\times 2N}] \tag{7.204}$$

then the matrix $[\tau]$ becomes symmetric and $[\tau] = [\tau]^T$. Using this property, the Lyapunov rate is expressed as

$$\dot{V}_e = -k_e\boldsymbol{e}^T[\tau^T[\tau]\boldsymbol{e} = -k_e\left([\tau]\boldsymbol{e}\right)^T[\tau]\boldsymbol{e} \leq 0 \tag{7.205}$$

which is negative semi-definite. Therefore the corresponding VSCMG null motion

$$\dot{\boldsymbol{\eta}} = k_e\left[[Q]^T\left([Q][Q]^T\right)^{-1}[Q] - [I_{N\times N}]\right][A]\begin{pmatrix}\Delta\boldsymbol{\Omega}\\ \Delta\boldsymbol{\gamma}\end{pmatrix} \tag{7.206}$$

is a globally stable motion. Numerical simulations show that the weight matrix $[\hat{W}]$ does not have to be restricted to the form in Eq. (7.204). Different weights can be applied to the individual CMG and RW modes with the resulting VSCMG null motion

$$\dot{\boldsymbol{\eta}} = k_e\left[[\hat{W}][Q]^T\left([Q][\hat{W}][Q]^T\right)^{-1}[Q] - [I_{N\times N}]\right][A]\begin{pmatrix}\Delta\boldsymbol{\Omega}\\ \Delta\boldsymbol{\gamma}\end{pmatrix} \tag{7.207}$$

remaining stable. Note however, that no guarantee as to asymptotic stability can be made. As was the case with the classical single-gimbal CMG null motion, it is still not possible to reorient between any two arbitrary sets of $\boldsymbol{\eta}$ vectors, since the internal momentum vector must be conserved. If the momentum is not conserved, then some torque acts on the spacecraft.

Example 7.15: A rigid spacecraft is equipped with four identical VSCMGs arranged in the standard pyramid configuration as shown in Figure 7.12. While maintaining a constant spacecraft attitude, it is desired to rearrange the current set of asymmetric gimbal angles to a new set of preferred symmetric angles $\boldsymbol{\gamma}_f$.

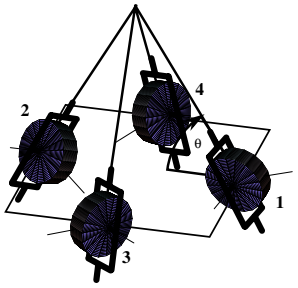


Figure 7.12: VSCMG Pyramid Configuration

This task is attempted both with the conventional CMG null motion and the VSCMG null motion, even though we know in advance that the CMG null motion will not be able to complete this task successfully without exerting a torque on the spacecraft. The VSCMG null motion in Eq. (7.206) is used in this study where equal weights are applied to the RW and CMG modes. The parameter a_{RW} is set to zero, indicating that the RW wheel speeds Ω_i can be changed by the VSCMG null motion as necessary to drive the gimbal angles towards their preferred values.

The simulation parameters are shown in Table 7.5. The spacecraft is at rest and the feedback control law is turned off. Both of the following null motions reconfigured the gimbal angles without changing the spacecraft attitude. The CMG and RW states for both the CMG and VSCMG null motion are illustrated in Figure 7.13. The states of the CMG null motion are indicated with a dashed line, while the states of the VSCMG null motion are depicted with a solid line. As predicted, both simulations remained stable. They differed however in their ability to complete the given task.

Figure 7.13(i) compares the gimbal angles of either null motion. As expected, the CMG null motion does a poor job in driving the initial gimbal angles $\boldsymbol{\gamma}(t_0)$ towards the computed set of symmetric preferred angles $\boldsymbol{\gamma}_f$. While two gimbal angles do approach the ± 45 degree marks, the other two remain far away. However, the VSCMG null motion is able to drive the gimbal angles fairly close to the preferred gimbal angles. The corresponding gimbal rates for either null motion are shown in Figure 7.13(ii) and are small for both cases.

Table 7.5: VSCMG Simulation Parameters

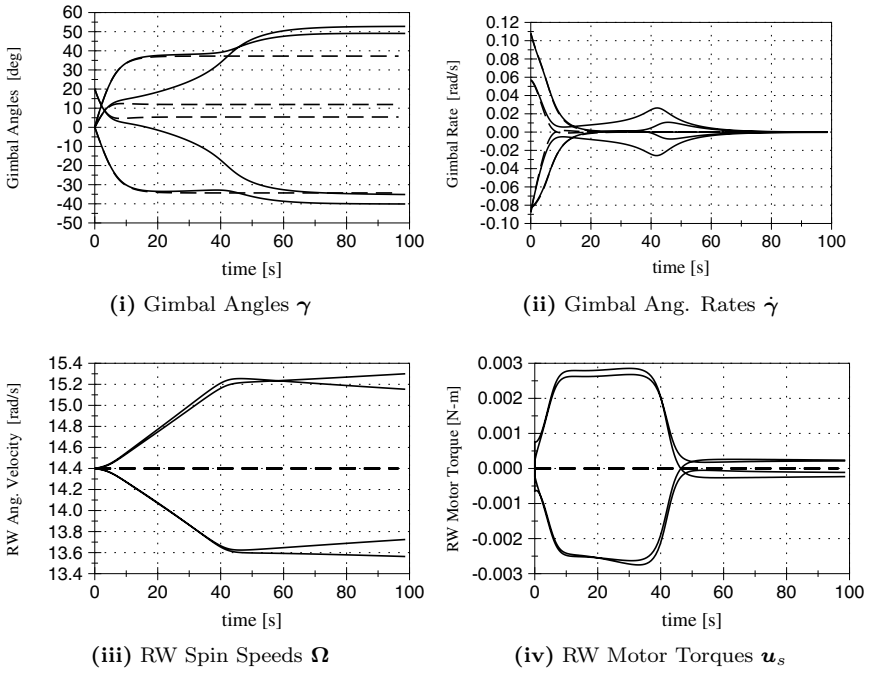
Parameter	Value	Units
I_{s_1}	86.215	kg-m^2
I_{s_2}	85.070	kg-m^2
I_{s_3}	113.565	kg-m^2
θ	54.75	degrees
J_s	0.13	kg-m^2
J_t	0.04	kg-m^2
J_g	0.03	kg-m^2
$\gamma_i(t_0)$	[0 0 0 20]	deg
$\dot{\gamma}_i(t_0)$	[0 0 0 0]	rad
$\Omega(t_0)$	14.4	rad/sec
k_e	0.1	

If the required gimbal rates were too high, then they could be reduced by lowering the gain k_e and therefore slowing down the null motion maneuver.

The VSCMG null motion comes much closer to achieving the task by being able to vary the RW spin speeds as shown in Figure 7.13(iii). During the first 40 seconds of the maneuver the Ω_i are changed steadily to counter the torque produced by rearranging the gimbal angles into a more symmetric configuration. From then on the Ω_i remain relatively constant. The corresponding RW motor torques are shown in Figure 7.13(iv). Again the most active region is during the initial 40 seconds of the maneuver. What is important to note is that the magnitudes of the required RW motor torques are very small. In fact, they are small enough to be feasible with existing RW motors on conventional CMGs. Using this type of null motion therefore only requires a change in the RW feedback control law, and not a change in the hardware design of the CMG itself. Changing the RW feedback law however allows for the gimbals to be rearranged in a much more general manner than is possible with the conventional CMG null motion.

Another use of the VSCMG/CMG null motion to avoid singularities is to use the redundant degrees of freedom to *continuously* rearrange the gimbals such that a singularity index is minimized. If a singular gimbal configuration is approached during a maneuver, then the gimbal angles are automatically rearranged using the VSCMG null motion to reduce the singularity index. This method has the advantage that no preferred $\mathbf{\Omega}$ and $\boldsymbol{\gamma}$ state vectors need to be computed prior to the maneuver. Since it is still theoretically possible to encounter a CMG singularity, this method is very useful when combined with the VSCMG steering law. While the VSCMG steering law may require very large RW motor torques to drive through a CMG singularity, using the VSCMG null motion early on to rearrange the gimbal angles to less singular configurations can lead to drastic reductions in the required RW motor torques.

In order to drive the gimbal configuration towards a "less singular" configuration, a measure of singularity proximity is needed. A gradient type method



(i) Gimbal Angles $\boldsymbol{\gamma}$ **(ii)** Gimbal Ang. Rates $\dot{\boldsymbol{\gamma}}$

(iii) RW Spin Speeds $\boldsymbol{\Omega}$ **(iv)** RW Motor Torques $\boldsymbol{u}_s$

Figure 7.13: CMG and VSCMG Null Motion Simulation

is developed below that provides the necessary state error vectors $\Delta\boldsymbol{\Omega}$ and $\Delta\boldsymbol{\gamma}$ for the VSCMG null motion in Eq. (7.207). The non-dimensional singularity indicator δ could be used here as the singularity measure of the VSCMG null motion steering law. However, since using the gradient method requires taking analytical partial derivatives of the singularity measure with respect to the gimbal angles, this would lead to very complex equations which have to be derived specifically for each physical system.

Instead, the condition number κ of the matrix $[D_1]$ is used as the singularity measure. Using a singular value decomposition (SVD), the $3 \times N$ matrix $[D_1]$ is decomposed as

$$[D_1] = [U][\Sigma][V]^T = \begin{bmatrix} \boldsymbol{u}_1 \boldsymbol{u}_2 \boldsymbol{u}_3 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & 0 & & 0 \\ 0 & \sigma_2 & 0 & \cdots & 0 \\ 0 & 0 & \sigma_3 & & 0 \end{bmatrix} \begin{bmatrix} \\ \boldsymbol{v}_1 \cdots \boldsymbol{v}_N \\ \\ \end{bmatrix}^T \tag{7.208}$$

where $[U]$ is a 3×3, $[\Sigma]$ is a $3 \times N$ and $[V]$ is a $N \times N$ matrix. Assume that the singular values have been arranged such that $\sigma_1 \geq \sigma_2 \geq \sigma_3$. The non-dimensional condition number κ is then defined as

$$\kappa = \frac{\sigma_1}{\sigma_3} \tag{7.209}$$

As the gimbal angles approaches a singular CMG configuration, the index κ would grow very large since $\sigma_3 \to 0$. The theoretically best possible matrix conditioning would be with $\sigma_1 = \sigma_3$ where $\kappa = 1$. The goal of the VSCMG null motion would be to minimize the singularity index κ during a maneuver. Let $\kappa(t)$ be the singularity index at the current time and let $\kappa(t^+)$ be the index after a discrete gimbal angle adjustment has been made. Using a Taylor series expansion of κ in terms of the gimbal angles γ_i yields

$$\kappa(t^+) = \kappa(t) + \frac{\partial \kappa}{\partial \boldsymbol{\gamma}}^T \Delta\boldsymbol{\gamma} \tag{7.210}$$

Since ideally $\kappa(t^+) = 1$, using a minimum norm inverse, the desired gimbal angle correction is given by

$$\Delta\boldsymbol{\gamma} = \alpha \frac{(1 - \kappa(t))}{|\frac{\partial \kappa}{\partial \boldsymbol{\gamma}}|^2} \frac{\partial \kappa}{\partial \boldsymbol{\gamma}} \tag{7.211}$$

where the positive scalar α scales the gradient step. As will be shown with numerical examples, Eq. (7.211) works well when the gimbal configuration is to be rearranged while the spacecraft attitude is held stationary. However, if Eq. (7.211) is used to drive the gimbal angles away from singular configurations during a maneuver, then the VSCMG null motion corrections become too "soft" as a singular configuration is rapidly approached. The $|\partial\kappa/\partial\boldsymbol{\gamma}|^2$ term in the denominator drives $\Delta\boldsymbol{\gamma}$ to zero as $\partial\kappa/\partial\boldsymbol{\gamma}$ becomes very large in the neighborhood of a singularity. To counter this softening effect, the following stiffer gimbal correction algorithm is proposed.

$$\Delta\boldsymbol{\gamma} = \alpha(1 - \kappa(t)) \frac{\partial \kappa}{\partial \boldsymbol{\gamma}} \tag{7.212}$$

Numerical studies show that the VSCMG null motion driven by this $\Delta\boldsymbol{\gamma}$ during a maneuver is more successful in keeping the gimbal angles away from singular configurations.

If $[D_1]$ is reasonably well conditioned, it is not desirable to have the VSCMG be active at this point and drive the gimbal angles to an even better conditioned configuration. Doing so would only unnecessarily waste fuel and energy. To stop the VSCMG null motion at some pre-determined singularity index κ_{db}, a deadband is introduced. Whenever $\kappa \leq \kappa_{db} > 1$, then we set $\alpha = 0$.

Using Eq. (7.209), the partial derivatives of κ with respect to the gimbal angles are found to be

$$\frac{\partial \kappa}{\partial \gamma_i} = \frac{1}{\sigma_3} \frac{\partial \sigma_1}{\partial \gamma_i} - \frac{\sigma_1}{\sigma_3} \frac{\partial \sigma_3}{\partial \gamma_i} \tag{7.213}$$

The partial derivatives of the singular values with respect to the gimbal angles are given by[8]

$$\frac{\partial \sigma_j}{\partial \gamma_i} = \boldsymbol{u}_j^T \frac{\partial [D_1]}{\partial \gamma_i} \boldsymbol{v}_j \tag{7.214}$$

The result in Eq. (7.214) may have to be modified if $\sigma_1 = \sigma_3$. However, this event will never be encountered if $\kappa_{db} > 1$ is adopted. Using Eq. (7.183b), the partial derivative of $[D_1]$ with respect to γ_i is readily found to be

$$\frac{\partial [D_1]}{\partial \gamma_i} = [\mathbf{0} \cdots \mathbf{0} \ \boldsymbol{\chi}_i \ \mathbf{0} \cdots \mathbf{0}] \tag{7.215}$$

where the i-th column vector $\boldsymbol{\chi}_i$ is defined to be

$$\boldsymbol{\chi}_i = \frac{\partial \hat{\boldsymbol{g}}_{t_i}}{\partial \gamma_i} J_{s_i} \left(\Omega_i + \omega_{s_i}\right) + \hat{\boldsymbol{g}}_{t_i} J_{s_i} \left(\Omega_i + \frac{\partial \hat{\boldsymbol{g}}_{s_i}}{\partial \gamma_i}^T \boldsymbol{\omega} \right) \tag{7.216}$$

Since the partial derivatives of the gimbal frame axes are given by

$$\frac{\partial \hat{\boldsymbol{g}}_{s_i}}{\partial \gamma_i} = \hat{\boldsymbol{g}}_{t_i} \tag{7.217a}$$

$$\frac{\partial \hat{\boldsymbol{g}}_{t_i}}{\partial \gamma_i} = -\hat{\boldsymbol{g}}_{s_i} \tag{7.217b}$$

the vector $\boldsymbol{\chi}_i$ is expressed compactly as

$$\boldsymbol{\chi}_i = -\hat{\boldsymbol{g}}_{s_i} J_{s_i} \left(\Omega_i + \omega_{s_i}\right) + \hat{\boldsymbol{g}}_{t_i} J_{s_i} (\Omega_i + \omega_{t_i}) \tag{7.218}$$

Substituting Eqs. (7.215) and (7.218) into Eq. (7.214) and carrying out the vector algebra, the partial derivatives of the singular values with respect to the gimbal angles are given by

$$\frac{\partial \sigma_j}{\partial \gamma_i} = \left(\boldsymbol{u}_j^T \boldsymbol{\chi}_i\right) [V_{ij}] \tag{7.219}$$

Note that these singular value sensitivities can be computed very quickly given the vectors $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ obtained from a numerical SVD of the local matrix $[D_1]$. Therefore the $\Delta\boldsymbol{\gamma}$ vector can be easily computed and fed to the VSCMG null motion in Eq. (7.207). As the following numerical simulation shows, the end result is convenient method to drive the gimbal angles away from singular neighborhood and maintain a well-conditioned control influence matrix $[D_1]$.

Example 7.16: The following numerical simulation illustrates both the use of the VSCMG steering law in Eq. (7.192) to produce the required torque even in singular CMG configurations, and the use of the VSCMG null motion steering law to drive the gimbal angles continuously away from singular CMG configurations. The simpler VSCMG null motion in Eq. (7.206) is used here which weighs the RW and CMG modes equally. A rigid spacecraft is reoriented from large initial displacements to coincide with the target attitude through the use of four VSCMGs. The VSCMGs are arranged in the standard CMG pyramid configuration shown in Figure 7.12. The spacecraft and VSCMG properties are given in Table 7.6.

Two simulations are performed. One simulation uses only the VSCMG steering law in Eq. (7.192). The second simulation superimposes onto this steering

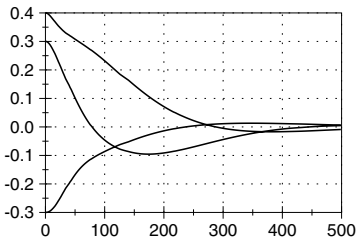
Table 7.6: Spacecraft and VSCMG Properties

Parameter	Value	Units
I_{s_1}	15053	kg-m^2/sec
I_{s_2}	6510	kg-m^2/sec
I_{s_3}	11122	kg-m^2/sec
N	4	
θ	54.75	degrees
J_s	0.70	kg-m^2
J_t	0.35	kg-m^2
J_g	0.35	kg-m^2
$\boldsymbol{\sigma}(t_0)$	$[0.4\ 0.3\ -0.3]$	
$\boldsymbol{\omega}(t_0)$	[0.0 0.0 0.0]	rad/sec
$\boldsymbol{\gamma}(t_0)$	$[45\ -45\ 45\ -45]$	deg
$\dot{\boldsymbol{\gamma}}(t_0)$	[0 0 0 0]	rad
$\Omega_i(t_0)$	628	rad/sec
Ω_f	628	rad/sec
$[P]$	[725 477 623]	kg-m^2/sec
K	35	kg-m^2/sec^2
$K_{\dot{\gamma}}$	1.0	sec^{-1}
$W_{s_i}^0$	40	
W_{g_i}	1.0	
μ	100	
κ_{db}	3	

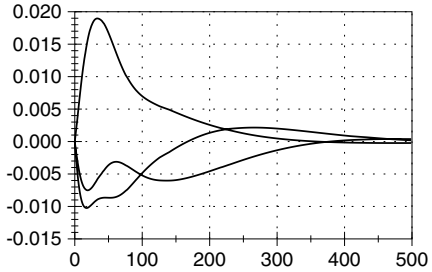
law the VSCMG null motion given Eq. (7.206) with the stiff gradient multiplier in Eq. (7.212) to continuously reconfigure the gimbal angles away from singular configurations. The vector $\boldsymbol{\Omega}_f$ is chosen to be the same as the initial $\boldsymbol{\Omega}(t_0)$ vector, which results in the null motion trying to keep the RW spin speeds as close to their original values as possible. The values of the diagonal angular velocity feedback gain matrix $[P]$ were chosen such that each mode of the linearized closed loop dynamics is critically damped.[12, 27] The null motion weights $\hat{W}_{s_i}$ and $\hat{W}_{g_i}$ are set equal in this simulation. Setting $\hat{W}_{s_i}$ equal to zero would have yielded a pure CMG null motion. This is typically the preferred setting. By having $\hat{W}_{s_i} = \hat{W}_{g_i}$ in this simulation, the null motion utilizes the RW mode very little. However, setting $\hat{W}_{s_i}$ equal to zero would restrict the types of null motion (therefore what types of gimbal angle reconfigurations) are possible.[12, 27]

The resulting numerical simulations are illustrated in Figure 7.14. Results obtained from the simulation that only utilized the steering law in Eq. (7.192) are indicated by a dashed line, results obtained from the simulation with VSCMG null motion added are indicated with a solid line. Figures 7.14(i) and 7.14(ii) are valid for both simulations and show that the closed loop dynamics is indeed asymptotically stable for both simulations.

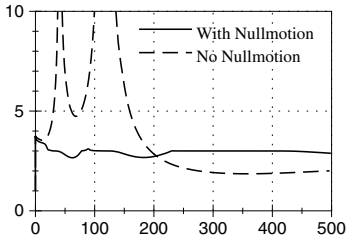
Figures 7.14(iii) and 7.14(iv) shows the singularity indices κ and δ for both simulations. Without the singularity-avoiding null motion added, the gimbal



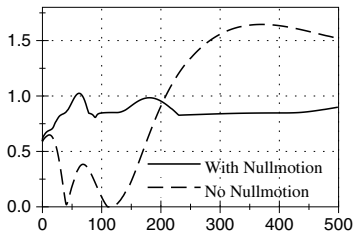
(i) MRP Attitude Vector $\boldsymbol{\sigma}$



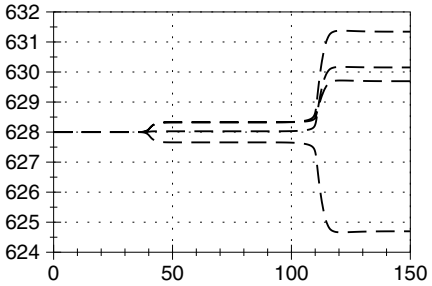
(ii) Angular Velocity Vector $\boldsymbol{\omega}$ (rad/s)



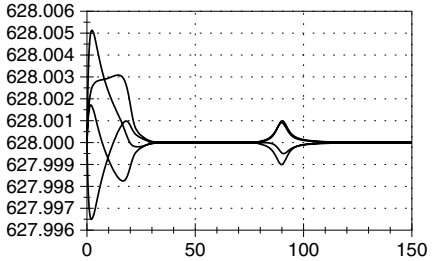
(iii) Singularity Index κ



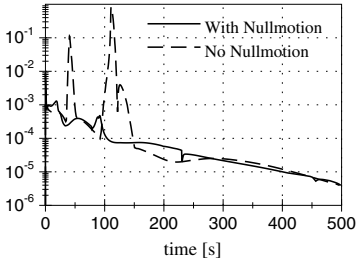
(iv) Singularity Index δ



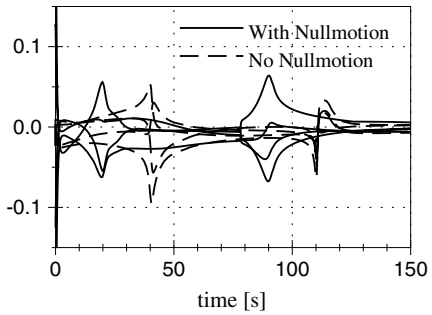
(v) $\boldsymbol{\Omega}$ Without Null Motion (rad/s)



(vi) $\boldsymbol{\Omega}$ With Null Motion (rad/s)



(vii) $\boldsymbol{u}_s$ Vector Magnitude (N-m)



(viii) Gimbal Rates Vector $\dot{\boldsymbol{\gamma}}$ (rad/s)

Figure 7.14: Comparison of Maneuvers With and Without VSCMG Null Motion

angles approach a singularity twice. During the second approach the non-dimensional determinant δ actually reaches zero and remains zero for a finite duration. Therefore it would be impossible to precisely perform this maneuver with the conventional CMG steering law. Some modifications would have to be used to produce and $approximate$ required torque in the neighborhood of this singular configuration. However, the VSCMG steering law is easily able to handle this singularity by temporarily using its RW modes. During both periods where $\delta \to 0$, the condition number κ grows very large as seen in Figure 7.14(iii). If the same maneuver is performed with the singularity avoiding VSCMG null motion added, the condition number κ is reduced from the outset and remains relatively low throughout the maneuver. Note that this index could have been reduced even more, but it remains essentially around the given condition number deadband value of 3. The trade off of lowering this deadband value is that the VSCMG null motion ends up reconfiguring the gimbals more often (i.e. using more energy).

One drawback of the VSCMG steering law as proposed in Ref. 27 is that for it to be able to drive through singular configurations, a relatively change in $\mathbf{\Omega}$ (i.e. large RW motor torque) is required. For this maneuver the associated $\mathbf{\Omega}$ changes are illustrated in Figure 7.14(v). Note that the time scale in this and some other Figures is changed to better illustrate the "interesting" regions. Using the VSCMG null motion to reconfigure the CMG cluster to preferred gimbal angles as was done in Example 7.15, it was found that the associated RW $\mathbf{\Omega}$ changes were rather small. The same is observed here where the null motion is performed during the maneuver itself as seen in Figure 7.14(vi).

The equivalent RW motor torque vector magnitudes $|\boldsymbol{u}_s|$ are plotted in Figure 7.14(vii). Note that classical CMGs already have an active RW control motor that simply maintains a constant wheel speed. The additional effort required by the VSCMG null motion is visible as small "humps" of the solid like at the beginning of the maneuver and before 100 seconds. What is very encouraging is that the magnitude of these humps is very small and still easily feasible with the standard existing RW torque motors. Conversely, the standard VSCMG steering law requires periodically RW torques that are much larger and would require some reengineering of the RW control motors.

The associated gimbal rates for both simulations are shown in Figure 7.14(viii). While the added VSCMG null motion does require periodically higher gimbal rates to reconfigure the gimbals, the overall control effort for the CMG mode is about the same. Again, the biggest difference in control effort between adding the VSCMG null motion or not to the VSCMG steering law manifests itself in the required RW control effort.

Problems

7.1 Given Euler's rotational equation of motion in Eq. (4.32) and (4.33).

a) Linearize them about $\boldsymbol{\omega}(t) = 0$.

b) Linearize them about $\boldsymbol{\omega}(t) = \boldsymbol{\omega}_r(t)$.

7.2 Examine the following functions $V(\boldsymbol{x})$ where $\boldsymbol{x} = (x_1, x_2)^T$. Find if they are positive (negative) definite or semi-definite functions. Also state the neighborhood $B_\delta(\boldsymbol{x}_r)$ about which this holds.

a) $V(\boldsymbol{x}) = \frac{1}{2}\left(x_1^2 + x_2^2\right)$

b) $V(\boldsymbol{x}) = \frac{1}{2}\left(x_1^2 - x_2^2\right)$

c) $V(\boldsymbol{x}) = \log\left(1 + x_1^2 + x_2^2\right)$

d) $V(\boldsymbol{x}) = \frac{1}{2}\left(x_1^2 + 4x_2^2\right)$

e) $V(\boldsymbol{x}) = \left(x_1^2 + 4x_2^2\right) e^{-\left(x_1^2 + 4x_2^2\right)}$

f) $V(\boldsymbol{x}) = x_1^2 - 2x_1 + x_2^2 - 4x_2 + 5$

7.3 Verify Eqs. (7.68) starting from the Lyapunov function definition in Eq. (7.67).

7.4 ♣ Several more HW problems and projects will be added to this chapter.

Bibliography

[1] Rugh, W. J., *Linear System Theory*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1993.

[2] Van de Vegte, J., *Feedback Control Systems*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 2nd ed., 1990.

[3] Slotine, J. E. and Li, W., *Applied Nonlinear Control*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1991.

[4] Wiggins, S., *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, Texts in Applied Mathematics 2, Springer Verlag, New York, 1990.

[5] Mohler, R. R., *Dynamics and Control*, Vol. 1 of *Nonlinear Systems*, Prentice Hall, Englewood Cliffs, New Jersey, 1991.

[6] Mukherjee, R. and Chen, D., "Asymptotic Stability Theorem for Autonomous Systems," *Journal of Guidance, Control, and Dynamics*, Vol. 16, Sept.–Oct. 1993, pp. 961–963.

[7] Mukherjee, R. and Junkins, J. L., "Invariant Set Analysis of the Hub-Appendage Problem," *Journal of Guidance, Control, and Dynamics*, Vol. 16, Nov.–Dec. 1993, pp. 1191–1193.

[8] Junkins, J. L. and Kim, Y., *Introduction to Dynamics and Control of Flexible Structures*, AIAA Education Series, Washington D.C., 1993.

[9] Robinett, R. D., Parker, G. G., Schaub, H., and Junkins, J. L., "Lyapunov Optimal Saturated Control for Nonlinear Systems," *35th Aerospace Sciences and Meeting*, Reno, Nevada, Jan. 1997, paper No. 97-0112.

[10] Junkins, J. L. and Bang, H., "Manuever and Vibration Control of Hybrid Coordinate Systems Using Lyapunov Stability Theory," *Journal of Guidance, Control and Dynamics*, Vol. 16, No. 4, Jul.–Aug. 1993, pp. 668–676.

[11] Schaub, H. and Junkins, J. L., "Feedback Control Law Using the Eigenfactor Quasi-Coordinate Velocity Vector," *Journal of the Chinese Society of Mechanical Engineers*, Vol. 19, No. 1, 1997, pp. 51–59.

[12] Schaub, H., *Novel Coordinates for Nonlinear Multibody Motion with Applications to Spacecraft Dynamics and Control*, Ph.D. thesis, Texas A&M University, College Station, TX, May 1998.

[13] Tsiotras, P., "A Passivity Approach to Attitude Stabilization Using Nonredundant Kinematic Parameterizations," *34th IEEE Conference on Decision and Control*, Princeton University, March 20-22 1996, pp. 1238–1243.

[14] Junkins, J. L. and Turner, J. D., *Optimal Spacecraft Rotational Maneuvers*, Elsevier Science Publishers, Amsterdam, Netherlands, 1986.

[15] 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.

[16] 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.

[17] Schaub, H., Robinett, R. D., and Junkins, J. L., "Globally Stable Feedback Laws for Near-Minimum-Fuel and Near-Minimum-Time Pointing Maneuvers for a Landmark-Tracking Spacecraft," *Journal of the Astronautical Sciences*, Vol. 44, No. 4, 1996, pp. 443–466.

[18] Krishnan, S. and Vadali, S. R., "An Inverse-Free Technique for Attitude Control of Spacecraft Using CMGs," *Acta Astronautica*, Vol. 39, No. 6, 1997, pp. 431–438.

[19] Schaub, H., Robinett, R. D., and Junkins, J. L., "Adaptive External Torque Estimation by Means of Tracking a Lyapunov Function," *Journal of the Astronautical Sciences*, Vol. 44, No. 3, July–Sept. 1996.

[20] Junkins, J. L. and Bang, H., "Lyapunov Optimal Control Law for Flexible Space Structure Maneuver and Vibration Control," *Journal of the Astronautical Sciences*, Vol. 41, No. 1, Jan.–Mar. 1993, pp. 91–118.

[21] Lee, B. and Grantham, W. J., "Aeroassisted Orbital Maneuvering Using Lyapunov Optimal Feedback Control," *Journal of Guidance, Control and Dynamics*, Vol. 12, No. 2, 1989, pp. 237–242.

[22] Kalman, R. E. and Bertram, J. E., "Control System Analysis and Design Via the "Second Method" of Lyapunov: Continous Time Systems," *Transactions of ASME: Journal of Basic Engineering*, June 1960.

[23] Paielli, R. A. and Bach, R. E., "Attitude Control with Realization of Linear Error Dynamics," *Journal of Guidance, Control and Dynamics*, Vol. 16, No. 1, Jan.–Feb. 1993, pp. 182–189.

[24] Schaub, H., Akella, M., and Junkins, J. L., "Adaptive Control of Nonlinear Attitude Motions Realizing Linear Closed Loop Dynamics," *Journal of Guidance, Control and Dynamics*, Vol. 24, No. 1, January–February 2001, pp. 95–100.

[25] Schaub, H., Akella, M., and Junkins, J. L., "Adaptive Control of Nonlinear Attitude Motions Realizing Linear Closed Loop Dynamics," *9-th AAS/AIAA Astrodynamics Specialist Conference*, Breckenridge, CO, Feb. 1999, Paper No. 151.

[26] Oh, H. S. and Vadali, S. R., "Feedback Control and Steering Laws for Spacecraft Using Single Gimbal Control Moment Gyros," *Journal of the Astronautical Sciences*, Vol. 39, No. 2, 1991, pp. 183–203.

[27] Schaub, H., R.Vadali, S., and Junkins, J. L., "Feedback Control Law for Variable Speed Control Moment Gyroscopes," *8th AAS/AIAA Space Flight Mechanics Meeting*, Monterey, California, Feb. 9–11 1998, Paper No. AAS 98-140.

[28] Nakamura, Y. and Hanafusa, H., "Inverse Kinematic Solutions with Singularity Robustness for Robot Manipulator Control," *Journal of Dynamic Systems, Measurement, and Control*, Vol. 108, Sept. 1986, pp. 164–171.

[29] Junkins, J. L., *An Introduction to Optimal Estimation of Dynamical Systems*, Sijthoff & Noordhoff International Publishers, Alphen aan den Rijn, Netherlands, 1978.

[30] Vadali, S. R., Oh, H. S., and Walker, S. R., "Preferred Gimbal Angles for Single-Gimbal Control Moment Gyros," *Journal of Guidance, Control and Dynamics*, Vol. 13, No. 6, 1990, pp. 1090–1095.

Part II

CELESTIAL MECHANICS

Chapter Eight

Classical Two-Body Problem

Amongst the most important historical development in the history of the scientific method is Newton's analytical solution of the two-body problem. In Astronomia Nova (1609), Kepler published the first essentially correct solution for the motion of planets around the Sun, where the solution was obtained by solving the inverse problem of "given the observed (measured) right ascension and declination history of a planet, determine a mathematical model which captures the behavior with sufficient rigor that predictions as accurate as the measurements can be made." Kepler was seeking to modify the historical work of Copernicus who modeled the planet paths as circular orbits about the Sun. Kepler was intrigued by Tycho Brahe's observations of the orbit of Mars which appeared to deviate significantly from the Copernican circular model, and in the course of developing a model which fit the data for Mars and all the other more circular planetary orbits, he conjectured that the motion was actually an ellipse with the Sun at a focus. Kepler's elegant geometric analysis amounted to an insightful, sophisticated curve-fitting operation, but was found to be in such excellent agreement with the measurements of the day, that his laws of planetary motion were considered exact.

Newton, in his quest to develop calculus, differential equations, and his famous laws of motion, was fascinated by the beauty and precision of Kepler's laws and set about the task of discovering what force law must be existing between bodies in the solar system to be consistent with Newton's laws of motion and Kepler's experimentally verified laws of planetary motion. From this analysis came Newton's discovery of the law of universal gravitation, and the ensuing analytical solution of the two-body problem. Because Newton's analytical solution for Keplerian motion was an immediate and convincing demonstration of the validity of Newton's calculus, differential equations, and laws of motion, the acceptance of Newton's many allied advancements were immediate. The acceleration of the evolution of science and mathematics and the consequences

since are simply immeasurable.

This chapter is essentially a modern rendition of Newton's analytical solution of the two-body problem, however, we make extensive use of matrices and other modern mathematical constructions not available to Newton. Because the subject relies heavily on geometry of conic sections, we will begin with a brief summary of this subject.

8.1 Geometry of Conic Sections

Kepler discovered that the orbit of one body about another is an ellipse, which is a special case of the intersection curve between a cone and a plane. More generally, we will see how Newton proved all of the conic sections are feasible orbits. Depending on the slope of the plane relative to the cone's axis of symmetry, the resulting curves of these conic intersections are either of an elliptical, parabolic or hyperbolic nature as illustrated in Figure 8.1. If the relative slope of the plane is less than that of the cone symmetry axis, then the resulting intersections form a *closed*, elliptic curve. An *open* (infinite ellipse) parabolic curve is the result of the limiting case where both plane and cone symmetry axis slope are equal. The hyperbolic curve occurs for plane slopes larger than the relative cone symmetry axis slope. Mastering the basic geometry of conic sections is of fundamental importance to understanding orbital mechanics. This section provides a terse review of some of the more important aspect of the geometry of elliptic, parabolic and hyperbolic orbits.

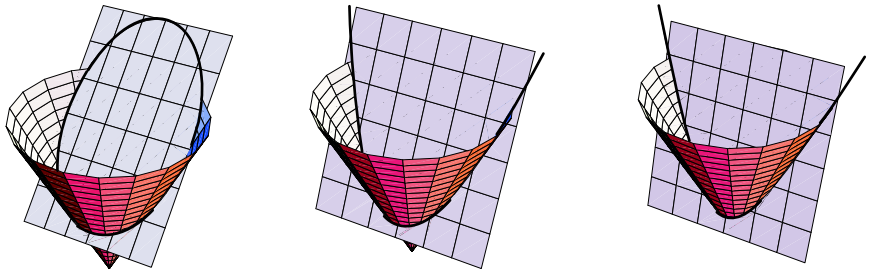


Figure 8.1: Illustration of Conic Intersections

A sample elliptical orbit is illustrated in Figure 8.2. The shape of an ellipse is defined through its *semi-major axis* a and *semi-minor axis* b where $a \geq b$. Let (X, Y) be the coordinates of a body performing an elliptical motion with the coordinate system origin chosen to be in the center of the ellipse. The standard rectangular coordinate description of an ellipse is given by

$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} = 1 \tag{8.1}$$

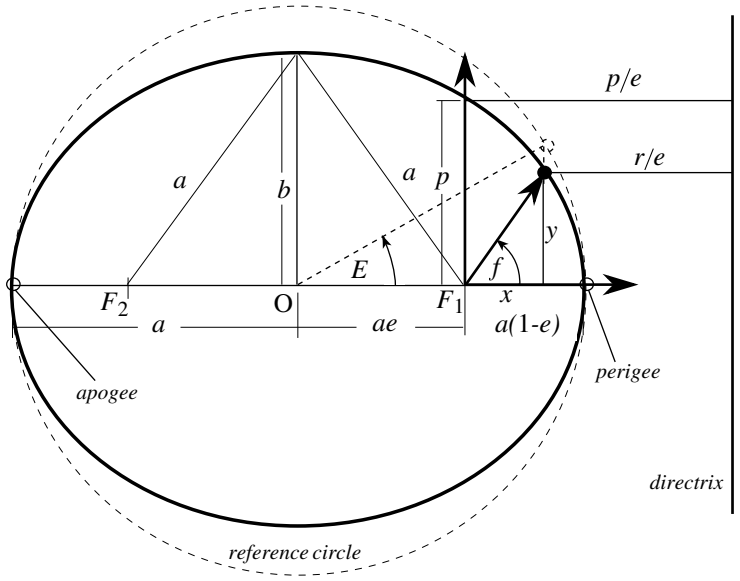


Figure 8.2: Geometry of an Elliptic Conic Section

Similarly, the rectangular coordinate description of a hyperbola is given by

$$\frac{X^2}{a^2} - \frac{Y^2}{b^2} = 1 \tag{8.2}$$

Instead of describing a location on the ellipse relative to the ellipse geometric center, this section develops a description which defines the conic section relative to a *focal* point.

Every ellipse has two focal points F_1 and F_2. For the special situation where the ellipse collapses to the circular case (i.e. $a = b$), the two focal points occupy the same point; this clearly corresponds to a planar section normal to the cone symmetry axis if Figure 8.1. A well known useful property of an ellipse is that the sum of the two radial distances from any point on the ellipse to each focal point is constant and equal to $2a$.

An important parameter that describes the shape of conic intersections is the non-dimensional constant e called the *eccentricity*. It indicates whether the conic intersection is elliptic, parabolic or hyperbolic. For ellipses the eccentricities range between $0 \leq e \leq 1$. Parabolas always have $e = 1$ and hyperbolas always have eccentricities great than 1.

With reference to Figure 8.2, we introduce the "directrix definition of a conic section." Begin with constructing two perpendicular lines. On the first line locate a point F. Designate the first line as *major axis* and the second perpendicular line as the *directrix*. The conic section is defined as the curve whose radial distance r from F to a typical point P on the curve has a constant

Table 8.1: Naming Convention of Periapses and Apoapses for Orbits about Various Celestial Bodies

Celestial Body	Periapses	Apoapses
Sun	Perihelion	Apohelion
Mercury		
Venus		
Earth	Perigee	Apogee
Moon	Periselenium	Aposelenium
Mars		
Jupiter	Perijove	Apojove
Saturn		
Uranus		
Neptune		
Pluto		

ratio c to the perpendicular distance from P to the directrix. While the directrix itself typically doesn't appear in the description of orbital motion, it plays a key role below in deriving several important conic intersection properties.

First, we derive another mathematical description of a conic section. While Eq. (8.1) and (8.2) each were only valid for one type of orbit, we now develop a description valid for any type of orbit. Let the vector $\boldsymbol{r}$ point from the focus F to the current orbit position with r being its magnitude. The distance p is the perpendicular distance (to the major axis) between the focus and the orbit and is called the *semilatus rectum* or simply *the parameter*.[1] The angle f measures the heading of the position vector $\boldsymbol{r}$ relative to the semi-major axis and is called the *true anomaly*. The cartesian $\boldsymbol{r}$ vector components (x, y) are expressed as

$$x = r\cos f \tag{8.3}$$

$$y = r\sin f \tag{8.4}$$

Using the property of the directrix, we can see from the geometry of Fig. 8.2 that the following statement must be true.

$$\frac{p}{e} = x + \frac{r}{e} \tag{8.5}$$

Using Eq. (8.3), this statement is rewritten to express the radial distance r in terms of the true anomaly f.

$$r = \frac{p}{1 + e\cos f} \tag{8.6}$$

Note that Eq. (8.3) not only holds for the elliptical case, but also describes parabolic and hyperbolic trajectories. Thus it forms a *universal* description of conic intersections.

The closest point on the ellipse to the focus is called *periapses* or *perifocus*, while the furthest point is called *apoapses* or *apofocus*. When orbiting about

certain celestial bodies, these terms are refined to reflect the fact that the orbit is about a particular body. For example, for an orbit about Earth the closed and farthest points are called *perigee* and *apogee*. Other naming conventions are found in Table 8.1. Note the closest point to the focus F occurs at $f = 0$, this gives the perifocus radius

$$r_p = \frac{p}{1+e} \tag{8.7}$$

For the case of closed orbits, investigate the point at $f = \pi$, this give the apofocus radius

$$r_a = \frac{p}{1-e} \tag{8.8}$$

It is clear that

$$2a = r_a + r_p = \frac{p}{1-e} + \frac{p}{1+e}$$

from which

$$2a = p\left(\frac{1+e+1-e}{1-e^2}\right)$$

or

$$p = a(1-e^2) \tag{8.9}$$

So we now see that

$$r_p = a(1-e) \tag{8.10}$$

$$r_a = a(1+e) \tag{8.11}$$

Further, since (from Fig. 8.2) $a = OF + r_p$, then $OF = a - r_p = ae$ is the distance from the ellipse center O to the focus F.

The semi-minor axis b can be expressed in terms of a and e as

$$b = a\sqrt{1-e^2} \tag{8.12}$$

Therefore, if $e \to 0$, then the orbit becomes circular and $b = a$. Having $e \to 1$ could indicate either one of two situations. The first case is that the conic section is becoming parabolic, in which case both a and b would grow infinitely large. However, it is also possible for e to approach 1 without the orbit becoming an open-pathed parabola. As Eq. (8.12) indicates, if the semi-minor axis b shrinks to zero for a fixed and finite semi-major axis a, then the corresponding eccentricity e would have to approach 1. At the limit where $e = 1$, the elliptic motion collapses down to a cyclic motion on a finite line segment. This case is referred to as the rectilinear motion case. Examples of orbits where $e \to 1$ without the flight path becoming parabolic are comets that typically are on a "skinny", near-parabolic orbit about the sun. As will be shown later on in

this chapter, the deciding factor whether an object is on an elliptic, parabolic or hyperbolic orbit is the object's energy.

For the rectilinear motion case where $e = 1$ and a is finite, the corresponding r_p is obviously zero. From this it is evident that a completely rectilinear orbit is not possible in reality. The celestial body about which the object is orbiting in this manner would have to have an infinitesimally small diameter. However, true rectilinear fractional-orbits are possible. To illustrate, throw a rock straight up in the air on a non-rotating earth. The resulting flight path before the objects impacts with the ground (or yourself if you do not step aside) is a true rectilinear ellipse. Using Eq. (8.10), the semi-minor axis b in Eq. (8.12) can be written in terms of r_p and e.

$$b = r_p \sqrt{\frac{1+e}{1-e}} \tag{8.13}$$

While the true anomaly f has a convenient direct geometric interpretation, it is not always mathematically convenient to express the current location through this angle. Instead, the *eccentric anomaly* E is often used. Imagine the ellipse being "stretched" along the semi-minor axis into the shape of a perfect reference circle. The angular position of the new orbit location relative to the ellipse center is the true anomaly E as shown in Figure 8.2. The following developments express various elliptic elements in terms on this eccentric anomaly E. While Eqs.(8.3), (8.4) and (8.6) are universally valid, the following expressions using the eccentric anomaly E are only valid for the elliptic special case. To write the radial distance r in terms of E instead of using the universally valid f, we use the directrix property to state that

$$ae + \frac{p}{e} = a\cos E + \frac{r}{e} \tag{8.14}$$

Substituting Eq. (8.9), the radial distance r is expressed as

$$r = a(1 - e\cos E) \tag{8.15}$$

Studying Figure 8.2, the semi-major axis component x is written as

$$x = a(\cos E - e) \tag{8.16}$$

The semi-minor axis component y is found by making use of $y = \sqrt{r^2 - x^2}$ and performing some trigonometric simplifications.

$$y = a\sqrt{1-e^2}\sin E \tag{8.17}$$

Finding a direct relationship between the true anomaly f and the eccentric anomaly E is less straight forward than the previous developments. In particular, it involves using half angle trigonometric identities whose use is initially not very intuitive. Using Eqs. (8.16) and (8.17), the sin and cos of f are written as

$$\sin f = 2\cos\frac{f}{2}\sin\frac{f}{2} = \frac{y}{r} = \frac{\sqrt{1-e^2}\sin E}{1 - e\cos E} \tag{8.18}$$

$$\cos f = \cos^2\frac{f}{2} - \sin^2\frac{f}{2} = \frac{x}{r} = \frac{\cos E - e}{1 - e\cos E} \tag{8.19}$$

where the first transformation was performed using standard half angle trigonometric identities. Making use of $\sin^2 \frac{f}{2} = 1 - \cos^2 \frac{f}{2}$, Eq. (8.19) is rewritten as

$$2\cos^2 \frac{f}{2} = \frac{(1-e)(1+\cos E)}{1 - e\cos E} \tag{8.20}$$

After dividing Eq. (8.18) by Eq. (8.20) and performing some simplifications, we find that

$$\tan \frac{f}{2} = \sqrt{\frac{1+e}{1-e}} \frac{\sin E}{1+\cos E} \tag{8.21}$$

The right hand side of Eq. (8.21) can be further simplified by again making use of the previous half angle trigonometric identities. The final result is a remarkably simple transformation between the true anomaly f and the eccentric anomaly E.

$$\tan \frac{f}{2} = \sqrt{\frac{1+e}{1-e}} \tan \frac{E}{2} \tag{8.22}$$

Note that quadrants are not an issue in the above anomaly transformation. With this mapping we will be able to exploit the simpler mathematical expressions in terms of E and then translate this angle into the geometrically more meaningful angle f for visualization purposes.

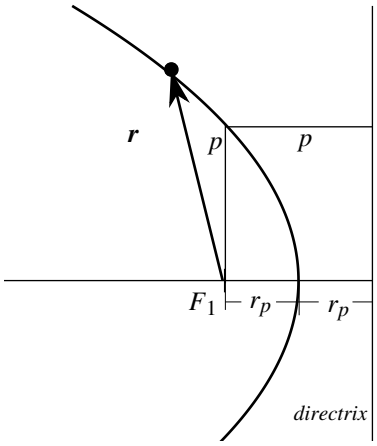


Figure 8.3: Geometry of a Parabolic Orbit

As the eccentricity e approaches 1 and the semi-major axis a grows to infinity, the orbit no longer remains a closed path. At the critical transition, the conic intersection shape is that of a parabola with the second focus point F_2 having moved off to infinity. Figure 8.3 illustrates such a parabolic orbit. Since $e = 1$, note that the distance from any point on the parabola to the focus F_1 is the

same as to the directrix. Since the distance r_p is finite, Eq. (8.13) indicates that the semi-minor axis b is infinitely large. Eq. (8.7) shows that the semilatus rectum for a parabola is simply

$$p = 2r_p \tag{8.23}$$

However, actual orbits are rarely parabolic since the eccentricity must be *precisely* equal to 1. Theoretical analysis therefore typically focuses on the more common cases of having either elliptical or hyperbolic orbits.

Figure 8.4 illustrates the geometry of a hyperbola. While a parabola has moved the second focus F_2 off to infinity, with hyperbolas this focus reappears on the other side of F_1. A common practice in celestial mechanics is to denote

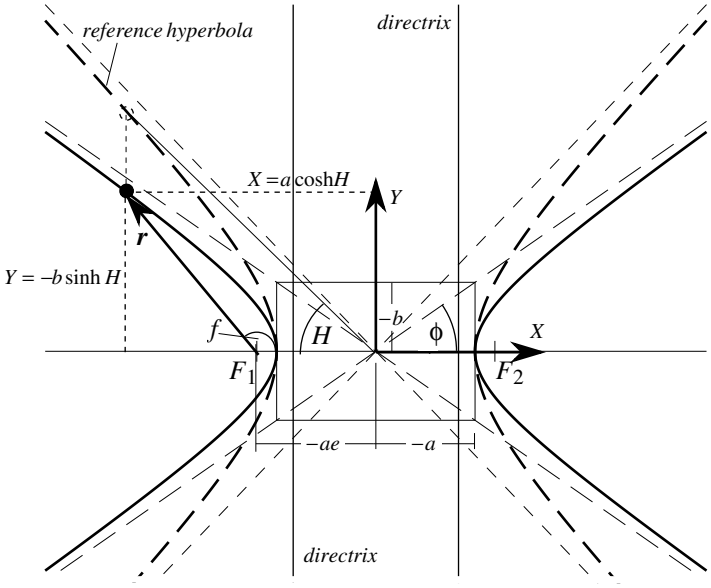


Figure 8.4: Geometry of a Hyperbolic Orbit

the semi-axes a and b as being negative quantities for a hyperbola. This results in many expressions for hyperbolic parameters being algebraically equivalent to their elliptic cousins. Also, this convention will allow us to express the orbit energy equation in one algebraic form that pertains to all three possible conic section cases. The distance between foci is $-2ae > 0$, similar as with an ellipse. However, the semi-axes a and b now have different geometric meanings. The distance between the two hyperbola periapses is $-2a$. Apoapses points don't make sense in this setting since the curve isn't closed. A hyperbolic curve will asymptotically approach a straight line motion as the true anomaly f grows sufficiently large. Unlike with ellipses and parabolas, the hyperbola is essentially only curved in the proximity of its focus. Let's assume a cartesian coordinate system is aligned with the semi-axes and has its center between the two foci.

The slope ϕ that the hyperbola will asymptotically approach is given by

$$\phi = \cos^{-1}\frac{1}{e} \tag{8.24}$$

The box of dimension $(-2b) \times (-2a)$ between the two perigee points has a diagonal of length $-2ae$. Therefore, the semi-minor axis b for a hyperbola can be expressed as

$$b = a\sqrt{e^2 - 1} \tag{8.25}$$

Substituting Eq. (8.24), the parameters a and b can be related through the slope angle ϕ as

$$b^2 = a^2 \tan^2 \phi \tag{8.26}$$

From the geometry of Figure 8.4, it is evident that

$$r_p = (-a)(e - 1) = a(1 - e) \tag{8.27}$$

Note that this expression is algebraically equivalent to the elliptic r_p expression in Eq. (8.10). Using Eqs. (8.6) and (8.27) and setting $f = 0$, the semilatus rectum p for a hyperbola is given by

$$p = (-a)(e^2 - 1) = a(1 - e^2) \tag{8.28}$$

or alternatively through

$$p = r_p(1 + e) \tag{8.29}$$

As is the case with elliptic orbits, it is convenient to express the location within the orbit through another anomaly angle. For hyperbolas the *hyperbolic anomaly* H is used. In this case the reference hyperbola is created using an eccentricity of $e = \sqrt{2}$ which corresponds to having an asymptotic slope of 45 degrees. Following similar steps as was done with the elliptic case, the parameters r, x and y are expressed in terms of H as

$$r = (-a)(e\cosh H - 1) = a(1 - e\cosh H) \tag{8.30}$$

$$x = (-a)(e - \cosh H) = a(\cosh H - e) \tag{8.31}$$

$$y = (-a)\sqrt{e^2 - 1}\sinh H \tag{8.32}$$

Finding the direct relationship between the true anomaly f and the hyperbolic anomaly H again involves half-angle identities. Besides the previously shown standard half-angle trigonometric identities, the following hyperbolic identities are used.

$$1 = \cosh^2 x - \sinh^2 x \tag{8.33}$$

$$\sinh x = 2\cosh\frac{x}{2}\sinh\frac{x}{2} \tag{8.34}$$

$$\cosh x = \cosh^2\frac{x}{2} + \sinh^2\frac{x}{2} \tag{8.35}$$

After following the same steps as were done in developing the relationship between f and E, the hyperbolic anomaly H is related to the true anomaly f through

$$\tan\frac{f}{2} = \sqrt{\frac{e+1}{e-1}}\tanh\frac{H}{2} \tag{8.36}$$

8.2 Relative Two-Body Equations of Motion

In celestial mechanics, bodies are often treated as particles with their rigid body motion neglected. The reason for this naturally being the typical spherical shape of massive heavenly bodies and the large relative distances involved. Also, we'll see mass elements of finite bodies can be considered particles and integration over the mass distribution gives various derived results for finite bodies. Assume two particles of mass m_1 and m_2 are moving generally in space. The only forces acting on them is the mutual gravitational attraction and some disturbance forces $\boldsymbol{f}_{d_1}$ and $\boldsymbol{f}_{d_2}$ as illustrated in Figure 8.5. The magnitude of the gravitational attraction is given by Newton's Law of Universal Gravitation in Eq. (2.4). The position vectors $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ are measured relative to an inertial reference frame $\mathcal{N}$.

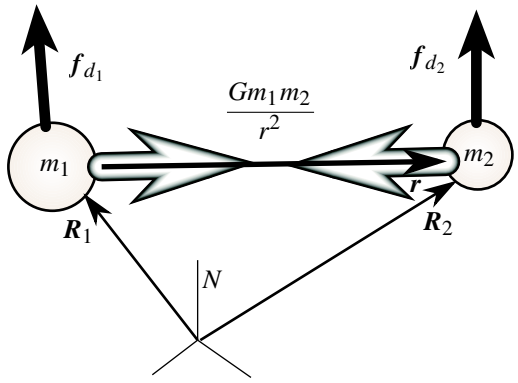


Figure 8.5: Gravity and Disturbance Forces Between Two Bodies

The disturbance forces $\boldsymbol{f}_{d_i}$ could be present for various reasons. In a Low Earth Orbit (LEO) the aerodynamic drag of the rarified atmosphere could affect the motion. Considering the Earth and its moon to be a two particle system, then both would experience another gravitational attraction with the sun which could be expressed as a disturbance force on the two-body Earth-Moon system. In the present discussion the inertial motion of each mass is of lesser importance. Instead, we would like to focus on the *relative* motion between the two bodies. The position of mass m_2 relative to mass m_1 is given through the vector

$$\boldsymbol{r} = \boldsymbol{R}_2 - \boldsymbol{R}_1 = r\hat{\boldsymbol{\imath}}_r \tag{8.37}$$

where $r = |\boldsymbol{r}|$ and $\hat{\boldsymbol{\imath}}_r = \boldsymbol{r}/r$. Using Newton's equations of motion in Eq. (2.2), the inertial equations of motion for each body are written as

$$m_1 \ddot{\boldsymbol{R}}_1 = \frac{Gm_1 m_2}{r^3}\boldsymbol{r} + \boldsymbol{f}_{d_1} \tag{8.38}$$

$$m_2 \ddot{\boldsymbol{R}}_2 = -\frac{Gm_1 m_2}{r^3}\boldsymbol{r} + \boldsymbol{f}_{d_2} \tag{8.39}$$

where G is the universal gravity constant. The *gravitational coefficient* μ is defined as

$$\mu = G(m_1 + m_2) \tag{8.40}$$

Note that for many systems $m_1 \gg m_2$ and μ can therefore be approximated as

$$\mu \approx Gm_1 \tag{8.41}$$

An example of this situation would be a satellite in Earth's orbit. The mass m_2 of the satellite would be negligible compared to the massive Earth. The practical reason for choosing to work with μ instead of G is that μ is more accurately understood for various systems than is the universal gravitational constant G. The product Gm_1 can be extracted from LEO satellite tracking data with a relatively high degree of accuracy. However, measuring G directly is much more challenging. More on this later. Taking the difference between Eq. (8.39) and (8.38), the equations of motion of m_2 relative to m_1 are found to be[2]

$$\ddot{\boldsymbol{r}} = -\frac{\mu}{r^3}\boldsymbol{r} + \boldsymbol{a}_d \tag{8.42}$$

where the disturbance acceleration vector $\boldsymbol{a}_d$ is defined as

$$\boldsymbol{a}_d = \frac{1}{m_2}\boldsymbol{f}_{d_2} - \frac{1}{m_1}\boldsymbol{f}_{d_1} \tag{8.43}$$

This vector differential equation in Eq. (8.42) is easily the most important result in celestial mechanics. It forms the basis for various developments. Note that the two disturbance accelerations in Eq. (8.43) are often a near cancellation. Again we consider the Earth-Moon system with the sun's gravitational attraction modeled as the external influence. Labeling the sun's mass as m_3, we express the disturbance acceleration $\boldsymbol{a}_d$ as

$$\boldsymbol{a}_d = \frac{1}{m_2}\frac{Gm_2 m_3}{|\boldsymbol{r}_{23}|^3}\boldsymbol{r}_{23} - \frac{1}{m_1}\frac{Gm_1 m_3}{|\boldsymbol{r}_{13}|^3}\boldsymbol{r}_{13} \approx 0 \tag{8.44}$$

since $\boldsymbol{r}_{23} \approx \boldsymbol{r}_{13}$. Thus, even though the sun's gravitational force itself is very large, it effect on the *relative* two-body motion is often negligible. Therefore the relative disturbance acceleration vector $\boldsymbol{a}_d$ is typically considered to be small or actually set equal to zero to obtain a good approximate solution. In this case, the relative equations of motion are written as

$$\ddot{\boldsymbol{r}} = -\frac{\mu}{r^3}\boldsymbol{r} \tag{8.45}$$

By defining the relative gravitational potential energy function V as

$$V(\boldsymbol{r}) = -\frac{\mu}{r} \tag{8.46}$$

the relative equations of motion can also be written using the $\nabla_{\boldsymbol{r}}$ operator as

$$\ddot{\boldsymbol{r}} = -\nabla_{\boldsymbol{r}} V(\boldsymbol{r}) \tag{8.47}$$

If the relative position vector $\boldsymbol{r}$ is expressed through the cartesian vector components (x, y, z), then the relative equations of motion are given by

$$\ddot{x} = -\frac{\mu}{r^3} x \tag{8.48}$$

$$\ddot{y} = -\frac{\mu}{r^3} y \tag{8.49}$$

$$\ddot{z} = -\frac{\mu}{r^3} z \tag{8.50}$$

which form a set of three *coupled* , nonlinear, scalar differential equations. The differential equations only decouple for the special cases of having a circular orbit where the radius r remains constant, or having a straight line orbit ($y = z = 0 \Rightarrow \ddot{x} = -\frac{\mu}{x^2}$).

8.3 Fundamental Integrals

Even though the relative equations of motion in Eq. (8.42) are nonlinear differential equation, it is remarkable that an exact analytical solution to them exists. This section will show several manipulations which each lead to perfect differentials. Integrating these differentials then leads to the *fundamental integrals* of an orbit. In the absence of disturbances, these parameters remain constant and provide an important, geometrically elegant way to describe an orbit. Further, from these fundamental integrals we are also able to verify analytically Kepler's three laws of planetary motion.

8.3.1 Conservation of Angular Momentum

In this section we study a variation of the standard angular momentum vector. The massless angular momentum vector $\boldsymbol{h}$ is defined as

$$\boldsymbol{h} = \boldsymbol{r} \times \dot{\boldsymbol{r}} = h\hat{\boldsymbol{\imath}}_h \tag{8.51}$$

A rotating coordinate system $\mathcal{M}$ is placed on the mass m_1 with the unit direction vectors $\{\hat{\boldsymbol{\imath}}_r, \hat{\boldsymbol{\imath}}_\theta, \hat{\boldsymbol{\imath}}_h\}$ as shown in Figure 8.6. To see what type of relative motion is possible with Eq. (8.42), we differentiate the massless angular momentum vector $\boldsymbol{h}$. Using the chain rule and the relative equations of motion in Eq. (8.42), the vector $\dot{\boldsymbol{h}}$ is written as

$$\dot{\boldsymbol{h}} = \dot{\boldsymbol{r}} \times \dot{\boldsymbol{r}} + \boldsymbol{r} \times \ddot{\boldsymbol{r}} = \boldsymbol{r} \times \left(-\frac{\mu}{r^3}\boldsymbol{r} + \boldsymbol{a}_d\right) = \boldsymbol{r} \times \boldsymbol{a}_d \tag{8.52}$$

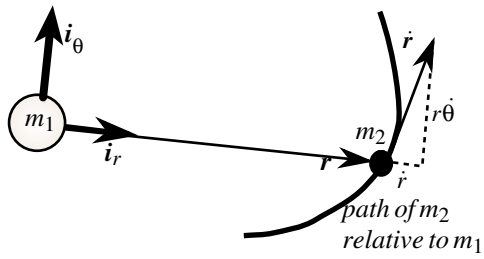


Figure 8.6: Planar Orbit Motion

Note that for the case where $\boldsymbol{a}_d \approx 0$, the angular momentum vector $\boldsymbol{h}$ remains inertially fixed since $\dot{\boldsymbol{h}} = 0$. This fundamental result is important since it states that all possible relative motions will lie in an inertially fixed plane perpendicular to $\hat{\boldsymbol{\imath}}_h$. Since all relative motion will occur in a plane, we introduce polar coordinates for the radial and transverse components of $\boldsymbol{r}$ and $\dot{\boldsymbol{r}}$ in this orbit plane. The velocity vector $\dot{\boldsymbol{r}}$ is then given by

$$\dot{\boldsymbol{r}} = \dot{r}\hat{\boldsymbol{\imath}}_r + r\dot{\theta}\hat{\boldsymbol{\imath}}_\theta \tag{8.53}$$

where $\dot{\theta}$ is the planar rotation rate of the position vector. Using Eqs. (8.37) and (8.53), the angular momentum vector $\boldsymbol{h}$ is written as

$$\boldsymbol{h} = \boldsymbol{r} \times \dot{\boldsymbol{r}} = (r\hat{\boldsymbol{\imath}}_r) \times (\dot{r}\hat{\boldsymbol{\imath}}_r + r\dot{\theta}\hat{\boldsymbol{\imath}}_\theta) = r^2\dot{\theta}\hat{\boldsymbol{\imath}}_h \tag{8.54}$$

Comparing this result with Eq. (8.51) we find

$$h = r^2\dot{\theta} \tag{8.55}$$

Since h is constant, we find *Kepler's second law of planetary motion* which states that the relative position vector $\boldsymbol{r}$ sweeps equal areas in equal equal times. Thus Kepler's second law is a geometric property of the conservation of angular momentum.

8.3.2 The Eccentricity Vector Integral

The following development will introduce the notion of eccentricity into the relative equations of motion. Also, it will be apparent that all relative motions between two bodies indeed either describe elliptic, parabolic or hyperbolic paths. Since the angular momentum $\boldsymbol{h}$ is perpendicular to both $\boldsymbol{r}$ and $\dot{\boldsymbol{r}}$, the vector $\dot{\boldsymbol{r}} \times \boldsymbol{h}$ lies in the orbit plane. Assuming $\boldsymbol{a}_d = 0$, then $\dot{\boldsymbol{h}} = 0$ and

$$\frac{d}{dt}(\dot{\boldsymbol{r}} \times \boldsymbol{h}) = \ddot{\boldsymbol{r}} \times \boldsymbol{h} \tag{8.56}$$

After substituting Eqs. (8.45) and (8.51), the derivative of $\dot{\boldsymbol{r}} \times \boldsymbol{h}$ is written as

$$\frac{d}{dt}(\dot{\boldsymbol{r}} \times \boldsymbol{h}) = -\frac{\mu}{r^3}\boldsymbol{r} \times (\boldsymbol{r} \times \dot{\boldsymbol{r}}) \tag{8.57}$$

Making use of the trigonometric identity

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) \equiv (\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{c} \tag{8.58}$$

and substituting the polar coordinate expressions in Eqs. (8.37) and (8.53), this derivative is rewritten as

$$\frac{d}{dt}(\dot{\boldsymbol{r}} \times \boldsymbol{h}) = \frac{\mu}{r^2}(r\dot{\boldsymbol{r}} - \dot{r}\boldsymbol{r}) \tag{8.59}$$

The key step in this development is being able to rewrite Eq. (8.59) in the form of a perfect differential:

$$\frac{d}{dt}(\dot{\boldsymbol{r}} \times \boldsymbol{h}) = \mu \frac{d}{dt}\left(\frac{\boldsymbol{r}}{r}\right) \tag{8.60}$$

This allows us to trivially integrate this expression by introducing the *constant* vector $\boldsymbol{c}$. The relative motion between two bodies must therefore satisfy the following constraint.

$$\boldsymbol{c} = \dot{\boldsymbol{r}} \times \boldsymbol{h} - \mu\left(\frac{\boldsymbol{r}}{r}\right) = constant \tag{8.61}$$

To gain further insight into the geometric meaning of the constant $\boldsymbol{c}$ vector, we perform the dot product between it and the position vector $\boldsymbol{r}$ to find

$$\boldsymbol{r} \cdot \boldsymbol{c} = \boldsymbol{r} \cdot \left(\dot{\boldsymbol{r}} \times \boldsymbol{h} - \mu\left(\frac{\boldsymbol{r}}{r}\right)\right) = h^2 - \mu r \tag{8.62}$$

where the trigonometric identity $\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}) \equiv \boldsymbol{c} \cdot (\boldsymbol{a} \times \boldsymbol{b})$ was used. Making use of the dot product definition

$$\boldsymbol{r} \cdot \boldsymbol{c} = r|\boldsymbol{c}| \cos(\angle\ \boldsymbol{r}, \boldsymbol{c}) \tag{8.63}$$

the scalar orbit radius r is expressed as

$$r = \frac{h^2/\mu}{1 + \frac{|\boldsymbol{c}|}{\mu}\cos(\angle\ \boldsymbol{r}, \boldsymbol{c})} \tag{8.64}$$

with the expression $(\angle\ \boldsymbol{r}, \boldsymbol{c})$ being the angle between the two vectors $\boldsymbol{r}$ and $\boldsymbol{c}$. Studying Eq. (8.6), it is evident that Eq. (8.64) geometrically describes a conic intersection. Thus we have proven *Kepler's first law of planetary motion* which states that all relative motions between two bodies are either elliptic, parabolic or hyperbolic in nature. Further, we can also express the semilatus rectum p in terms of the angular momentum magnitude h as

$$h^2 = \mu p \tag{8.65}$$

The angle $(\angle\ \boldsymbol{r}, \boldsymbol{c})$ is now exposed as being the true anomaly f. This implies that the constant vector $\boldsymbol{c}$ is aligned with the semi-major axis and points toward

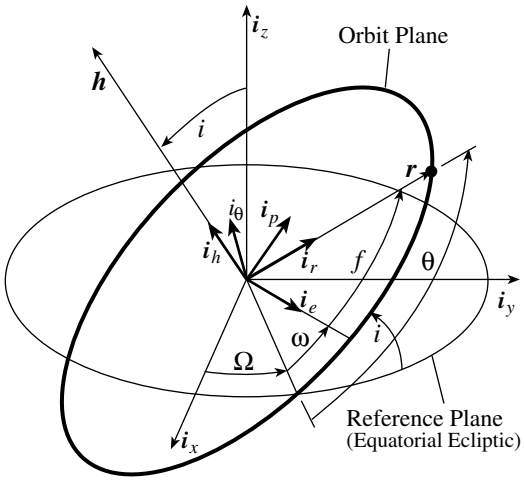


Figure 8.7: (3-1-3) Euler Angle Description of the Orbit Plane

periapses, since f is measured from this axis. Comparing Eqs. (8.6) and (8.64), we finally write the *eccentricity vector* $\boldsymbol{c}$ as

$$\boldsymbol{c} = \mu e \hat{\boldsymbol{i}}_e \tag{8.66}$$

Given current position and velocity vectors of mass m_2 relative to m_1, Eq. (8.61) can be used to compute the periapses direction of the conic path.

The eccentricity and momentum unit direction vectors $\hat{\boldsymbol{i}}_e$ and $\hat{\boldsymbol{i}}_h$ respectively are illustrated in Figure 8.7. The orientation of the orbit plane $\mathcal{O} : \{\hat{\boldsymbol{i}}_e, \hat{\boldsymbol{i}}_p, \hat{\boldsymbol{i}}_h\}$ relative an inertial reference frame $\mathcal{N} : \{\hat{\boldsymbol{i}}_x, \hat{\boldsymbol{i}}_y, \hat{\boldsymbol{i}}_z\}$ is typically given in terms of the (3-1-3) Euler angles Ω (longitude of the ascending node), i (inclination angle) and ω (argument of the perihelion). If the orbit inclination i goes to zero, then this orbit plane orientation description is non-unique. Figure 8.7 also illustrates well the angle θ, which is defined as

$$\theta = \omega + f \tag{8.67}$$

Therefore, if the orbit plane orientation is inertially fixed, then $\dot{\omega} = 0$ and $\dot{f} = \dot{\theta}$. Another useful reference frame is the one that tracks the position of the mass m itself and is given by $\mathcal{M} : \{\hat{\boldsymbol{i}}_r, \hat{\boldsymbol{i}}_\theta, \hat{\boldsymbol{i}}_h\}$.

Example 8.1: The orbit period P can be found through a direct application of Kepler's second law. Separating variables in Eq. (8.55) we find

$$h dt = r^2 d\theta$$

After integrating both sides of the above equation over one orbit, and recognizing that the right hand side computes the area A of an ellipse, we find that

$$hP = A = 2\pi ab$$

Using Eqs. (8.12), (8.9) and (8.65), the orbit period P is then expressed in terms of the semi-major axis a and the gravitational coefficient μ as

$$P = 2\pi\sqrt{\frac{a^3}{\mu}} \tag{8.68}$$

Eq. (8.68) verifies *Kepler's third law of planetary motion* which states that the term P/a^3 is a constant.

8.3.3 Conservation of Energy

For a dynamical system containing two masses and gravity being the only force present, it is clear that the sum of the system kinetic and potential energy will remain constant (i.e. the system is conservative). We would like to investigate whether a similar "conservation of energy principle" holds for the *relative motion description* we have adopted. The following analysis is performed assuming $\boldsymbol{a}_d = 0$. The inertial kinetic energy of m_1 is

$$T_1 = \frac{1}{2} m_1 \dot{\boldsymbol{r}}_1 \cdot \dot{\boldsymbol{r}}_1 \tag{8.69}$$

with the kinetic rate expressed as

$$\dot{T}_1 = m_1 \ddot{\boldsymbol{r}}_1 \cdot \dot{\boldsymbol{r}}_1 \tag{8.70}$$

This motivates us to examine the scalar energy rate like quantity $\ddot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}}$ for the relative motion description. The relative equations of motion in Eq. (8.47) can be written as

$$\ddot{\boldsymbol{r}} = -\frac{\partial V}{\partial \boldsymbol{r}} = -\frac{\partial V}{\partial r}\frac{\partial r}{\partial \boldsymbol{r}} = -\frac{\partial V}{\partial r}\frac{1}{r}\boldsymbol{r} \tag{8.71}$$

where the relative potential energy function V is defined in Eq. (8.46). Using Eqs. (8.37) and (8.53), the quantity $\ddot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}}$ is written as

$$\ddot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}} = -\frac{\partial V}{\partial r}\frac{1}{r}(r\hat{\boldsymbol{\imath}}_r) \cdot (\dot{r}\hat{\boldsymbol{\imath}}_r + r\dot{\theta}\hat{\boldsymbol{\imath}}_\theta) = -\frac{\partial V}{\partial r}\frac{1}{r}\dot{r} \tag{8.72}$$

Both sides of this equation are then written as the prefect differentials

$$\frac{d}{dt}\left(\frac{1}{2}\dot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}}\right) = -\frac{dV}{dt} \tag{8.73}$$

After integrating both sides we conclude that the total relative energy for Keplerian motion is conserved.

$$\frac{1}{2}\dot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}} + V = \text{constant} \tag{8.74}$$

Introducing the scalar constant $\alpha/2$, we obtain the famous *energy integral*:

$$\frac{1}{2}\dot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}} = \frac{\mu}{r} + \frac{\alpha}{2} \tag{8.75}$$

The expression $\frac{1}{2}\dot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}}$ is defined to be the *relative* kinetic energy per unit mass of m_2 relative to m_1. The expression μ/r represents a gravity potential like function per unit mass. For the special case were $m_2 \ll m1$, this expression is related to the standard gravity potential function $V(r)$ through

$$-\frac{\mu}{r} = -\frac{G(m_1 + m_2)}{r} \approx -\frac{Gm_1}{r} = \frac{V(r)}{m_2} \tag{8.76}$$

The approximation made when we assume that $m_2 \ll m_1$ implies that the mass m_1 is inertially fixed. The motion of m_2 about m_1 causes negligible acceleration of m_1. To visualize such a situation, consider the space shuttle in Earth's orbit. While the shuttle motion will in theory perturb the Earth's motion, for practical purposes this effect can be neglected. However, if the we study the Earth-Moon system, then $m_2 \not\ll m_1$ and the approximation in Eq. (8.76) would not be valid. Setting $v^2 = \dot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}}$, the energy integral of the *relative motion* is written in its most popular form called the *vis-viva equation*.

$$v^2 = \mu\left(\frac{2}{r} - \alpha\right) \tag{8.77}$$

In essence, this equation relates the instantaneous scalar position and velocity of a body at any point on the orbit through the energy constant α. To express α in terms of conic intersection parameters, we examine the orbit radius r and velocity v at periapses. First, we develop α for the elliptic case. The periapses radius r_p is given in Eq. (8.10). The periapses velocity v_p is found by making use of Eqs. (8.10), (8.7) and (8.55).

$$v_p^2 = r_p^2\dot{\theta}_p^2 = \frac{r_p^4\dot{\theta}_p^4}{r_p^2} = \frac{h^2}{r_p^2} = \frac{\mu a(1-e^2)}{a^2(1-e)^2} = \frac{\mu(1+e)}{a(1-e)} \tag{8.78}$$

Using Eq. (8.77), the energy constant α is expressed as

$$\alpha_{elliptic} = \frac{2}{r} - \frac{v^2}{\mu} = \frac{2}{a(1-e)} - \frac{1+e}{a(1-e)} = \frac{(1-e)}{a(1-e)} = \frac{1}{a} \tag{8.79}$$

Therefore, for the elliptic case, the energy constant α is simply the inverse of the semi-major axis a. Since $a \to \infty$ for the parabolic case, the corresponding energy constant α goes to zero.

$$\alpha_{parabolic} = 0 \tag{8.80}$$

Example 8.2: The energy equations in Eq. (8.77) can be used to readily compute various critical velocities. The minimum velocity necessary to escape

Table 8.2: Approximate Astrometric Data for the Planets

Body	Symbol	Average Radius [km]	μ [km^3/sec^2]
Sun	$\odot$	$696\ 10^3$	$1.326\ 10^{11}$
Mercury	$\mercury$	$2.43\ 10^3$	$2.208\ 10^4$
Venus	$\venus$	$6.07\ 10^3$	$3.248\ 10^5$
Earth	$\oplus$	$6.37\ 10^3$	$3.986\ 10^5$
Moon	$\leftmoon$	$1.74\ 10^3$	$4.902\ 10^3$
Mars	$\mars$	$3.40\ 10^3$	$4.282\ 10^4$
Jupiter	$\jupiter$	$71.3\ 10^3$	$1.267\ 10^8$
Saturn	$\saturn$	$60.1\ 10^3$	$3.795\ 10^7$
Uranus	$\uranus$	$24.5\ 10^3$	$5.796\ 10^6$
Neptune	$\neptune$	$25.1\ 10^3$	$6.870\ 10^6$
Pluto	$\pluto$	$2.90\ 10^3$	$4.402\ 10^4$

the gravitational pull of a celestial body is called the *escape velocity*, which corresponds to the object being on a parabolic orbit.

Various astrometric data is given for our solar system in Table 8.2. Ignoring the atmospheric drag, the critical escape velocity magnitude v for a body on the Earth's surface would be

$$v = \sqrt{\frac{2\mu_{ext\oplus}}{r_\oplus}} = 11.06 \ \text{km/sec}$$

On the Moon, how fast would one have to propel an object horizontally such that it would never hit the Moon? This would correspond to a circular orbit just above the Moon surface (we are ignoring the lunar craters and mountains here). For a circular orbit the semi-major axis a is equal to $r_{\leftmoon}$. The corresponding velocity magnitude v is then given by

$$v = \sqrt{\frac{\mu_{\leftmoon}}{r_{\leftmoon}}} = 1.68 \ \text{km/sec}$$

which corresponds to propelling an object at roughly 1 mile per second.

For the hyperbolic case, the energy constant is rewritten using the same steps as with the elliptic case. Using Eqs. (8.27) and (8.28) along with Eq. (8.55), the constant *alpha* is expressed for the hyperbolic case as

$$\alpha_{hyperbolic} = \frac{1}{a} \tag{8.81}$$

Note that by adopting the convention that $a < 0$ for hyperbolic orbits, the algebraic expression for α is the same for a hyperbola as it is for an ellipse. The consequence of this sign convention is that we are able to write a *universally valid* vis-viva equation in terms of a for the three conic section cases.

$$v^2 = \mu\left(\frac{2}{r} - \frac{1}{a}\right) \tag{8.82}$$

The various energy levels in Eq. (8.82) clearly illustrate if the orbit trajectories of an object form a closed path or not. Note that the quantities v^2 and a are always positive quantities for the elliptic case. Therefore it is impossible for the elliptic orbit radius r to go to infinity. As the energy level increases and a grows to infinity, at the limiting parabolic case it is possible for the object to fly infinitely far away. However, it will only barely be able to reach infinity and have no remaining velocity when it gets there. An object on an hyperbolic orbit ($a < 0$) is able to "fly to infinity and beyond" since as r grows infinitely large, the object retains a positive escape velocity $v^2 = -\mu/a$.

Sometimes it is convenient to write the velocity vector $\boldsymbol{v}$ in terms of its radial and tangential components $v_r = \dot{r}$ and $v_\theta = r\dot{\theta}^2$ respectively. Using Eqs. (8.55) and (8.65), v_θ is written as

$$v_\theta^2 = \frac{\mu p}{r^2} \tag{8.83}$$

Using the energy equation in Eq. (8.82), the corresponding radial velocity components is given by

$$v_r^2 = v^2 - v_\theta^2 = \mu\left(\frac{2r-p}{r^2} - \frac{1}{a}\right) \tag{8.84}$$

Example 8.3: We would like to study the ballistic missile problem of launching an unpowered object and hitting a specific spot somewhere else on Earth. For this, we revisit Example 2.2 where a mass is launched under a *constant* gravity field. There, for a given initial velocity v_0, there were typically two corresponding initial firing angles that would match the target condition. When there was only one possible initial angle γ_0, then the *minimum* initial velocity (energy) was being used.

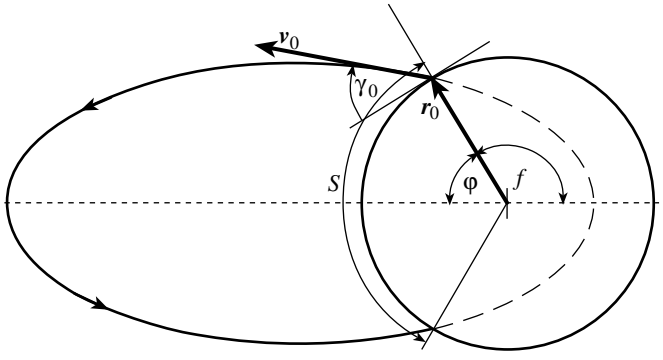


Figure 8.8: Ballistic Missile Trajectory

Similar results are found when a mass is subjected to an *inverse-squared* gravity field such as is the case with sub-orbital flights. Figure 8.8 illustrates one suborbital trajectory. Ignoring Earth's rotation, we can assume Keplerian motion. Clearly the initial missile velocity v_0 would need to be less than

the corresponding Earth's escape velocity, otherwise the missile would never return back to Earth. Since the missile mass is much less than Earth's mass, the resulting elliptic motion will have its focus at Earth's center which can also serve as an inertial origin. The desired range S is related to the *semi-range angle* φ through[3]

$$\varphi = \frac{1}{2}\frac{S}{R_\oplus}$$

For the following analysis, it is easier to work with the semi-range angle φ than with the true anomaly f. First we have to relate φ with the initial launch states $\boldsymbol{r}_0$ and $\boldsymbol{v}_0$. We do this using the orbit elements a and e. Since $\varphi = \pi - f$, we can use Eq. (8.6) and launch conditions to find

$$\cos\varphi = -\cos f = \frac{1}{e}\left(1 - \frac{p}{R_\oplus}\right)$$

The initial position and velocity vector are written in $\mathcal{O}$ frame components as

$$\begin{aligned}\boldsymbol{r}_0 &= R_\oplus \hat{\boldsymbol{\imath}}_r\\ \boldsymbol{v}_0 &= v_0 \sin\gamma_0 \hat{\boldsymbol{\imath}}_r + v_0\cos\gamma_0 \hat{\boldsymbol{\imath}}_\theta\end{aligned}$$

Using $\boldsymbol{h} = \boldsymbol{r}\times\boldsymbol{v}$, we express the constant scalar angular momentum h as

$$h = R_\oplus v_0 \cos\gamma_0$$

The semi-latus rectum is then found to be

$$p = \frac{h^2}{\mu} = R_\oplus \frac{v_0^2}{\mu/R_\oplus}\cos^2\gamma_0 = R_\oplus \nu_0^2\cos^2\gamma_0$$

where ν_0 is the initial velocity normalized by Earth's circular orbit speed. Having $\nu_0^2 = 1$ means the object has enough energy to be able to achieve a circular orbit. Having $\nu_0^2 \geq 2$ means the the object has exceeded the escape velocity and is leaving Earth on either a parabolic or hyperbolic trajectory. Using the energy equation at launch conditions and the definition for ν_0, the elliptic semi-major axis a is expressed as

$$a = \frac{R_\oplus}{2-\nu_0^2}$$

Finally, using Eq. (8.9), the eccentricity e is expressed in terms of initial launch conditions as

$$e = \sqrt{1 - \nu_0^2(2-\nu_0^2)\cos^2\gamma_0}$$

Using these orbit elements we are able to express $\cos\varphi$ as

$$\cos\varphi = \frac{1-\nu_0^2\cos^2\gamma_0}{\sqrt{1-\nu_0^2(2-\nu_0^2)\cos^2\gamma_0}}$$

Using standard trigonometric identities, we can convert this expression into

$$\sin\varphi = \frac{\nu_0^2\sin 2\gamma_0}{\sqrt{2\left(2+\nu_0^2(2-\nu_0^2)(\cos 2\gamma_0 - 1)\right)}}$$

Either of these expressions relate initial launch conditions γ_0 and ν_0 to the desired semi-range angle φ. Note however that while $\nu_0^2 \geq 2$ still mathematically provide corresponding semi-range angles, since escape velocity is achieved, this situation corresponds to having the object fly through the Earth as is illustrated in the dashed line in Figure 8.8. Having a clear flight path is not enforced in these equations and must be checked separately. Combining the $\sin$ and $\cos$ expressions of φ, we are able to obtain a very compact expression for the semi-range angle.[4]

$$\tan\varphi = \frac{\nu_0^2 \tan\gamma_0}{1 - \nu_0^2 + \tan^2\gamma_0} \tag{8.85}$$

Assume that the initial velocity ν_0 is fixed and we intend to maximize the range φ. After taking the derivative of Eq. (8.85) with respect to γ_0 and setting it to zero, we find that the corresponding optimal initial launch angle is given through

$$\gamma_{0_{opt}} = \cos^{-1}\left(\pm\sqrt{\frac{1}{2-\nu_0^2}}\right) \tag{8.86}$$

Note that Eq. (8.86) only provides real answers for $0 \leq \nu_0^2 \leq 1$. The reason for this is that once a circularizing orbit speed is achieved, *any* location on Earth could be reached. Maximizing the range for velocities beyond this has no meaning. For very small initial velocities we can assume that $\nu_0^2 \approx 0$. This simplifies Eq. (8.86) to the constant gravity field case and yields the well-known optimal launch angles 45 and 135 degrees, depending on which direction one is launching the projectile. As ν_0^2 reaches 1 and the projectile achieves Earth's circularizing orbit speed, any point on Earth can be reached. The optimal launch angles for this case are γ_0 being either 0 or 180 degrees. The corresponding maximum semi-range angle for a given ν_0^2 is is found by back substituting $\gamma_{0_{opt}}$ into Eq. (8.85).

$$\tan\varphi_{max} = \frac{\nu_0^2}{2\sqrt{1-\nu_0^2}} \tag{8.87}$$

For a given feasible ν_0 and φ there are generally two possible initial launch angles γ_0. Solving Eq. (8.85) for γ_0 we find

$$\tan\gamma_0 = \frac{\nu_0^2 \pm \sqrt{\nu_0^4 - 4(1-\nu_0^2)\tan^2\varphi}}{2\tan\varphi} \tag{8.88}$$

Only one launch angle exists if the discriminant in Eq. (8.88) is zero. This condition represents a minimum energy trajectory to achieve the desired range. Setting the discriminant equal to zero we are able to retrieve Eq. (8.87). Solving this equation for ν_0^2 we find an expression for the minimum velocity necessary to achieve a given φ.

$$\nu_{0_{min}}^2 = 2\tan^2\varphi\left(\frac{1}{|\sin\varphi|} - 1\right) \tag{8.89}$$

For given initial velocities ν_0^2 and desired semi-range angles φ, Figure 8.9 illustrates the various corresponding initial launch angles γ_0. As ν_0^2 increases

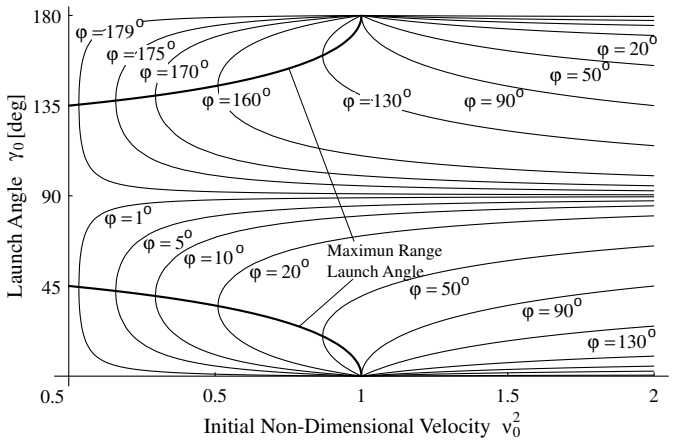


Figure 8.9: Comparison of Initial Velocities and Corresponding Launch Angles for Various Ranges

up to an value of 1, the corresponding launch angles for maximum range decrease from 45 degrees to zero degrees. For $\nu_0^2 \leq 1$ there are always two possible trajectories for a given velocity, a "high" and a "low" path. However, once $\nu_0^2 \geq 1$ the two possible trajectories each lie on opposite sides of the Earth.

The total relative energy of a body is universally expressed from the vis-viva equation in Eq. (8.82) as

$$\frac{v^2}{2} - \frac{\mu}{r} = -\frac{2\mu}{a} \tag{8.90}$$

where $-2\mu/a$ is referred to as the *total energy constant*. Observe that elliptic orbits with $a > 0$ have a negative total energy. Parabolic orbits have zero total energy since $a \to \infty$ and hyperbolic orbits with $a < 0$ have positive orbits. To minimize the total energy of an elliptic orbit, the semi-major axis a would need to made as small as possible.

8.4 Classical Solutions

The previous fundamental integrals are all used to describe the instantaneous state of an orbit. What is lacking is a method to determine the location of an object within the orbit itself at any instance of time. This section presents various classical solutions to the problem of solving the nonlinear relative differential equations of motion in Eq. (8.45).

8.4.1 Kepler's Equation

To determine the angular orbit position at any instance of time, we rewrite the angular momentum expression into a form that can easily be integrated. First, a frontal assault to this problem is presented that illustrates why using the true anomaly f is not attractive for this task. For the case where no disturbance acceleration $\boldsymbol{a}_d$ is present, the orbit angular momentum magnitude h is constant and $\dot{f} = \dot{\theta}$. Using Eq. (8.55) we find

$$h = r^2\dot{\theta} = r^2\dot{f} \tag{8.91}$$

which is rearranged into the form

$$hdt = r^2 df \tag{8.92}$$

Substituting Eqs. (8.6) and (8.65) into the above expression, we find the differential equation

$$\sqrt{\frac{\mu}{p^3}}dt = \frac{df}{(1+e\cos f)^2} \tag{8.93}$$

which is integrated from the initial time t_0 to another time t_1.

$$\sqrt{\frac{\mu}{p^2}}(t_1 - t_0) = \int_{f_0}^{f_1} \frac{df}{(1+e\cos f)^2} \tag{8.94}$$

The left hand side of Eq. (8.93) is easily be integrated. However, analytically solving the right hand side involves finding a solution to a non-standard elliptic integral; clearly not a very attractive proposition. By describing the angular position within the orbit through the eccentric anomaly E instead of the true anomaly f, we are able to replace the differential equation in Eq. (8.93) with an equivalent expression which can then be easily integrated. Once again we start with the massless angular momentum vector expression

$$\boldsymbol{h} = \boldsymbol{r} \times \dot{\boldsymbol{r}} \tag{8.95}$$

which is inertially fixed for $\boldsymbol{a}_d = 0$. The position vector $\boldsymbol{r}$ and velocity vector $\dot{\boldsymbol{r}}$ are written in $\mathcal{O}$ frame components as

$$\boldsymbol{r} = x\hat{\boldsymbol{\imath}}_e + y\hat{\boldsymbol{\imath}}_p \tag{8.96}$$

$$\dot{\boldsymbol{r}} = \dot{x}\hat{\boldsymbol{\imath}}_e + \dot{y}\hat{\boldsymbol{\imath}}_p \tag{8.97}$$

since the $\mathcal{O}$ frame unit vectors are inertially fixed. Substituting Eqs. (8.96) and (8.97) into Eq. (8.95) yields

$$\boldsymbol{h} = (x\dot{y} - y\dot{x})\hat{\boldsymbol{\imath}}_h = h\hat{\boldsymbol{\imath}}_h \tag{8.98}$$

Substituting the $x(E)$ and $y(E)$ expressions in Eq. (8.16) and (8.17) along with their derivatives into the h expression in Eq. (8.98), we find that

$$h = a^2\sqrt{1-e^2}\left(\cos^2 E + \sin^2 E - e\cos E\right)\frac{dE}{dt} \tag{8.99}$$

Using the standard trigonometric identity $\cos^2 E + \sin^2 E = 1$ and Eqs. (8.9) and (8.65), we are now able to replace Eq. (8.93) with the more attractive differential equation

$$\sqrt{\frac{\mu}{a^3}}dt = (1 - e\cos E)dE \tag{8.100}$$

This differential equation can be rewritten to provide a convenient expression for $\dot{E}$.

$$\frac{dE}{dt} = \dot{E} = \sqrt{\frac{\mu}{a^3}}\frac{1}{(1 - e\cos E)} = \frac{1}{r}\sqrt{\frac{\mu}{a}} \tag{8.101}$$

Integrating Eq. (8.100) we obtain the famous *Kepler's equation*.

$$\sqrt{\frac{\mu}{a^3}}(t_1 - t_0) = (E - e\sin E)|_{E_0}^{E_1} \tag{8.102}$$

Given some initial time t_0, eccentric anomaly E_0 and a current time t_1, Kepler's equation is solved for the current eccentric anomaly E using a numerical method such as Newton's method. Setting $E_0 = 0$ and $E_1 = 2\pi$ we are able to verify the orbit period equation in Eq. (8.68). Let us introduce the *mean anomaly* M as

$$M = M_0 + n(t - t_0) \tag{8.103}$$

with $0 \leq M \leq 2\pi$ and the *mean angular motion* n as

$$n = \sqrt{\frac{\mu}{a^3}} = \frac{2\pi}{P} \tag{8.104}$$

where P is the orbit period, then Kepler's equation of Eq. (8.102) is written in its classical form as

$$M = M_0 + n(t_1 - t_0) = E - e\sin E \tag{8.105}$$

Clearly it is more convenient to use the eccentric or mean anomaly instead of the true anomaly to describe a position within the orbit. Using Eq. (8.22) and the eccentricity e, the eccentric anomaly E can always be translated back into the true anomaly f if necessary.

Example 8.4: The following example illustrates how well Newton's method is suited to numerically solve Kepler's equation for the eccentric anomaly E that corresponds to a given mean anomaly M. Once a mean anomaly M is computed using Eq. (8.105) for a given mean angular motion n and time t_1, we must solve the nonlinear equation

$$f(E) = M - (E - e\sin E) = 0$$

for the corresponding eccentric anomaly E. Given an initial guess $\hat{E}$ for the true E, Newton's method computes the step correction ΔE through

$$\Delta E = -\frac{f(\hat{E})}{f'(\hat{E})}$$

where $f'(\hat{E})$ is given by

$$f'(\hat{E}) = \left.\frac{df}{dE}\right|_{\hat{E}} = -(1 - e\cos\hat{E}) = -\frac{r(\hat{E})}{a}$$

Typically, setting the initial value of $\hat{E}$ equal to M provides a good starting point for the numerical iteration and results in a fast convergence rate. The reason for this is evident in Figure 8.10 which compares the mean versus the eccentric anomaly for eccentricities ranging from $e = 0$ (circular case) to $e = 1$ (parabolic case). For circular or near-circular orbits, assuming $M \approx E$ is clearly a very good initial guess. However, even for the limiting parabolic case do the eccentric anomalies remain relatively close to the mean anomalies.

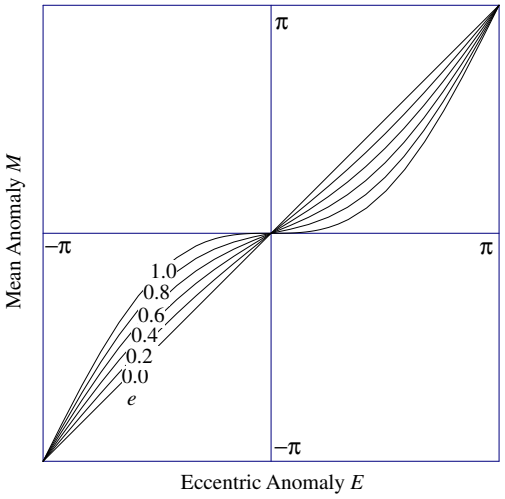


Figure 8.10: Mean Anomaly versus Eccentric Anomaly for Various Eccentricities

To demonstrate the good convergence characteristics of Newton's method, we numerically solve for the eccentric anomaly E corresponding to a mean anomaly M of 1.5. The eccentricity e is set to 0.8, a rather large, near-parabolic value. For near-circular cases with small values of e, Figure 8.10 already illustrates that $M \approx E$.

Table 8.3 shows the convergence for each iteration step. Even though the eccentricity e is rather large, after *only* three integration steps the eccentric anomaly estimate $\hat{E}$ is already accurate up to 5 significant digits. After just two more iteration steps, the estimate is sufficiently accurate for double precision arithmetic.

Table 8.3: Iteration Steps of Applying Newton's Method to Numerically Solve Kepler's Equation

Iteration Step	ΔE	$\hat{E}$
0		1.500000000000000
1	8.45863 10^{-1}	2.345863185046372
2	-1.75896 10^{-1}	2.169967661788265
3	-6.42592 10^{-3}	2.163541744525667
4	-9.44057 10^{-6}	2.163532303960638
5	-2.04353 10^{-11}	2.163532303940202
6	-1.53462 10^{-16}	2.163532303940202

To find Kepler's equation for a hyperbolic orbit, we substitute the x and y definitions in terms of the hyperbolic anomaly H in Eqs. (8.31) and (8.32) along with their derivatives

$$\dot{x} = -a \sinh H \dot{H} \tag{8.106}$$

$$\dot{y} = a\sqrt{e^2 - 1} \cosh H \dot{H} \tag{8.107}$$

into the angular momentum expression in Eq. (8.98). After making use of Eqs. (8.28) and (8.65), the hyperbolic anomaly derivative is given by

$$\frac{dH}{dt} = \frac{n}{e \cosh H - 1} = \frac{1}{r}\sqrt{\frac{\mu}{a}} \tag{8.108}$$

Kepler's equation is then found after separating variables in Eq. (8.108) and integrating both sides to be

$$n(t - t_0) = (e \sinh H - H)|_{H_0}^{H_1} = N(t_1) - N(t_0) \tag{8.109}$$

where $N = e \sinh H - H$ is the hyperbolic equivalent of the elliptic mean anomaly.

8.4.2 Orbit Elements

Given some initial conditions $\boldsymbol{r}(t_0)$ and $\dot{\boldsymbol{r}}(t_0)$ and a current time t_1, the second order relative differential equation of motion in Eq. (8.45) can be solved for any current $\boldsymbol{r}(t_1)$ and $\dot{\boldsymbol{r}}(t_1)$ vectors. Note that the six scalar vector components of the initial conditions are invariants of the solution, similar to the constant fundamental integrals introduced in the previous section. These orbit constants determine the size and shape of the orbit trajectory. The time t_1 indicates where an object is within the orbit.

This behavior is universally true. Any two-body orbit geometry can be described through six scalar constants with the body position within the orbit described through a time-like variable. Therefore, instead of having six degrees of freedom in (8.45), there is really only one degree of freedom in a fixed orbit. By choosing other system invariants than $\boldsymbol{r}(t_0)$ and $\dot{\boldsymbol{r}}(t_0)$, the differential equations

of motion in Eq. (8.45) can be replaced with simpler expressions For example, given the orbit semi-major axis a and eccentricity e, Kepler's equation replaces the second order relative differential equation of motion with a scalar (i.e. one degree of freedom), algebraic relationship between the time t_1 and the mean anomaly M.

The six orbit invariants are called the *orbit elements*. Any six orbit constants can be used for this purpose. A commonly used set of orbit elements are

$$\{a, e, i, \Omega, \omega, M_0\} \tag{8.110}$$

The first two invariants a and e determine the orbit size and shape. The following three scalars Ω, i and ω are the (3-1-3) Euler angles which define the orbit plane orientation. Finally, the mean anomaly M_0 specifies where the object is within the orbit trajectory at time t_0. To translate $\boldsymbol{r}(t_0)$ and $\dot{\boldsymbol{r}}(t_0)$ into the orbit elements in Eq. (8.110), the following steps are taken. The semi-major axis a is found by first finding

$$r_0 = \sqrt{\boldsymbol{r}(t_0) \cdot \boldsymbol{r}(t_0)} \tag{8.111}$$

$$v_0^2 = \dot{\boldsymbol{r}}(t_0) \cdot \dot{\boldsymbol{r}}(t_0) \tag{8.112}$$

and then using the energy equation:

$$\frac{1}{a} = \frac{2}{r_0} - \frac{v_0^2}{\mu} \tag{8.113}$$

The eccentricity e is found by first computing the constant vector c

$$\boldsymbol{c} = \dot{\boldsymbol{r}}_0 \times \boldsymbol{h} - \mu \frac{\boldsymbol{r}_0}{r_0} \tag{8.114}$$

and then calculating

$$e = \frac{|\boldsymbol{c}|}{\mu} \tag{8.115}$$

Let the direction cosine matrix $[C]$ map inertial $\mathcal{N}$ frame vectors into orbit frame $\mathcal{O}$ vectors. Using Eq. (3.5), the $\mathcal{O}$ frame unit vectors are found through

$$\hat{\boldsymbol{i}}_e = \boldsymbol{c}/\mu e = C_{11}\hat{\boldsymbol{i}}_x + C_{12}\hat{\boldsymbol{i}}_y + C_{13}\hat{\boldsymbol{i}}_z \tag{8.116}$$

$$\hat{\boldsymbol{i}}_p = \hat{\boldsymbol{i}}_h \times \hat{\boldsymbol{i}}_e = C_{21}\hat{\boldsymbol{i}}_x + C_{22}\hat{\boldsymbol{i}}_y + C_{23}\hat{\boldsymbol{i}}_z \tag{8.117}$$

$$\hat{\boldsymbol{i}}_h = \boldsymbol{h}/h = C_{31}\hat{\boldsymbol{i}}_x + C_{32}\hat{\boldsymbol{i}}_y + C_{33}\hat{\boldsymbol{i}}_z \tag{8.118}$$

Given the direction cosine matrix elements, the corresponding (3-1-3) Euler angles are found using Eq. (3.36).

$$\Omega = \tan^{-1}\left(\frac{C_{31}}{-C_{32}}\right) \tag{8.119}$$

$$i = \cos^{-1}(C_{33}) \tag{8.120}$$

$$\omega = \tan^{-1}\left(\frac{C_{13}}{C_{23}}\right) \tag{8.121}$$

To find the initial mean anomaly M_0, we define $\sqrt{\mu}\boldsymbol{\sigma} = \boldsymbol{r} \cdot \dot{\boldsymbol{r}}$, so the constant σ_0 is

$$\sigma_0 \equiv \frac{\boldsymbol{r}_0 \cdot \dot{\boldsymbol{r}}_0}{\sqrt{\mu}} = \frac{r_0 \dot{r}_0}{\sqrt{\mu}} \tag{8.122}$$

Using Eqs. (8.15) and (8.101), σ_0 can also be written as

$$\sigma_0 = \sqrt{a} e \sin E_0 \tag{8.123}$$

However, if Eq. (8.123) were solved directly for E_0, then we would have to deal with the quadrant issues of the $\sin^{-1}$ function. Instead, we use Eqs. (8.15) and (8.101)to express the initial eccentric anomaly E_0 in terms of σ_0 and r_0 through

$$E_0 = \tan^{-1}\left(\frac{\sigma_0/\sqrt{a}}{1 - r_0/a}\right) \qquad -\pi \leq E_0 \leq \pi \tag{8.124}$$

By making use of the numerical function $\tan2(x, y)$, no quadrant problems arise with this formula. The initial mean anomaly is then found through

$$M_0 = E_0 - e \sin E_0 \tag{8.125}$$

The reverse process of this orbit element transformation is posed as an exercise problem at the end of this chapter.

Example 8.5: The scalar parameter $\sigma(t)$ is defined through

$$\sigma(t) = \frac{\boldsymbol{r}(t) \cdot \dot{\boldsymbol{r}}(t)}{\sqrt{\mu}} \tag{8.126}$$

and provides a measure of orthogonality between the instantaneous position and velocity vector. Note that $\sigma(t)$ is zero at apoapses and periapses of elliptic orbits and at any point of a circular orbit. The second derivative of σ assumes a very familiar form. Differentiating σ we find

$$\dot{\sigma} = \frac{\dot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}}}{\sqrt{\mu}} + \frac{\boldsymbol{r} \cdot \ddot{\boldsymbol{r}}}{\sqrt{\mu}}$$

After substituting Eqs. (8.45) and (8.113), the σ rate expression reduces to

$$\dot{\sigma} = \sqrt{\mu}\left(\frac{v^2}{\mu} - \frac{1}{r}\right) = \sqrt{\mu}\left(\frac{1}{r} - \frac{1}{a}\right)$$

Differentiating $\dot{\sigma}$ we find

$$\ddot{\sigma} = -\sqrt{\mu}\frac{\dot{r}}{r^2} = -\frac{\mu}{r^3}\frac{r\dot{r}}{\sqrt{\mu}} = -\frac{\mu}{r^3}\frac{\boldsymbol{r} \cdot \dot{\boldsymbol{r}}}{\sqrt{\mu}}$$

which, after using Eq. (8.126), leads to the familiar algebraic expression

$$\ddot{\sigma} = -\frac{\mu}{r^3}\sigma \tag{8.127}$$

Therefore the scalar $\ddot{\sigma}$ differential equation has the same algebraic form as the relative equations of motion in Eq. (8.45). Any scalar differential equation of the algebraic form given in Eq. (8.127) automatically has a corresponding invariant vector. The vector $\boldsymbol{c}_2$, defined through

$$\boldsymbol{c}_2 = \dot{\sigma}\boldsymbol{r} - \sigma\dot{\boldsymbol{r}}$$

is inertially fixed since

$$\dot{\boldsymbol{c}}_2 = \dot{\sigma}\dot{\boldsymbol{r}} + \ddot{\sigma}\boldsymbol{r} - \dot{\sigma}\dot{\boldsymbol{r}} - \sigma\ddot{\boldsymbol{r}} = 0$$

This fixed vector $\boldsymbol{c}_2$ can be found for any scalar parameter which satisfies Eq. (8.127). To geometrically interpret this vector, we perform the dot product of $\boldsymbol{c}_2$ with the position vector $\boldsymbol{r}$.

$$\boldsymbol{r}\cdot\boldsymbol{c}_2 = \boldsymbol{r}\cdot(\dot{\sigma}\boldsymbol{r} - \sigma\dot{\boldsymbol{r}}) = \left(\sqrt{\mu}\left(\frac{v^2}{\mu} - \frac{1}{r}\right)r^2 - \frac{(r\dot{r})^2}{\sqrt{\mu}}\right)$$

Evaluating the expression $\boldsymbol{r}\cdot(\dot{\boldsymbol{r}}\times\boldsymbol{h})$ it can be shown that

$$(r\dot{r})^2 = v^2r^2 - h^2$$

Using this identity and the fact that $\boldsymbol{r}\cdot\boldsymbol{c}_2 = r|\boldsymbol{c}_2|\cos(\angle\boldsymbol{r},\boldsymbol{c}_2)$, the orbit radius r is expressed as

$$r = \frac{h^2/\mu}{1 + \frac{|\boldsymbol{c}_2|}{\sqrt{\mu}}\cos(\angle\boldsymbol{r},\boldsymbol{c}_2)}$$

Comparing this expression to Eq. (8.6), it is evident that $\boldsymbol{c}_2$ must be given by

$$\boldsymbol{c}_2 = e\sqrt{\mu}\,\hat{\boldsymbol{i}}_e$$

It is related to the eccentricity vector $\boldsymbol{c}$ in Eq. (8.66) in that it also points towards periapses, but has a different vector magnitude.

Besides the set presented in Eq. (8.110), many other orbit element sets are possible. For example, another feasible orbit element set could be given by

$$\{h, e, \Omega, i, \omega, f\} \tag{8.128}$$

with the true anomaly f acting as the time-like parameter. The orbit shape is determined through the parameters h and e. The orbit plane orientation is again determined through the (3-1-3) Euler angle set $\{\Omega, i, \omega\}$. Note that there is no sixth parameter explicitly specifying the initial position within the orbit. Since f forms in essence our time variable, implicitly when $f = 0$ then we have $M_0 = E_0 = 0$. Let the frame $\mathcal{M} = \{\hat{\boldsymbol{i}}_r, \hat{\boldsymbol{i}}_\theta, \hat{\boldsymbol{i}}_h\}$ have its first direction unit vector $\hat{\boldsymbol{i}}_r$ track the position of the body m as shown in Figure 8.7. The relative position vector $\boldsymbol{r}$ is expressed in $\mathcal{M}$ frame components as

$$\boldsymbol{r} = {}^{\mathcal{M}}\begin{pmatrix} r \\ 0 \\ 0 \end{pmatrix} \tag{8.129}$$

The orientation of the $\mathcal{M}$ frame relative to the inertial $\mathcal{N}$ frame is given through the (3-1-3) Euler angle set $\{\Omega, i, \theta\}$ with $\theta = \omega + f$. Using the (3-1-3) Euler angle parameterization of the direction cosine matrix in Eq. (3.35), the inertial vector components of the position vector $\boldsymbol{r}$ are given by

$$\boldsymbol{r} = r \begin{pmatrix} \cos\Omega\cos\theta - \sin\Omega\sin\theta\cos i \\ \sin\Omega\cos\theta + \cos\Omega\sin\theta\cos i \\ \sin\theta\sin i \end{pmatrix}^{\mathcal{N}} \tag{8.130}$$

Therefore, given the five constant orbit elements in Eq. (8.128) and a true anomaly f, Eq. (8.130) provides the current inertial position vector components directly without having to perform any numerical iterations. Kepler's equation is used later to correlate the current f with a corresponding time. Using Eqs. (8.6) and (8.65), the derivative of Eq. (8.130) is written in the relatively simple form

$$\dot{\boldsymbol{r}} = -\frac{\mu}{h} \begin{pmatrix} \cos\Omega(\sin\theta + e\sin\omega) + \sin\Omega(\cos\theta + e\cos\omega)\cos i \\ \sin\Omega(\sin\theta + e\sin\omega) - \cos\Omega(\cos\theta + e\cos\omega)\cos i \\ -(\cos\theta + e\cos\omega)\sin i \end{pmatrix}^{\mathcal{N}} \tag{8.131}$$

Again, given a true anomaly f, the corresponding inertial velocity vector is readily computed.

Given the Eqs. (8.3) and (8.4), the position vector can $\boldsymbol{r}$ is written as a direct function of the true anomaly f as

$$\boldsymbol{r} = r\cos f\hat{\boldsymbol{i}}_e + r\sin f\hat{\boldsymbol{i}}_p \tag{8.132}$$

where the unit vectors $\hat{\boldsymbol{i}}_e$ and $\hat{\boldsymbol{i}}_p$ are illustrated in Figure 8.7. Using Eqs. (8.55) and (8.65), the velocity vector $\boldsymbol{v}$ is then expressed as a direct function of the true anomaly f as

$$\boldsymbol{v} = -\frac{\mu}{h}\sin f\hat{\boldsymbol{i}}_e + \frac{\mu}{h}(e + \cos f)\hat{\boldsymbol{i}}_p \tag{8.133}$$

Another popular set of orbit elements is the Delaunay variable set given by

$$\{l, g, h, L, G, H\} \tag{8.134}$$

where small letters indicate orientation-type quantities and the capital letters represent the corresponding generalized momentas. In terms of the previous orbit elements, they are defined as[5]

$l \equiv$ Mean Anomaly M
$g \equiv$ Argument of the Pericenter ω
$h \equiv$ Longitude of Ascending Node Ω
$L = \sqrt{\mu a}$
$G = L\sqrt{1-e^2} = \sqrt{\mu a(1-e^2)} \equiv$ Angular Momentum Magnitude
$H = G\cos i \equiv$ Ang. Mom. Component Normal to Equitorial Plane

These variables are popular because they are *canonical variables* that abide by the differential equations

$$\frac{dL}{dt} = \frac{\partial \mathcal{H}}{\partial l} \qquad \frac{dl}{dt} = -\frac{\partial \mathcal{H}}{\partial L} \tag{8.135a}$$

$$\frac{dG}{dt} = \frac{\partial \mathcal{H}}{\partial g} \qquad \frac{dg}{dt} = -\frac{\partial \mathcal{H}}{\partial G} \tag{8.135b}$$

$$\frac{dH}{dt} = \frac{\partial \mathcal{H}}{\partial h} \qquad \frac{dh}{dt} = -\frac{\partial \mathcal{H}}{\partial H} \tag{8.135c}$$

Where the scalar $\mathcal{H}$ is the system Hamiltonian. For a perfectly spherical planet, the Hamiltonian of a small satellite orbiting this body is given by

$$\mathcal{H} = \frac{\mu}{2a} = \frac{\mu^2}{2L^2} \tag{8.136}$$

The beauty of these variables is that their simple differential equations in Eq. (8.135) can easily be modified to encompass situations other than orbits about spherical bodies. For example, it is possible to extend the Hamiltonian $\mathcal{H}$ to incorporate Earth oblateness effects.

To verify the differential equations of Eq. (8.135), let us verify that all the orbit elements except for the mean anomaly l are constant for the case of a spherical Earth. Since the Hamiltonian in Eq. (8.136) only depends on L, then $\dot{L}$, $\dot{G}$, $\dot{H}$, $\dot{g}$ and $\dot{h}$ are zero. The only non-zero quantity is $\dot{l}$ given by

$$\dot{l} = -\frac{\partial \mathcal{H}}{\partial L} = \frac{\mu^2}{L^3} = \sqrt{\frac{\mu}{a^3}} = n \tag{8.137}$$

which agrees with Kepler's equation. If oblateness effects were included in the Hamiltonian expression, than we would find that other variables would be time varying too.

While the physical interpretations of the classical orbit elements shown in Eq. (8.110) are easy to visualize, this set of orbit elements often lead to singular equations as the eccentricity to the orbit inclination angle tend to zero. Prof. Roger A. Broucke developed a set of orbit elements called the *equinoctial variables* which are non-singular and don't lead to any mathematical singularities for any eccentricity or orbit inclination angle. To do so, we define the *longitude of pericenter* as

$$\varpi = \omega + \Omega \tag{8.138}$$

Instead of using the mean or true anomaly as the time dependent quantity, the *mean longitude* ϑ is used instead. The mean longitude is defined here as the sum of the argument of perigee, ascending node and the mean anomaly.

$$\vartheta = \omega + \Omega + M = \varpi + M \tag{8.139}$$

The equinoctial element set is given by the parameters

$$\{a, P_1, P_2, Q_1, Q_2, \vartheta\} \tag{8.140}$$

where the elements P_i and Q_i are defined in terms of the classical orbit elements as

$$P_1 = e\sin\varpi \qquad P_2 = e\cos\varpi \tag{8.141}$$

$$Q_1 = \tan\frac{i}{2}\sin\Omega \qquad Q_2 = \tan\frac{i}{2}\cos\Omega \tag{8.142}$$

The inverse transformation is given by

$$e = \sqrt{P_1^2 + P_2^2} \tag{8.143}$$

$$i = 2\tan^{-1}\left(Q_1^2 + Q_2^2\right) \tag{8.144}$$

$$\varpi = \tan^{-1}\left(\frac{P_1}{P_2}\right) \tag{8.145}$$

$$\Omega = \tan^{-1}\left(\frac{Q_1}{Q_2}\right) \tag{8.146}$$

$$M = \vartheta - \tan^{-1}\left(\frac{P_1}{P_2}\right) \tag{8.147}$$

8.4.3 Lagrange/Gibbs F and G Solution

The orbit plane is defined through the two initial condition vectors $\boldsymbol{r}(t_0)$ and $\dot{\boldsymbol{r}}(t_0)$ as shown in Figure 8.11. Since *any* orbit position vector $\boldsymbol{r}(t)$ and velocity vector $\dot{\boldsymbol{r}}(t)$ will lie in the orbit plane, they can be expressed as a linear combination of the initial condition vectors as

$$\boldsymbol{r}(t) = F\boldsymbol{r}(t_0) + G\dot{\boldsymbol{r}}(t_0) \tag{8.148}$$

$$\dot{\boldsymbol{r}}(t) = \dot{F}\boldsymbol{r}(t_0) + \dot{G}\dot{\boldsymbol{r}}(t_0) \tag{8.149}$$

where F and G are yet to be determined functions. Substituting Eqs. (8.148) and (8.149) into the relative differential equations of motion in Eq. (8.45) and making use of the fact that the initial position and velocity vectors are arbitrary, one immediate property of these functions is found to be

$$\ddot{F} = -\frac{\mu}{r^3}F \tag{8.150}$$

$$\ddot{G} = -\frac{\mu}{r^3}G \tag{8.151}$$

with the initial conditions being

$$F(t_0) = 1 \qquad \dot{F}(t_0) = 0 \tag{8.152}$$

$$G(t_0) = 0 \qquad \dot{G}(t_0) = 1 \tag{8.153}$$

Note that the second order differential equations for both F and G have the same algebraic form as the relative equations of motion.

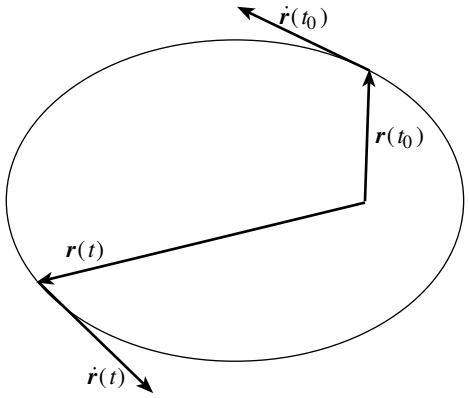


Figure 8.11: Orbit Plane Illustration

A brute force approach to solving for the F and G functions would be to attempt a power series solution of the form

$$\boldsymbol{r}(t) = \boldsymbol{r}(t_0) + \sum_{n=1}^{\infty} \frac{(t-t_0)^n}{n!} \left.\frac{d^n \boldsymbol{r}}{dt^n}\right|_{t_0} \tag{8.154}$$

Battin shows in Ref. 2 that it is indeed possible to find an algebraic recursive solution for the power series coefficients. While the development of this method yields many interesting insights, the power series solution has the drawback of having a slow convergence rate and has therefore less practical value. Instead, the following method provides exact analytic expressions for both F and G for arbitrary $(t - t_0)$. Using the $\mathcal{O}$ frame position vector components (x, y) in Eq. (8.96), we can write Eq. (8.148) as

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} x_0 & \dot{x}_0 \\ y_0 & \dot{y}_0 \end{bmatrix} \begin{pmatrix} F \\ G \end{pmatrix} \tag{8.155}$$

Solving Eq. (8.155) for F and G we find

$$\begin{pmatrix} F \\ G \end{pmatrix} = \frac{1}{(x_0\dot{y}_0 - y_0\dot{x}_0)} \begin{bmatrix} \dot{y}_0 & -\dot{x}_0 \\ -y_0 & x_0 \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{8.156}$$

Using Eqs. (8.65) and (8.98), the determinant is rewritten as

$$x_0\dot{y}_0 - y_0\dot{x}_0 = h = \sqrt{\mu p} \tag{8.157}$$

Therefore the F and G functions are given explicitly through

$$F = \frac{1}{\sqrt{\mu p}}(x\dot{y}_0 - y\dot{x}_0) \tag{8.158}$$

$$G = \frac{1}{\sqrt{\mu p}}(yx_0 - xy_0) \tag{8.159}$$

In Eqs. (8.16) and (8.17) we have already found expressions of position vector orbit plane components x and y in terms of the orbit elements a, e and the eccentric anomaly E. This allows us to find analytical expressions of F and G in terms of E. Using the $\dot{E}$ expression in Eq. (8.101), the velocity components $\dot{x}$ and $\dot{y}$ are found to be

$$\dot{x} = -\frac{\sqrt{\mu a}}{r} \sin E \tag{8.160}$$

$$\dot{y} = \frac{\sqrt{\mu a(1-e^2)}}{r} \cos E \tag{8.161}$$

Upon substituting Eqs. (8.16), (8.17), (8.160) and (8.161) into Eq. (8.156) we find the desired exact, analytical solution of the function F to be

$$F = 1 - \frac{a}{r_0}(1 - \cos \hat{E}) \tag{8.162}$$

where $\hat{E} \equiv E - E_0$ is defined as the change in eccentric anomaly. Using similar steps, the function G is expressed at first as

$$G = \sqrt{\frac{a^3}{\mu}} \left[\sin \hat{E} - e(\sin E - \sin E_0) \right] \tag{8.163}$$

To write the second part of the above expression in terms of $\hat{E}$ instead of E and E_0, Kepler's equation is written at time t and t_0 as

$$M = M_0 + \sqrt{\frac{\mu}{a^3}}(t - t_0) = E - e \sin E \tag{8.164}$$

$$M_0 = E_0 - e \sin E_0 \tag{8.165}$$

Subtracting one equation from the other leads to the expression

$$-e(\sin E - \sin E_0) = \sqrt{\frac{\mu}{a^3}}(t - t_0) - \hat{E} \tag{8.166}$$

which allows G to be written in the final form

$$G = (t - t_0) + \sqrt{\frac{a^3}{\mu}} \left(\sin \hat{E} - \hat{E} \right) \tag{8.167}$$

The next step is to find analytical expressions for $\dot{F}$ and $\dot{G}$. Using the $\dot{E}$ expression in Eq. (8.101), Eqs. (8.162) and (8.167) are differentiated to yield

$$\dot{F} = -\frac{\sqrt{\mu a}}{r r_0} \sin \hat{E} \tag{8.168}$$

$$\dot{G} = 1 - \frac{a}{r}(1 - \cos \hat{E}) \tag{8.169}$$

Since and F and G functions along with their derivatives are written in terms of $\hat{E}$, we would like to write the current orbit radius r, which appears in Eqs. (8.168) and (8.169), also in terms of $\hat{E}$. Substituting the trigonometric identity

$$\cos E = \cos E_0 \cos \hat{E} - \sin E_0 \sin \hat{E} \tag{8.170}$$

into Eq. (8.15) we find

$$r = a \left(1 - e \cos E_0 \cos \hat{E} + e \sin E_0 \sin \hat{E}\right) \tag{8.171}$$

Using Eqs. (8.15) and (8.122), $r = r(\hat{E})$ is written as

$$r = a + (r_0 - a) \cos \hat{E} + \sqrt{a}\sigma_0 \sin \hat{E} \tag{8.172}$$

To find a "modified" Kepler's equation in terms of $\hat{E}$, we recognize first that

$$\frac{d\hat{E}}{dt} = \dot{E} = \frac{1}{r}\sqrt{\frac{\mu}{a}} \tag{8.173}$$

After separating variables, this leads to

$$\sqrt{\frac{\mu}{a}} dt = r d\hat{E} \tag{8.174}$$

Substituting Eq. (8.172) and diving both sides by a leads to the perfect differential equation

$$\sqrt{\frac{\mu}{a^3}} dt = \left(1 - (1 - \frac{r_0}{a}) \cos \hat{E} + \frac{\sigma_0}{\sqrt{a}} \sin \hat{E}\right) d\hat{E} \tag{8.175}$$

Integrating this equation leads to the desired "modified" Kepler's equation

$$\hat{M} \equiv \sqrt{\frac{\mu}{a^3}}(t - t_0) = \hat{E} - (1 - \frac{r_0}{a}) \sin \hat{E} - \frac{\sigma_0}{\sqrt{a}}(\cos \hat{E} - 1) \tag{8.176}$$

The F and G solution provided in Eqs. (8.148) and (8.149) provide a direct mapping of the initial position and velocity vectors into corresponding vectors at the current time t. Note that the inverse transformation is provided by the remarkably simple expression:

$$\boldsymbol{r}(t_0) = \quad \dot{G}\boldsymbol{r}(t) - G\dot{\boldsymbol{r}}(t) \tag{8.177}$$

$$\dot{\boldsymbol{r}}(t_0) = -\dot{F}\boldsymbol{r}(t) + F\dot{\boldsymbol{r}}(t) \tag{8.178}$$

This inverse mapping is achieved using the F and G solution property

$$F\dot{G} - G\dot{F} = 0 \tag{8.179}$$

This property can be verified by back substituting the expressions found for F and G, along with their derivatives.

Another useful form of the F and G solution is to express the solution in terms of a true anomaly difference $\hat{f} = f - f_0$. To make use of the expressions for F and G in Eqs. (8.158) and (8.159), we write the position vector $\boldsymbol{r}(t)$ velocity vector $\dot{\boldsymbol{r}}(t)$ as

$$\boldsymbol{r} = \underbrace{r\cos f}_{x}\,\hat{\boldsymbol{i}}_e + \underbrace{r\sin f}_{y}\,\hat{\boldsymbol{i}}_p \tag{8.180}$$

$$\dot{\boldsymbol{r}} = \underbrace{-\frac{h\sin f}{p}}_{\dot{x}}\,\hat{\boldsymbol{i}}_e + \underbrace{\frac{h}{p}\left(e+\cos f\right)}_{\dot{y}}\,\hat{\boldsymbol{i}}_p \tag{8.181}$$

with the orbit radius rate $\dot{r}$ being given by

$$\dot{r} = \frac{he\sin f}{p} \tag{8.182}$$

Substituting Eq. (8.180) and (8.181) into Eqs. (8.158) and (8.159) and simplifying using trigonometric identities yields[2]

$$F = 1 - \frac{r}{p}\left(1-\cos\hat{f}\right) \tag{8.183}$$

$$G = \frac{rr_0}{h}\sin\hat{f} \tag{8.184}$$

Substituting Eqs. (8.180) and (8.181) into Eq. (8.122), the parameter σ is given by

$$\sigma = \frac{\boldsymbol{r}\cdot\dot{\boldsymbol{r}}}{\sqrt{\mu}} = \frac{hre\sin f}{\sqrt{\mu}p} = \frac{re\sin f}{\sqrt{p}} \tag{8.185}$$

Differentiating the F and G expressions in Eqs. (8.183) and (8.184) with respect to time and making use of the σ expression, the $\dot{F}$ and $\dot{G}$ are found to be

$$\dot{F} = -\frac{\sqrt{\mu}}{pr}\left(\sigma\left(1-\cos\hat{f}\right) + \sqrt{p}\sin\hat{f}\right) \tag{8.186}$$

$$\dot{G} = 1 - \frac{r_0}{p}\left(1-\cos\hat{f}\right) \tag{8.187}$$

The orbit radius is expressed in terms of $\hat{f}$ through the orbit equation

$$r = \frac{p}{1+e\cos(f_0+\hat{f})} \tag{8.188}$$

Expanding the cos term and making use of the σ definition and the fact that $r_0 = p/(1+e\cos f_0)$, the orbit radius can also be written in the form[2]

$$r = \frac{pr_0}{r_0 + (p-r_0)\cos\hat{f} - \sqrt{p}\sigma_0\sin\hat{f}} \tag{8.189}$$

Using Eq. (8.185) we can relative the σ parameters at time t_0 and t through

$$\frac{\sigma}{r} = \frac{\sigma_0}{r_0}\cos\hat{f} + \frac{\sin\hat{f}}{r_0\sqrt{p}}(p - r_0) \tag{8.190}$$

Substituting Eqs. (8.189) and (8.190) into Eq. (8.186), we are able to express the $\dot{F}$ rate expression in terms of initial states states only.[2]

$$\dot{F} = \frac{\mu}{r_0 p}\left(\sigma_0(1 - \cos\hat{f}) - \sqrt{p}\sin\hat{f}\right) \tag{8.191}$$

Problems

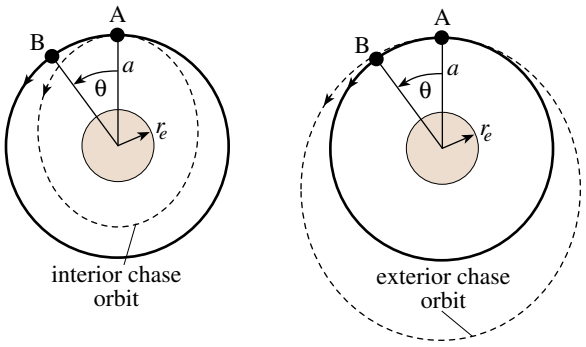
8.1 Integrate Kepler's second law in Example 8.1 directly without using the geometrical insight that the right hand side computes the area of an ellipse.

8.2 Starting with Eqs. (8.18) and (8.19), derive the partial derivative expression of f with respect to E given by

$$\frac{\partial f}{\partial E} = \frac{b}{r} \tag{8.192}$$

8.3 Specify the steps necessary to translate the six orbit elements in Eq. (8.110) given at time t_1 into the corresponding inertial position and velocity vectors $\boldsymbol{r}(t_1)$ and $\dot{\boldsymbol{r}}(t_1)$.

8.4 Verify that Eq. (8.131) is indeed the derivative of Eq. (8.130).



(i) Interior Chase Orbit **(ii)** Exterior Chase Orbit

Figure 8.12: Illustration of Two Rendezvous Chase Orbit Options.

8.5 Consider two spacecraft (A and B) in the same circular orbit of radius a. Spacecraft B is initially θ radians of true anomaly ahead of A. It is desired that the spacecraft A "catch up" (or rendezvous) with B by transferring temporarily onto a "chase" orbit, then transferring back onto the original circular orbit. Referring to Figure 8.12, two options are being considered:
Option 1: Use an Interior Orbit Spacecraft A decreases its velocity (by amount Δv_1), so that it transfers at apogee onto a judicious chase orbit. Upon return to apogee, it increases its velocity by Δv_1 to rendezvous with spacecraft B and maintain again a circular orbit of radius a.
Option 2: Use an Exterior Orbit Spacecraft A increases its velocity (by an amount Δv_2), so that it transfers at perigee onto a judicious chase orbit. Upon returning to perigee, it decreases its velocity by Δv_2 to rendezvous with spacecraft B.
Assume all velocity changes are instantaneous and tangential to the orbit. Assume rendezvous occurs, for both options, after one orbit of A on the chase orbit.

a) Determine the required velocity increments (Δv_1 and Δv_2) as functions of a, θ and μ

b) Discuss the relative advantages of these two options, and also discuss any circumstances which would make the solutions unfeasible.

8.6 Write subroutines that map between Cartesian orbit position coordinates and the orbit element set shown in Eq. (8.110) using the following steps.

a) Write a subroutine that maps the Cartesian position and velocity coordinates of a spacecraft to the corresponding orbit elements and current mean anomaly.

b) Write a subroutine that maps current orbit elements and time since last perigee passage to corresponding Cartesian position and velocity coordinates. Verify that this numerical mapping is the precise inverse of task a).

c) Write a numerical simulation that integrates the differential equations of motion in Eq. (8.45) using a 4-th order Runge Kutta integration scheme. Using the subroutine of task b), compare the answer of the numerical integration to the analytical two-body solution.

8.7 Program the F and G solution to the two-body problem. Verify the answer by comparing it to a numerical integration of the differential equations of motion in Eq. (8.45).

8.8 Consider the two-body equations of motion in Eq. (8.45) and the F and G solution in Eqs. (8.148) and (8.149).

a) Prove that F and G satisfy $\ddot{F} = -\mu/r^3 F$ and $\ddot{G} = -\mu/r^3 G$.

b) Prove that $H = F\dot{G} - G\dot{F}$ is a constant of the solution. Evaluate H as a function of the initial conditions.

Bibliography

[1] Greenwood, D. T., *Principles of Dynamics*, Prentice-Hall, Inc, Englewood Cliffs, New Jersey, 2nd ed., 1988.

[2] Battin, R. H., *An Introduction to the Mathematics and Methods of Astrodynamics*, AIAA Education Series, New York, 1987.

[3] Nelson, W. C. and Loft, E. E., *Space Mechanics*, Prentice Hall, Englewood Cliffs, New Jersey, 1962.

[4] Junkins, J. L., "A Closed Form Method for Obtaining Projectile Impact Range Deviation Due to Errors in Magnitude and Orientation of the Initial Velocity Vector," Tech. rep., Missle & Space Systems Division, Douglas Aircraft Company, Inc., 1966.

[5] Brouwer, D., "Solution of the Problem of Artificial Satellite Theory Without Drag," *The Astronautical Journal*, Vol. 64, No. 1274, 1959, pp. 378–397.

Chapter Nine

Restricted Three-Body Problem

WHILE the Keplerian two-body problem has a very elegant analytical solution, the general three-body problem increases the level of complexity to such a degree to make an analytical solution intractable. While no general solutions exist for this 18-th order system, ten exact analytical integrals are possible for the general case, corresponding to conservation of angular momentum (3 integrals), energy (1 integral), and motion of the system mass center (6 integrals). These ten integrals, together with imposing other special case conditions, permit considerable additional analytical progress to be made. Virtually all of this progress stems from the work of the brilliant French astronomer and mathematician, Lagrange. A familiar three-body problem is the Sun-Earth-Moon system. While the moon does orbit the Earth in a near-elliptical manner, to account for some of the deviations of its orbit relative to Earth, the gravitational effect of the sun must also be taken into account. This is one reason why a precise description of the lunar orbit is very complicated.

While Newton was the first to study the combined motion of several celestial objects accounting for their mutual gravitational attraction, it was Lagrange in 1772 who submitted his memoir *Essai sur le Problème des Trois Corps* to the Paris Academy that demonstrated analytical solutions do exist for the three-body problem if certain restrictions are imposed. These restrictions force the three bodies to remain in an equilateral triangle or collinear formation. This chapter studies these motions and shows their properties. Of particular interest is the circular restricted three-body problem. Here two larger, spherical bodies are restricted to follow a circular, Keplerian motion, while a third body of relatively infinitesimal mass is moving among them in a general fashion. A good way to visualize this is to think of the Apollo program where a small space vehicle is flying under the gravitational influence of the uniformly rotating Earth-Moon system.

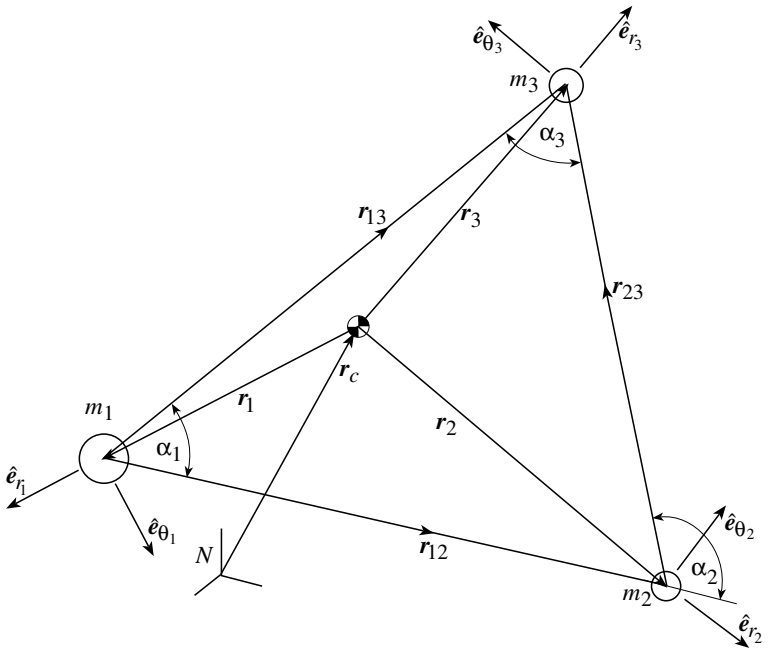


Figure 9.1: Illustration of Three-Body Problem

9.1 Lagrange's Three-Body Solution

Lagrange showed that it is possible to find solutions of the three-body problem where the shape of the three-body formation does not change in time. First, the more general case is studied where the size or orientation of the fixed three-body formation is free to vary with time. Then, the special case is studied where all three bodies are assumed to rotate about their center of mass at a common angular rate. In both studies the mass of each body is assumed to be sufficiently large to affect the motion of the remaining two bodies.

9.1.1 General Conic Solutions

Assume the position of the three masses m_1, m_2 and m_3 are located relative to their center of mass through the vectors $\boldsymbol{r}_1$, $\boldsymbol{r}_2$ and $\boldsymbol{r}_3$ as shown in Figure 9.1. Since external forces are neglected, we find

$$M\ddot{\boldsymbol{r}}_c = \boldsymbol{F}_{external} \approx 0 \tag{9.1}$$

where the total system mass M is defined as

$$M = m_1 + m_2 + m_3 \tag{9.2}$$

and $\boldsymbol{r}_c$ is the inertial position vector of the system center of mass. We see that the motion of the mass center is a constant velocity straight line motion.

Since $\ddot{\boldsymbol{r}}_c = 0$, we adopt the center of mass as an *inertial* origin, then the three equations of motion are given by

$$m_i \ddot{\boldsymbol{r}}_i = G \sum_{j=1}^{3} \frac{m_i m_j}{r_{ij}^3} \boldsymbol{r}_{ij} = \boldsymbol{F}_i \qquad \text{for } i = 1, 2, 3 \quad i \neq j \tag{9.3}$$

where G is the universal gravitational constant, $\boldsymbol{F}_i$ is the net resultant force acting on each mass m_i, $\boldsymbol{r}_{ij}$ is the relative position vector defined as

$$\boldsymbol{r}_{ij} = \boldsymbol{r}_j - \boldsymbol{r}_i \tag{9.4}$$

and the scalar distances r_{ij} are computed as

$$r_{ij} = r_{ji} = \sqrt{\boldsymbol{r}_{ij} \cdot \boldsymbol{r}_{ij}} \tag{9.5}$$

Since the position vectors $\boldsymbol{r}_i$ are defined relative to the center of mass, then according to Eq. (2.43)

$$m_1 \boldsymbol{r}_1 + m_2 \boldsymbol{r}_2 + m_3 \boldsymbol{r}_3 = 0 \tag{9.6}$$

must be true. Using Eq. (9.2) and adding and subtracting appropriate terms, Eq. (9.6) is written in three different forms:

$$M \boldsymbol{r}_1 = -m_2 \boldsymbol{r}_{12} - m_3 \boldsymbol{r}_{13} \tag{9.7a}$$

$$M \boldsymbol{r}_2 = m_1 \boldsymbol{r}_{12} - m_3 \boldsymbol{r}_{23} \tag{9.7b}$$

$$M \boldsymbol{r}_3 = m_1 \boldsymbol{r}_{13} + m_2 \boldsymbol{r}_{23} \tag{9.7c}$$

Squaring these equations we find the following useful *scalar* relationships

$$M^2 r_1^2 = m_2^2 r_{12}^2 + m_3^2 r_{13}^2 + 2 m_2 m_3 \boldsymbol{r}_{12} \cdot \boldsymbol{r}_{13} \tag{9.8a}$$

$$M^2 r_2^2 = m_1^2 r_{12}^2 + m_3^2 r_{23}^2 - 2 m_1 m_3 \boldsymbol{r}_{12} \cdot \boldsymbol{r}_{23} \tag{9.8b}$$

$$M^2 r_3^2 = m_1^2 r_{13}^2 + m_2^2 r_{23}^2 + 2 m_1 m_2 \boldsymbol{r}_{13} \cdot \boldsymbol{r}_{23} \tag{9.8c}$$

Following Lagrange's historic conjectures, the key assumption in this development is that we are seeking solutions to the three-body problem which retain the shape of the original three-body configuration. For this to be true, the three relative distances r_{12}, r_{23} and r_{13} must all evolve in the same manner. Let $f(t)$ be some generic time varying function with $f(0) = 1$. Then the relative distances r_{ij} must satisfy

$$\frac{r_{12}}{r_{12_0}} = \frac{r_{13}}{r_{13_0}} = \frac{r_{23}}{r_{23_0}} = f(t) \tag{9.9}$$

where r_{ij_0} is the initial relative distance. Since the formation shape is fixed, the angles α_i between the relative distance vectors are constant. Substituting Eq. (9.9) into (9.8a) we find

$$M^2 r_1^2 = f(t)^2 \left(m_2^2 r_{12_0}^2 + m_3^2 r_{13_0}^2 + 2 m_2 m_3 r_{12_0} r_{13_0} \cos \alpha_1 \right) \tag{9.10}$$

Since $f(0) = 1$, we are able to express r_{1_0} as

$$r_{1_0} = \frac{\sqrt{m_2^2 r_{12_0}^2 + m_3^2 r_{13_0}^2 + 2m_2 m_3 r_{12_0} r_{13_0} \cos\alpha_1}}{M} \tag{9.11}$$

Eq. (9.11) is alternatively derived immediately from Eq. (9.8a) by applying initial conditions. Using this definition in Eq. (9.10), we are able to express $r_1(t)$ as

$$r_1(t) = r_{1_0} f(t) \tag{9.12}$$

This states that the radial distance of the mass to the center of mass will evolve in the same manner as the relative distances. Similarly, we can express the other two radial distances as

$$r_2(t) = r_{2_0} f(t) \tag{9.13}$$

$$r_3(t) = r_{3_0} f(t) \tag{9.14}$$

Since the three-body configuration shape must remain invariant, all three angular velocity vectors must be equal (but not necessarily constant):

$$\boldsymbol{\omega}_1 = \boldsymbol{\omega}_2 = \boldsymbol{\omega}_3 = \boldsymbol{\omega} = \omega \hat{\boldsymbol{e}}_3 \tag{9.15}$$

The angular momentum vector $\boldsymbol{H}$ of the general three-body system about the center of mass is constant for this zero-external force system and is defined as

$$\boldsymbol{H} = \sum_{i=1}^{3} \boldsymbol{r}_i \times m_i \dot{\boldsymbol{r}}_i = constant \tag{9.16}$$

It is convenient to simplify the discussion by introducing a fundamental invariant plane whose normal is the constant angular momentum vector $\boldsymbol{H}$. If we further restrict attention to the case that all three sets of position and velocity vectors lie in this invariant plane at some initial instant, then the motion of (m_1, m_2, m_2) will remain co-planar forever, because all of the forces can be shown to lie in this same plane. Therefore we are able to treat this shape invariant three-body problem as a planar motion problem. Expressing the position vectors $\boldsymbol{r}_i$ and velocity vectors $\dot{\boldsymbol{r}}_i$ components in rotating reference frames $\mathcal{E}_i = \{\hat{\boldsymbol{e}}_{r_i}, \hat{\boldsymbol{e}}_{\theta_i}, \hat{\boldsymbol{e}}_{3_i}\}$ shown in Figure 9.1, we find

$$\boldsymbol{r}_i = r_i \hat{\boldsymbol{e}}_{r_i} \tag{9.17}$$

$$\dot{\boldsymbol{r}}_i = \dot{r}_i \hat{\boldsymbol{e}}_{r_i} + r_i \omega \hat{\boldsymbol{e}}_{\theta_i} \tag{9.18}$$

$$\ddot{\boldsymbol{r}}_i = (\ddot{r}_i - r_i \omega^2) \hat{\boldsymbol{e}}_{r_i} + (2\dot{r}_i \omega + r_i \dot{\omega}) \hat{\boldsymbol{e}}_{\theta_i} \tag{9.19}$$

Using Eqs. (9.12) through (9.15), we write the angular momentum vector as

$$\boldsymbol{H} = \sum_{i=1}^{3} \left(m_i r_{i_0}^2\right) f^2 \omega \hat{\boldsymbol{e}}_3 \tag{9.20}$$

Because $\boldsymbol{H}$ is constant, Eq. (9.20) implies that the product $f^2(t)\omega(t)$ must also be constant. Since the angular momentum vector of each particle is given by

$$\boldsymbol{H}_i = \boldsymbol{r}_i \times m_i \dot{\boldsymbol{r}}_i = m_i r_{i_0}^2 f^2 \omega \hat{\boldsymbol{e}}_3 \tag{9.21}$$

this implies that the angular momentum of each mass m_i is also constant. Taking the derivative of the constant vector $\boldsymbol{H}_i$ we find

$$\dot{\boldsymbol{H}}_i = \boldsymbol{r}_i \times m_i \ddot{\boldsymbol{r}}_i = \boldsymbol{r}_i \times \boldsymbol{F}_i = 0 \tag{9.22}$$

The resulting condition $\boldsymbol{r}_i \times \ddot{\boldsymbol{r}}_i$ dictates that the acceleration vector $\ddot{\boldsymbol{r}}_i$, and therefore also the i-th net force vector $\boldsymbol{F}_i$, must be parallel at all times to the radial position vector $\boldsymbol{r}_i$. The conclusion is one condition for the three-body configuration shape to remain fixed is that the resultant force vector $\boldsymbol{F}_i$ on each mass must be a radial force passing through the system center of mass, and can be written as

$$\boldsymbol{F}_i = F_i \hat{\boldsymbol{e}}_{r_i} \tag{9.23}$$

Substituting Eq. (9.19) and (9.23) into the equations motion in Eq. (9.3) we find

$$F_i = m_i(\ddot{r}_i - r_i \omega^2) \tag{9.24}$$

Using Eqs. (9.12)-(9.14) this is rewritten as

$$\frac{F_i}{m_i} = \ddot{f} r_{i_0} - r_i \omega^2 = r_i \left(\frac{\ddot{f}}{f} - \omega^2 \right) \tag{9.25}$$

Rearranging Eq. (9.25) we find

$$\frac{F_i}{m_i r_i} = \frac{\ddot{f}}{f} - \omega^2 = A(t) \tag{9.26}$$

which states that the ratio of the net resultant force over the radial distance to the center of mass and body mass m_i will remain the same for all three masses. Therefore

$$F_i(t) = A(t) r_i(t) m_i \tag{9.27}$$

The condition in Eq. (9.22) requires that $\boldsymbol{r}_i \times \ddot{\boldsymbol{r}}_i = 0$. As we will show next, this is only possible for our restricted three-body system for two types of configurations. Setting $i = 1$ in Eq. (9.3) and taking its cross product with $\boldsymbol{r}_1$ leads to

$$\boldsymbol{r}_1 \times \left(m_2 \frac{\boldsymbol{r}_2}{r_{12}^3} + m_3 \frac{\boldsymbol{r}_3}{r_{13}^3} \right) = 0 \tag{9.28}$$

The center of mass definition in Eq. (9.6) is rewritten as

$$m_3 \boldsymbol{r}_3 = -m_1 \boldsymbol{r}_1 - m_2 \boldsymbol{r}_2 \tag{9.29}$$

Substituting Eq. (9.29) into (9.28) leads to the condition

$$m_2 \boldsymbol{r}_1 \times \boldsymbol{r}_2 \left(\frac{1}{r_{12}^3} - \frac{1}{r_{13}^3} \right) = 0 \tag{9.30}$$

Similarly, for the other two cases we find the necessary conditions

$$m_1 \boldsymbol{r}_2 \times \boldsymbol{r}_3 \left(\frac{1}{r_{23}^3} - \frac{1}{r_{12}^3} \right) = 0 \tag{9.31}$$

$$m_3 \boldsymbol{r}_3 \times \boldsymbol{r}_1 \left(\frac{1}{r_{13}^3} - \frac{1}{r_{23}^3} \right) = 0 \tag{9.32}$$

There are only two geometric configurations that satisfy Eqs. (9.30)-(9.32). The first configuration found by Lagrange is that of an equilateral triangle since

$$r_{12} = r_{23} = r_{13} = \rho \tag{9.33}$$

The second possible configuration has all three bodies on a straight line in a collinear formation:

$$\boldsymbol{r}_1 \times \boldsymbol{r}_2 = \boldsymbol{r}_2 \times \boldsymbol{r}_3 = \boldsymbol{r}_3 \times \boldsymbol{r}_1 = 0 \tag{9.34}$$

It is remarkable to note that these are the *only* two possible three-body configurations which will maintain constant formation shapes. The necessary initial conditions for the three-body motion to be shape-invariant are summarized as:[1]

1. The net resultant force $\boldsymbol{F}_i$ on each mass must pass through the system center of mass.

2. The net resultant force $\boldsymbol{F}_i$ is along the radial vector locating each mass relative to the system center of mass.

3. The initial velocity vectors are proportional in magnitude to the respective distances of the masses to the system center of mass.

4. The initial velocity vectors make equal angles with the radial position vectors to the system center of mass

Equilateral Triangle Solution

Again, following the insights of Lagrange, we investigate the special case in which the masses lie at the vertices of a rotating equilateral triangle. For this case, the equations of motion of each individual mass take on a surprisingly simple and familiar form. We note for the most general class of equilateral motions, the triangle is rotating at some variable angular velocity (to be determined) and

the size of the equilateral triangle may be time varying. Substituting Eq. (9.33) into the three-body equations of motion in Eq. (9.3) we find the general form

$$m_1\ddot{\boldsymbol{r}}_1 = \frac{Gm_1}{\rho^3}(m_2\boldsymbol{r}_{12} + m_3\boldsymbol{r}_{13}) \tag{9.35a}$$

$$m_2\ddot{\boldsymbol{r}}_2 = \frac{Gm_2}{\rho^3}(-m_1\boldsymbol{r}_{12} + m_3\boldsymbol{r}_{23}) \tag{9.35b}$$

$$m_3\ddot{\boldsymbol{r}}_3 = -\frac{Gm_3}{\rho^3}(m_1\boldsymbol{r}_{13} + m_2\boldsymbol{r}_{23}) \tag{9.35c}$$

Substituting the center of mass conditions in Eq. (9.7), these equations of motion are written compactly as

$$\ddot{\boldsymbol{r}}_i + \frac{GM}{\rho^3}\boldsymbol{r}_i = 0 \qquad \text{for } i = 1, 2, 3 \tag{9.36}$$

Specializing these general equations in Eq. (9.35), we make use of the fact that for the equilateral triangle special class of motions $\alpha_1 = \alpha_3 = 60^\circ$ and $\alpha_2 = 120^\circ$ and substitute Eqs. (9.8), the three equations of motion can be written in the *decoupled* form as

$$\ddot{\boldsymbol{r}}_i + \frac{GM_i}{r_i^3}\boldsymbol{r}_i = \ddot{\boldsymbol{r}}_i + \frac{\mu_i}{r_i^3}\boldsymbol{r}_i = 0 \qquad \text{for } i = 1, 2, 3 \tag{9.37}$$

with the equivalent effective masses M_i defined as

$$M_1 = \frac{1}{M^2}(m_2^2 + m_3^2 + m_2m_3)^{3/2} \tag{9.38}$$

$$M_2 = \frac{1}{M^2}(m_1^2 + m_3^2 + m_1m_3)^{3/2} \tag{9.39}$$

$$M_3 = \frac{1}{M^2}(m_1^2 + m_2^2 + m_1m_2)^{3/2} \tag{9.40}$$

and $\mu_i = GM_i$. Note that the equations of motion in Eq. (9.37) are of the identical form as the relative, two-body equations of motion derived in Eq. (8.45). This implies that for the equilateral triangle three-body solution, each mass body behaves as if it were only attracted by a mass M_i placed at the center of mass of the system. Whether these orbits are elliptical, parabolic or hyperbolic depends on the energy of the system.

Example 9.1: To illustrate the general equilateral solution of the three-body problem, the motion of the following three-body system is numerically solved using Eq. (9.37). The masses are $m_1 = 5.967\ 10^{23}$kg (1/10th of Earth's mass), $m_2 = 7.35\ 10^{22}$kg (Moon's mass) and $m_3 = 3.675\ 10^{22}$kg (half of Moon's mass). Each side of the equilateral triangle has an initial length of 10^9m. The initial velocity vector of each mass forms a 40° angle with the respective radial position vector and has the magnitudes $|\dot{\boldsymbol{r}}_1| = 29.8659$m/s, $|\dot{\boldsymbol{r}}_2| = 189.181$m/s and $|\dot{\boldsymbol{r}}_3| = 195.552$m/s. The resulting motion is shown in Figure 9.2 has seen by a non-rotating frame.

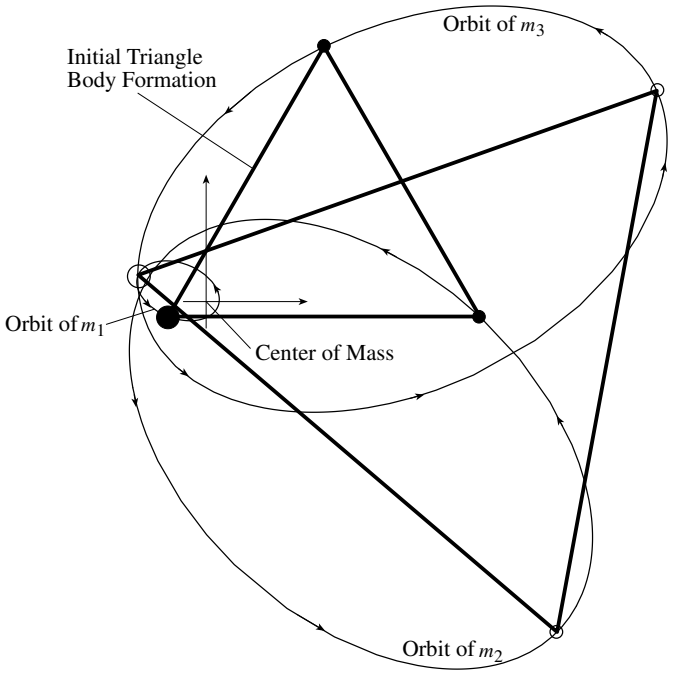


Figure 9.2: Illustration of General Equilateral Triangle Solution of the Three-Body Problem

The triangular configuration is highlighted at the initial time and at another time during the motion. Clearly the shape of the equilateral triangle is invariant, while its size and orientation changes with time. With the given initial energy, each mass follows an elliptic orbit with the system center of mass located on one of its foci.

Collinear Solution

The second invariant shape of the three-body problem is that of a straight line. Here the ratio of the distances between the bodies remains constant. Assuming the three bodies are aligned along a rotating straight line, then the rotating $\hat{\boldsymbol{e}}_{r_i}$ vectors will be collinear. The three position vectors $\boldsymbol{r}_i$ are now given by

$$\boldsymbol{r}_i = x_i \hat{\boldsymbol{e}}_r \qquad \text{for } i = 1, 2, 3 \tag{9.41}$$

Substituting these specific position vectors into the equations of motion, Eq. (9.3) and making use of Eqs. (9.23) and (9.27), the scalar force components F_i are expressed as

$$F_1 = Ax_1 m_1 = m_1 m_2 \frac{x_2 - x_1}{x_{12}^3} + m_1 m_3 \frac{x_3 - x_1}{x_{13}^3} \tag{9.42}$$

$$F_2 = Ax_2 m_2 = m_2 m_3 \frac{x_3 - x_2}{x_{23}^3} - m_1 m_2 \frac{x_2 - x_1}{x_{12}^3} \tag{9.43}$$

$$F_3 = Ax_3 m_3 = -m_1 m_3 \frac{x_3 - x_1}{x_{13}^3} - m_2 m_3 \frac{x_3 - x_2}{x_{23}^3} \tag{9.44}$$

where A was previously found to be a scalar quantity common to all three bodies. Note that since $r_i(t) = r_{i_0} f(t)$, then

$$F_i = \frac{constant}{f^2} \tag{9.45}$$

Since f is proportional to the radial distance to the center of mass, each mass is subject to an inverse-square-law attraction and therefore describes elliptic, parabolic or hyperbolic trajectories.[1] These three equations in (9.42) through (9.44) must be solved for the relative distances x_{12}, x_{13} and x_{23}. The order of the particles within the line is arbitrary and three configurations 123, 132 and 312 are possible. We will solve for the relative distances for the first case which is illustrated in Figure 9.3. The second and third case solutions can be found by appropriately rearranging the indices, since the choice of these labels is obviously arbitrary.

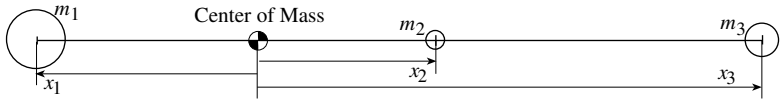


Figure 9.3: Collinear Three-Body Sequence Illustration

To simplify solving for the relative distances, the scalar quantity χ is introduced as

$$\chi = \frac{x_3 - x_2}{x_2 - x_1} = \frac{x_{23}}{x_{12}} \tag{9.46}$$

Note that

$$\frac{x_{13}}{x_{12}} = 1 + \chi \tag{9.47}$$

Observe, if any of the distances (x_{12}, x_{13}, x_{23}) are known, then finding χ determines the system configuration. Subtracting Eq. (9.43) from (9.42) and Eq. (9.44) from (9.43) yields

$$Ax_{12} = -\frac{m_1 + m_2}{x_{12}^2} + m_3 \left(\frac{1}{x_{23}^2} - \frac{1}{x_{13}^2} \right) \tag{9.48}$$

$$Ax_{23} = -\frac{m_2 + m_3}{x_{23}^2} + m_1 \left(\frac{1}{x_{12}^2} - \frac{1}{x_{13}^2} \right) \tag{9.49}$$

Substituting the definition of χ we are able to rearrange these two equations into the forms

$$Ax_{12}^3 = -(m_1 + m_2) + m_3 \left(\frac{1}{\chi^2} - \frac{1}{(1+\chi)^2} \right) \tag{9.50}$$

$$Ax_{12}^3 = -(m_2 + m_3)\frac{1}{\chi^3} + m_1 \left(\frac{1}{\chi} - \frac{1}{\chi(1+\chi)^2} \right) \tag{9.51}$$

Equating Eqs. (9.50) and (9.51), Lagrange's famous quintic equation for the admissible configuration is found:

$$\begin{aligned}(m_1 + m_2)\chi^5 + (3m_1 + 2m_2)\chi^4 + (3m_1 + m_2)\chi^3 \\ - (m_2 + 3m_3)\chi^2 - (2m_2 + 3m_3)\chi - (m_2 + m_3) = 0\end{aligned} \tag{9.52}$$

Since the polynomial coefficients only change sign once, so there is only one positive real root to this fifth-order polynomial equation. Therefore Eq. (9.52) uniquely defines the relative distances of the three bodies. Given χ and one of the relative distances between two bodies, the remaining two relative distances can be computed using Eqs. (9.46) and (9.47).

Example 9.2: To illustrate the general invariant collinear three-body solution, the three-body system from Example 9.1 is used with a different initial configuration. With the specified masses, solving Lagrange's quintic equation for scalar parameter χ we find

$$\chi = 0.451027$$

The scalar distance x_{12} is chosen to be 10^9m. Using Eqs. (9.46) and (9.47), the other two relative distances are

$$x_{23} = 4.510273\ 10^8\text{m}$$

$$x_{13} = 1.451027\ 10^9\text{m}$$

Each velocity vector initially forms a 40° angle with the respective radial position vector from the system center of mass. The initial speeds are $|\dot{\boldsymbol{r}}_1| = 32.9901$m/s, $|\dot{\boldsymbol{r}}_2| = 150.903$m/s and $|\dot{\boldsymbol{r}}_3| = 233.844$m/s. The resulting motion is shown in Figure 9.4 as seen by a non-rotating reference frame.

The collinear three-body configuration is highlighted at three distinct times. As predicted, the ratios of the relative distances remain the same, while the size and orientation of the configuration varies with time. With the given initial energy, all three orbits are elliptical. Again the system center of mass lies in the common foci of each ellipse. The geometric size of each orbit in these three-body problems is always an indication of the mass of that particular object. In this setting, the first mass m_1 is the most massive object. Therefore the center of mass point is located close to it and its trajectory describes the smallest ellipse about this point. On the other end of the spectrum, the mass m_3 is the lightest of the three and is therefore the furthest removed from the center of mass and with the largest elliptic orbit.

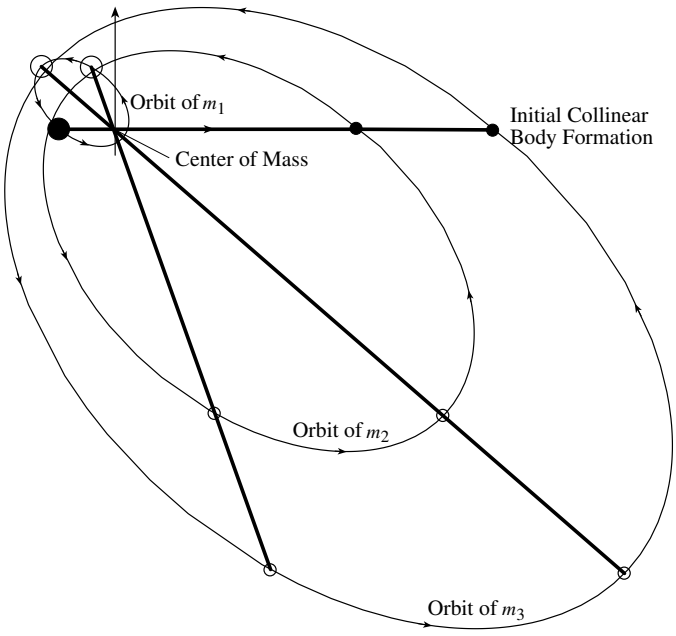


Figure 9.4: Illustration of General Invariant Collinear Solution of the Three-Body Problem

9.1.2 Circular Orbits

Instead of allowing the three orbits to be either elliptical, parabolic or hyperbolic, we now constrain them to be circular. For the three-body configuration shape to remain invariant under this condition, all three orbits must have the same constant angular velocity vector $\boldsymbol{\omega}$. The three bodies will now move in coplanar orbits with the orbit center being the system center of mass. Since this is a special case of the general conic orbit solutions, the invariant three-body configuration shapes will again be the equilateral triangle and collinear formations. Because the orbit radius r_i is now fixed, Eq. (9.24) shows that the scalar gravitational force F_i must equal the mass times the centrifugal acceleration.

$$F_i = -m_i r_i \omega^2 \tag{9.53}$$

We now derive specific conditions that must be met for such shape-invariant, circular three-body orbits to exist. Note that since the three-body formation shape will not grow with circular orbits, the masses m_i *will appear to remain fixed* when viewed from a frame rotating with an angular velocity $\boldsymbol{\omega}$. These specific mass locations are also referred to as *stationary points* of Lagrange's restricted three-body problem. As is seen later, examining these points is very useful when studying spacecraft orbits in the vicinity of two massive celestial bodies.

The three differential equations in Eq. (9.3) define the motion of any three-

body system. We are able to eliminate one second order differential equation by making use of the center of mass condition in Eq. (9.6). Substituting Eq. (9.53) into Eq. (9.3) we find the following three algebraic equations which must hold when all three orbits are circular.

$$\underbrace{\begin{bmatrix} \left(\frac{\omega^2}{G} - \frac{m_2}{r_{12}^3} - \frac{m_3}{r_{13}^3}\right) & \frac{m_2}{r_{12}^3} & \frac{m_3}{r_{13}^3} \\ \frac{m_1}{r_{12}^3} & \left(\frac{\omega^2}{G} - \frac{m_1}{r_{12}^3} - \frac{m_3}{r_{23}^3}\right) & \frac{m_3}{r_{23}^3} \\ m_1 & m_2 & m_3 \end{bmatrix}}_{[B]} \begin{bmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \\ \boldsymbol{r}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \tag{9.54}$$

For non-trivial solutions to exist, the determinant of $[B]$ must be equal to zero. Setting $r_{12} = r_{13} = r_{23} = \rho$ for the equilateral triangle case, the determinant of $[B]$ is found to be of the remarkably simple form

$$\det([B]) = m_3 \left(\frac{\omega^2}{G} - \frac{m}{\rho^3}\right) \tag{9.55}$$

Setting this determinant equal to zero, the following necessary condition is found for the equilateral triangle shaped three-body configuration to maintain circular orbits about the system mass center:

$$\rho^3 \omega^2 = GM \tag{9.56}$$

This equation is a close cousin to Kepler's third law of planetary motion. In fact, if we set m_3 equal to zero, then we obtain the same circular orbit angular velocity as is found studying Keplerian motion. Computing the orbit Period P for any of our three bodies we find

$$P = 2\pi \sqrt{\frac{\rho^3}{GM}} \tag{9.57}$$

By substituting Eq. (9.53) into (9.36) the same circular orbit condition is obtained from the more general conic formulation. This illustrates that the current development is a special case of the more general conic development, and that both are closely related to the corresponding Keplerian motion results.

Example 9.3: The numerical simulation in Example 9.1 is repeated with almost identical initial conditions. The only difference is that the angle of the initial velocity vector to the radial position vectors to the center of mass is now uniformly 90°. The velocities chosen in Example 9.1 were such that the necessary condition in Eq. (9.56) is satisfied. By having all initial velocity vectors be normal to the position vectors, circular orbits are achieved instead of elliptical orbits. The resulting motion is shown in Figure 9.5 as seen by a non-rotating coordinate frame.

With these initial conditions, the equilateral triangle shape of the three-body configuration rotates with a constant angular rate ω, but remains fixed in size. If the motion were shown as seen by a frame rotating with an angular rate

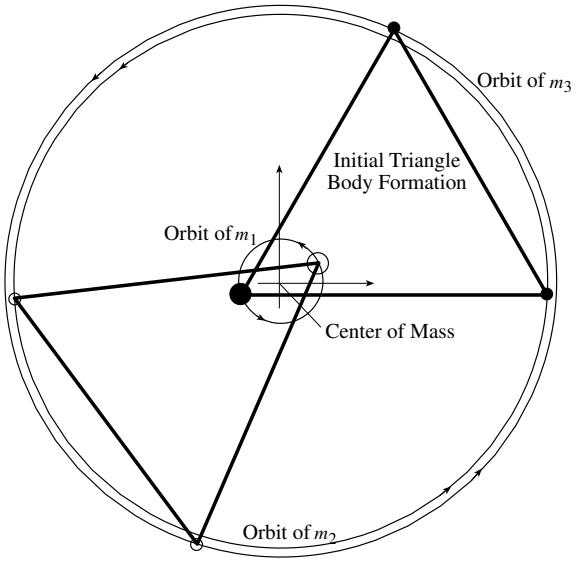


Figure 9.5: Illustration of Equilateral Triangle Solution of the Three-Body Problem with Circular Orbits

ω, the equilateral triangle would appear fixed in both size and orientation. Studying the three-body dynamics as seen in the rotating frame can be quite useful and is done extensively in the next section.

To find the collinear solution for the circular orbit case, we write the three position vectors as

$$\boldsymbol{r}_1 = x_1 \hat{\boldsymbol{e}}_r \tag{9.58a}$$

$$\boldsymbol{r}_2 = (x_1 + x_{12}) \hat{\boldsymbol{e}}_r \tag{9.58b}$$

$$\boldsymbol{r}_3 = (x_1 + x_{12} + x_{23}) \hat{\boldsymbol{e}}_r \tag{9.58c}$$

where the masses are aligned as was done in the general conic orbit development in Figure 9.3. Note that the scalar position coordinate x_1 of the first mass relative to the system center of mass is a negative quantity in this setting. In the following development, we seek expressions to solve for the geometric quantities x_1 and x_{23} and the kinetic quantity ω^2, while assuming that the relative distance x_{12} between the first and second mass is given. In essence, given a two-body problem consisting of m_1 and m_2, we are seeking to determine where we need to place the third mass relative to the second mass such that the new three-body system will maintain circular orbits about the center of mass with an invariant collinear mass configuration. Instead of placing the third mass, we could just as easily have chosen to place the first or second mass. Substituting the definitions

in Eq. (9.58) into Eq. (9.54), the following three *scalar* equations are found:

$$\frac{\omega^2}{G}x_1 + \frac{m_2}{x_{12}^2} + \frac{m_3}{(x_{12}+x_{23})^2} = 0 \tag{9.59a}$$

$$\frac{\omega^2}{G}(x_{12}+x_1) - \frac{m_1}{x_{12}^2} + \frac{m_3}{x_{23}^2} = 0 \tag{9.59b}$$

$$Mx_1 + m_2x_{12} + m_3(x_{12}+x_{23}) = 0 \tag{9.59c}$$

As is the case in solving the general collinear solution, it is convenient to rewrite these equations in terms of the scalar quantity χ defined in Eq. (9.46). Making use of this definition, the three scalar equations are written as

$$\frac{\omega^2}{G}x_{12}^2x_1 + m_2 + \frac{m_3}{(1+\chi)^2} = 0 \tag{9.60a}$$

$$\frac{\omega^2}{G}x_{12}^2(x_{12}+x_1) - m_1 + \frac{m_3}{\chi^2} = 0 \tag{9.60b}$$

$$\frac{M}{x_{12}}x_1 + m_2 + (1+\chi)m_3 = 0 \tag{9.60c}$$

Recall that the relative distance x_{12} is assumed to be a fixed parameter. Therefore Eq. (9.60a) can be solved directly for the necessary angular velocity magnitude ω in terms of χ.

$$\omega^2 = \frac{GM}{x_{12}^3(1+\chi)^2}\frac{m_2(1+\chi)^2 + m_3}{m_2 + (1+\chi)m_3} \tag{9.61}$$

Note that if m_3 is set to zero, then the standard two-body circular orbit speed condition is retrieved. The orbit Period P for this configuration is given by

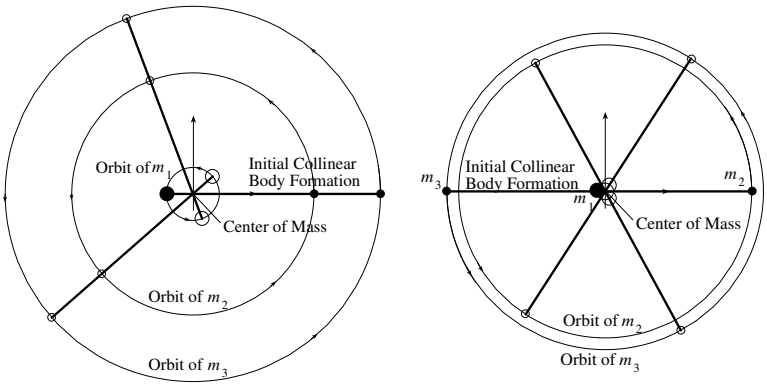
$$P = 2\pi\sqrt{\frac{x_{12}^3(1+\chi)^2}{GM}\frac{m_2 + (1+\chi)m_3}{m_2(1+\chi)^2 + m_3}} \tag{9.62}$$

Eq. (9.60a) provides an expression for the radial distance x_1 of the mass m_1 relative to the system center of mass in terms of χ.

$$x_1 = -\frac{x_{12}}{M}(m_2 + (1+\chi)m_3) \tag{9.63}$$

Substituting Eqs. (9.61) and (9.63) into Eq. (9.60b) leads us back to Lagrange's quintic polynomial equation provided Eq. (9.52). Given the three masses m_1, m_2 and m_3, this polynomial is solved numerically for its one real root. Given χ, we are then able to compute the remaining x_1 and x_{23} quantities.

Example 9.4: The numerical simulation in Example 9.2 is repeated with slightly different initial conditions. The angle of the initial velocity vector relative to the radial position vectors from the center of mass is now uniformly $90°$. The velocities chosen in Example 9.2 were such that the necessary



(i) m_3 On the Far Side of m_2 **(ii)** m_3 On the Far Side of m_1

Figure 9.6: Illustration of Collinear Solutions of the Three-Body Problem with Circular Orbits

condition in Eq. (9.61) is satisfied. The resulting motion is shown Figure 9.6 as seen by a non-rotating coordinate frame.

Clearly the initially collinear three-body formation remains collinear with these initial conditions. As seen by a reference frame rotating with an angular velocity ω, all three masses would appear to remain stationary along a straight line. Figure 9.6(ii) shows a similar configuration where m_3 is placed on the far side of m_1. Since m_1 is much larger than both m_2 and m_3, its orbits is a tight circle about the center of mass point.

9.2 Circular Restricted Three-Body Problem

In the circular restricted three-body problem we assume that both m_1 and m_2 are very massive objects compared to the third mass m_3. In this restricted problem, the Keplerian motion of the first two masses is determined through their respective inverse-square gravitational attraction by neglecting the effect of the relatively small third mass on the first two masses. From here on, we will drop the letter "3" subscript on the small mass and simply call it m. Therefore the masses m_1 and m_2 affect the motion of m, without in return being affected by m themselves. The bodies m_1 and m_2 are assumed be in circular orbits about their mutual center of mass. This is a good approximation for several celestial couples like Earth-Moon, Sun-Earth, Sun-Jupiter, ... The small mass m could then be the Apollo spacecraft flying between Earth and Moon or some asteroids moving under the influence of the sun and some planet.

For bodies in circular orbits about the system center of mass, Lagrange found five distinct three-body formations which are invariant when viewed from the rotating reference frame. We will verify these elegant results below. With

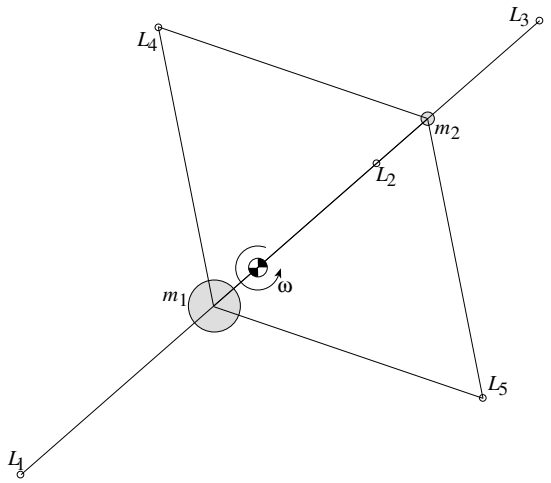


Figure 9.7: Stationary Lagrange Points

the motion of m_1 and m_2 being restricted to a circular orbit, the five possible locations for m for which its location appears invariant or stationary as seen by the rotating frame are called the Lagrange libration points L_1, L_2, L_3, L_4 and L_5 illustrated in Figure 9.7. We can see that the "straight line" points L_1, L_2 and L_3 are evident from the above analysis (Figure 9.6), whereas the points L_4 and L_5 are evident from the equilateral triangle solution shown in Figure 9.5. Their existence was thought to be of purely academic interest when Lagrange first presented these results. However, the Trojan asteroids were subsequently (1906) discovered which oscillate about the L_4, L_5 Sun-Jupiter Lagrange libration points. Also, the Earth-Moon Lagrange points have been studied as possible locations for large "space colony" space stations. Motions near L_4 and L_5 are neutrally stable; these stationary points have become popularly known as "Lunar Libration Points," for the case of the Earth-Moon system.

Please note that the labeling of the collinear libration points is not consistent across different text books. The notation adopted here labels the libration point on the far side of m_1 as L_1, the point between m_1 and m_2 as L_2 and the libration point on the far side of m_2 as L_3 as is illustrated in Figure 9.7. This notation makes geometric sense since the straight line points are labeled in ascending order from one end of the line to the other. While there is no consensus, a second popular alternative is to label the L_2 point as L_1, L_3 as L_2, and L_1 as L_3. The reasoning for this choice is based on relative energy state arguments of bodies at these libration points. However, the L_4 and L_5 are uniformly labeled in the literature as shown here.

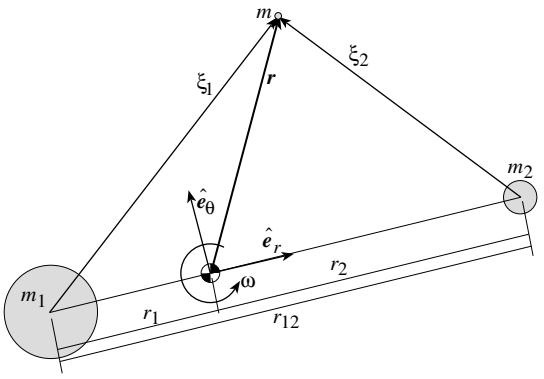


Figure 9.8: Illustration of Circular Restricted Three-Body Problem

9.2.1 Jacobi Integral

To develop the equations of motion of mass m near the circularly orbiting m_1 and m_2, we express the inertial position vector $\boldsymbol{r}$ of m with components taking in a rotating reference frame $\mathcal{F} : \{\hat{\boldsymbol{e}}_r, \hat{\boldsymbol{e}}_\theta, \hat{\boldsymbol{e}}_3\}$. The origin of $\mathcal{F}$ is at the system center of mass as shown in Figure 9.8. Since $m_3 \ll m_1, m_2$, from Eqs. (9.56) or (9.61) the *constant* angular velocity magnitude of the m_1-m_2 system is given by

$$\omega^2 = \frac{G(m_1 + m_2)}{r_{12}^3} \tag{9.64}$$

The angular velocity vector of the $\mathcal{F}$ frame relative to some inertial frame is $\boldsymbol{\omega} = \omega\hat{\boldsymbol{k}}$. The position vector $\boldsymbol{r}$ is expressed with $\mathcal{F}$ frame components as

$$\boldsymbol{r} = r_x\hat{\boldsymbol{e}}_r + r_y\hat{\boldsymbol{r}}_\theta + r_z\hat{\boldsymbol{e}}_3 \tag{9.65}$$

Note that while both m_1 and m_2 perform planar, circular motions, the mass m is able to move both within the orbit plane and perpendicular to it. Taking two inertial derivatives of $\boldsymbol{r}$ while keeping in mind that $\mathcal{F}$ is a rotating reference frame, the inertial acceleration vector of $\boldsymbol{r}$ is expressed as

$$\ddot{\boldsymbol{r}} = (\ddot{r}_x - 2\dot{r}_y\omega - r_x\omega^2)\hat{\boldsymbol{e}}_r + (\ddot{r}_y + 2\dot{r}_x\omega - r_y\omega^2)\hat{\boldsymbol{e}}_\theta + \ddot{r}_z\hat{\boldsymbol{e}}_3 \tag{9.66}$$

The gravitational force $\boldsymbol{F}$ acting on m due to m_1 and m_2 is expressed in $\mathcal{F}$ frame components as

$$\boldsymbol{F} = -G\begin{pmatrix} \frac{m_1}{\xi_1^3}(r_x - r_1) + \frac{m_2}{\xi_2^3}(r_x - r_2) \\ \left(\frac{m_1}{\xi_1^3} + \frac{m_2}{\xi_2^3}\right) r_y \\ \left(\frac{m_1}{\xi_1^3} + \frac{m_2}{\xi_2^3}\right) r_z \end{pmatrix} \tag{9.67}$$

where the relative distances ξ_i of m to m_i are given by

$$\xi_i = \sqrt{(r_x - r_i)^2 + r_y^2 + r_z^2} \tag{9.68}$$

Combining Eqs. (9.66) and (9.67), the equations of motion of m can be written as three scalar, coupled differential equations.

$$\ddot{r}_x - 2\omega\dot{r}_y - \omega^2 r_x + G\left(\frac{m_1}{\xi_1^3}(r_x - r_1) + \frac{m_2}{\xi_2^3}(r_x - r_2)\right) = 0 \tag{9.69a}$$

$$\ddot{r}_y + 2\omega\dot{r}_x - \omega^2 r_y + G\left(\frac{m_1}{\xi_1^3} + \frac{m_2}{\xi_2^3}\right) r_y = 0 \tag{9.69b}$$

$$\ddot{r}_z + G\left(\frac{m_1}{\xi_1^3} + \frac{m_2}{\xi_2^3}\right) r_z = 0 \tag{9.69c}$$

Let the potential function $U(r_x, r_y, r_z)$ be defined as

$$U(r_x, r_y, r_z) = \frac{\omega^2}{2}(r_x^2 + r_y^2) + \frac{Gm_1}{\xi_1} + \frac{Gm_2}{\xi_2} \tag{9.70}$$

Let the time derivative as seen by the $\mathcal{F}$ frame be labeled as

$$\frac{^{\mathcal{F}}d}{dt}\boldsymbol{x} = \boldsymbol{x}' \tag{9.71}$$

Then the velocity and acceleration vectors of m as seen by $\mathcal{F}$ are given by

$$\boldsymbol{r}' = \begin{pmatrix} \dot{r}_x \\ \dot{r}_y \\ \dot{r}_z \end{pmatrix} \qquad \boldsymbol{r}'' = \begin{pmatrix} \ddot{r}_x \\ \ddot{r}_y \\ \ddot{r}_z \end{pmatrix} \tag{9.72}$$

Using the potential function U and the local velocity and acceleration vectors, we are able to write the equations of motion of m in a compact vector form.

$$\boldsymbol{r}'' + 2\boldsymbol{\omega} \times \boldsymbol{r}' = \begin{pmatrix} \frac{\partial U}{\partial r_x} \\ \frac{\partial U}{\partial r_y} \\ \frac{\partial U}{\partial r_z} \end{pmatrix} = \nabla_{\boldsymbol{r}} U \tag{9.73}$$

By performing the vector dot product of Eq. (9.73) with $\boldsymbol{r}'$ we find the following perfect differential equation:

$$(\boldsymbol{r}'' + 2\boldsymbol{\omega} \times \boldsymbol{r}') \cdot \boldsymbol{r}' = \boldsymbol{r}'' \cdot \boldsymbol{r}' = \frac{1}{2}\frac{d}{dt}(\boldsymbol{r}' \cdot \boldsymbol{r}') = \frac{\partial U}{\partial \boldsymbol{r}} \cdot \boldsymbol{r}' = \frac{dU}{dt} \tag{9.74}$$

Integrating this equation with respect to time yields a perfect integral of the relative equations of motion.

$$v^2 = \boldsymbol{r}' \cdot \boldsymbol{r}' = 2U - C \tag{9.75}$$

Substituting the definition of U we find *Jacobi's Integral* for the circular restricted three-body problem.

$$v^2 = \omega^2(r_x^2 + r_y^2) + 2\frac{Gm_1}{\xi_1} + 2\frac{Gm_2}{\xi_2} - C \tag{9.76}$$

where the scalar constant C is determined through the initial conditions. Think of C as a negative, relative energy measure. The larger C is, the less relative energy the mass m has. This perfect integral of the relative equations of motion is used to study what trajectories of m are feasible given some initial energy state and as a means to verify the accuracy of a numerical integration. At any point in time of the motion governed by Eqs. (9.73), the Jacobi integral in Eq. (9.76) must be satisfied. We mention that Jacobi's Integral is simply the classical energy integral ($T + V =$ constant), expressed in rotating coordinates.

The equations of motion in Eq. (9.69) can be written in a convenient *non-dimensional* form. To do so, we introduce the non-dimensional time variable τ as

$$\tau = \omega t \tag{9.77}$$

Time derivatives with respect to this new time variable are denoted with the "o" symbol as

$$\overset{\text{o}}{x} = \frac{dx}{d\tau} \tag{9.78}$$

The non-dimensional time derivative $\overset{\text{o}}{x}$ is related to the previous time derivative $\dot{x}$ through

$$\dot{x} = \frac{dx}{dt} = \frac{dx}{d\tau}\frac{d\tau}{dt} = \overset{\text{o}}{x}\,\omega \tag{9.79}$$

Any scalar distances are non-dimensionalized by dividing them with the *constant* relative m_1-m_2 distance r_{12} as

$$x = \frac{r_x}{r_{12}} \qquad y = \frac{r_y}{r_{12}} \qquad z = \frac{r_z}{r_{12}} \qquad x_1 = \frac{r_1}{r_{12}} \qquad x_2 = \frac{r_2}{r_{12}} \tag{9.80}$$

Note that with the new non-dimensional coordinates the masses m_1 and m_2 are a unit distance apart and that therefore

$$x_2 - x_1 = 1 \tag{9.81}$$

The mass quantities are non-dimensionalized by introducing the scalar parameter μ as

$$\mu = \frac{m_2}{m_1 + m_2} = \frac{1}{\frac{m_1}{m_2} + 1} \tag{9.82}$$

Since the designation of m_1 and m_2 is typically chosen such that $m_2 \leq m_1$, we note that $\mu \leq 0.5$. Using these non-dimensional quantities, the center of mass condition in Eq. (9.6) is rewritten for the current setting as

$$(1 - \mu)x_1 + \mu x_2 = 0 \tag{9.83}$$

Using Eqs. (9.81) and (9.29) we are able to express the non-dimensional coordinates of m_1 and m_2 in terms of the mass ratio μ.

$$x_1 = -\mu \tag{9.84}$$
$$x_2 = 1 - \mu \tag{9.85}$$

Combining all these definitions, we now are able to rewrite the equations of motion of m in Eq. (9.69) into the following non-dimensional form:

$$\overset{\circ\circ}{x} - 2\,\overset{\circ}{y} = x - (1-\mu)\frac{x-x_1}{\rho_1^3} - \mu\frac{x-x_2}{\rho_2^3} = -\frac{\partial U}{\partial x} \tag{9.86a}$$
$$\overset{\circ\circ}{y} + 2\,\overset{\circ}{x} = \left(1 - \frac{1-\mu}{\rho_1^3} - \frac{\mu}{\rho_2^3}\right) y = -\frac{\partial U}{\partial y} \tag{9.86b}$$
$$\overset{\circ\circ}{z} = -\left(\frac{1-\mu}{\rho_1^3} + \frac{\mu}{\rho_2^3}\right) z = -\frac{\partial U}{\partial z} \tag{9.86c}$$

where the non-dimensional relative distance ρ_i is defined as

$$\rho_i = \sqrt{(x-x_i)^2 + y^2 + z^2} \tag{9.87}$$

and the corresponding non-dimensional potential function $U(x, y, z)$ is given by the expression

$$U(x, y, z) = \frac{1}{2}(x^2 + y^2) + \frac{1-\mu}{\rho_1} + \frac{\mu}{\rho_2} \tag{9.88}$$

Following similar steps as were done with the dimensional equations of motion, the non-dimensional Jacobi integral takes on the form

$$v^2 = (\overset{\circ}{x}{}^2 + \overset{\circ}{y}{}^2) = (x^2 + y^2) + 2\frac{1-\mu}{\rho_1} + 2\frac{\mu}{\rho_2} - C \tag{9.89}$$

Setting the relative velocities and accelerations in Eq. (9.86) equal to zero we find conditions which are satisfied by the *stationary points* of the circular restricted three-body problem. As will soon be evident, these stationary points are precisely the five Lagrange Libration points, L_i, specialized to the case at had. Studying Eq. (9.86c) we see that all stationary points have $z = 0$ and therefore must lie in the rotating m_1-m_2 plane. Eq. (9.86b) is only equal to zero for the two known geometric configurations. Either $y = 0$ which corresponds to the *collinear* solution with all bodies aligned with the rotating $\hat{\boldsymbol{e}}_r$ axis, or $\rho_1 = \rho_2$ which corresponds to the *equilateral triangle* solution. To solve for the scalar coordinate x for the collinear L_1, L_2 and L_3 libration points, Eq. (9.86a) is set equal to zero. Notice in Eq. (9.86a) the final two terms, for $y = z = 0$, simplify to

$$-(1-\mu)\frac{(x-x_1}{|x-x_1|^3} - \mu\frac{(x-x_2}{|x-x_2|^3}$$

while one is tempted to merely cancel in these fractions, care must be taken to obtain the correct signs because obviously

$$(x - x_i) = -|x - x_i| \qquad \text{if } (x - x_i) < 0$$
$$(x - x_i) = +|x - x_i| \qquad \text{if } (x - x_i) > 0$$

Using the libration point labeling shown in Figure 9.7, for the L_1 point it is clear that $x - x_1 < 0$ and $x - x_2 < 0$. Using these facts and Eqs. (9.84), (9.85) and (9.87) we find from Eq. (9.86a) an explicit condition for the L_1 position coordinate in terms of the mass ratio μ.

$$L_1 : \qquad x + \frac{1-\mu}{(\mu + x)^2} + \frac{\mu}{(x - 1 + \mu)^2} = 0 \tag{9.90}$$

This equation in essence replaces Lagrange's quintic equation for the circular restricted cases where $m_3 \ll m_1, m_2$. Examining the L_2 Lagrange libration point we find that $x - x_1 > 0$ and $x - x_2 < 0$. This leads to the necessary condition

$$L_2 : \qquad x - \frac{1-\mu}{(\mu + x)^2} + \frac{\mu}{(x - 1 + \mu)^2} = 0 \tag{9.91}$$

For L_3 we find that $x - x_1 > 0$ and $x - x_2 > 0$ and therefore

$$L_3 : \qquad x - \frac{1-\mu}{(\mu + x)^2} - \frac{\mu}{(x - 1 + \mu)^2} = 0 \tag{9.92}$$

Eqs. (9.90) through (9.92) provide three simplified, explicit relationships to solve for the three collinear Lagrange Libration point x coordinates. The advantage of these compared to solving the Lagrange quintic equation is that there is no need to reorder the indices to obtain the three possible solutions. By making use of the *known* sign of the $x - x_i$ terms for each libration point we are able to find these simplified expressions.

Example 9.5: For the Earth-Moon system assuming the approximation that their paths both describe circular orbits about the common center of mass, we can make use of the circular restricted three-body problem to compute the straight line libration points. For this system the mass ratio μ is given by

$$\mu = \frac{1}{81.3 + 1}$$

Solving Eqs. (9.90) through (9.92) for the non-dimensional x coordinates we find

$$L_1 : \quad x = -1.00506 \qquad L_2 : \quad x = 0.836915 \qquad L_3 : \quad x = 1.15568$$

Remember that the radial Earth-Moon distance is equal to 1 with these non-dimensional coordinates. As expected, the L_2 point is between the Earth-Moon system, the L_1 is on the "back-side" of Earth and L_3 is on the "back-side" of the Moon.

9.2.2 Zero Relative Velocity Surfaces

Jacobi's integral in Eq. (9.76) or (9.89) provides a very interesting exact integral of the relative equations of motion of m. The accuracy of a numerical simulation can be verified by checking that the constant C indeed remains invariant during a simulation. Another popular use of the Jacobi integral is to establish regions around m_1 and m_2 within which m may travel given its initial states. For unpowered flight, the initial position and velocity determine the resulting trajectory. Through the Jacobi integral, the initial states also determine the value of the constant C. Recall that C provides a negative relative energy measure of the body m. Extreme points on a trajectory are encountered whenever the velocity magnitude v goes to zero. Therefore, setting $v = 0$ in the Jacobi integral *for a given* energy constant C provides an algebraic expression of all such feasible (x, y, z) "apogee-like" locations.

$$(x^2 + y^2) + 2\frac{1-\mu}{\rho_1} + 2\frac{\mu}{\rho_2} = C \tag{9.93}$$

The surfaces described by Eq. (9.93) determine the geometric extremes possible

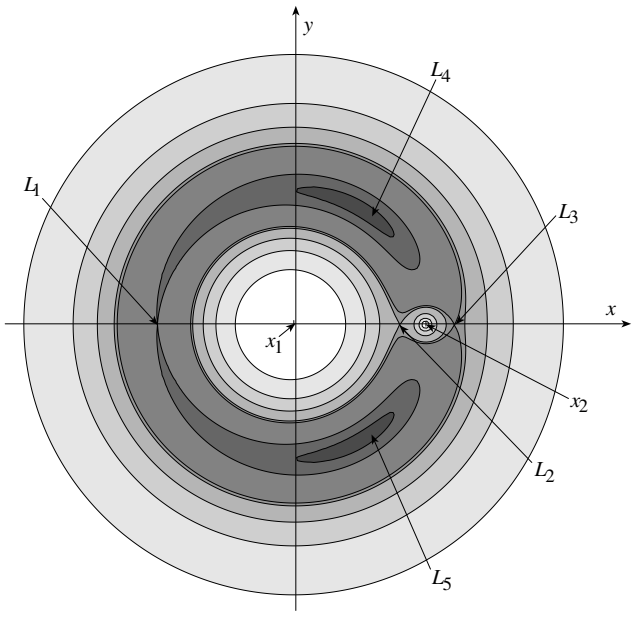


Figure 9.9: Zero Relative Velocity Surface Contours of the Earth-Moon System in the $x - y$ Plane

for a given relative energy state. For example, studying these surfaces allows us to quickly determine if it is possible for a body to travel from Earth to Moon or beyond. Note that when a mass m is located on this surface, this only implies that its velocity *relative to the rotating* $\mathcal{F}$ frame is zero, not that its inertial

velocity is zero. However, studying the motion of m as seen by the rotating reference frame is very convenient when exploring possible trajectories near two orbiting celestial bodies. Selected zero relative velocity surface contours for the Earth (mass m_1) and Moon (mass m_2) system are shown in Figure 9.9. The darker the coloring, the more energy (relative to the rotating frame) is required to enter these areas. When an object has a low relative energy state (i.e. the constant C is large), it may be in one of three areas. The areas are the immediate vicinity around either Earth or Moon, or far removed from the Earth-Moon system. Studying Eq. (9.93) it is evident that if (x, y) and C are large, then ρ_1 and ρ_2 are also large and the zero velocity surface expression is dominated by the quadratic terms in Eq. (9.93) and can be approximated as

$$x^2 + y^2 = C \tag{9.94}$$

This implies that away from the circularly orbiting two body system, the zero velocity surface becomes a circular cylinder with the symmetry axis aligned with the rotation axis $\hat{\boldsymbol{e}}_3$. In the planar cross section shown in Figure 9.9, this is visible as circular constant energy contours. If the body m is close to either m_1 or m_2 while C is large, then ρ_1 or ρ_2 respectively will become small and the dominance of either the second or third term in Eq. (9.93) will result in the zero velocity surface being approximated either by

$$\rho_1^2 = (x - x_1)^2 + y^2 + z^2 = \frac{4(1-\mu)^2}{C^2} \tag{9.95}$$

for ρ_1 being small or by

$$\rho_2^2 = (x - x_2)^2 + y^2 + z^2 = \frac{4\mu^2}{C^2} \tag{9.96}$$

for ρ_2 being small. For this limiting case the zero relative velocity surfaces shapes converge to perfect spheres about either the locations of either m_1 or m_2. Figure 9.9 illustrates this behavior through the white circular contours around both Earth and Moon. As the energy of m increases, more volume becomes accessible and the closed regions near the Earth or Moon ultimately open such that motion is not confined to remain near either massive body; at these energy stages "interchange" orbits are feasible. The last two excluded regions correspond to the minimum C regions around the L_4 and L_5 Lagrange libration points. Figure 9.10 shows individual zero velocity surface contours in the $x-y$ plane for various critical energy state. Regions that cannot be reached by m with the current energy state are grayed out with the same darkness as they have in Figure 9.9. The first critical state is where m has just enough energy to reach the L_2 point between m_1 and m_2. Increasing the energy state of m infinitesimally beyond this state opens up a corridor between the two orbiting bodies, making it theoretically possible for an object to pass from one body to another. However, it is still impossible for an object near m_1 or m_2 to leave the two body system. The next critical energy state is where the zero velocity surface reaches the L_3 point. Any additional energy now makes it possible for

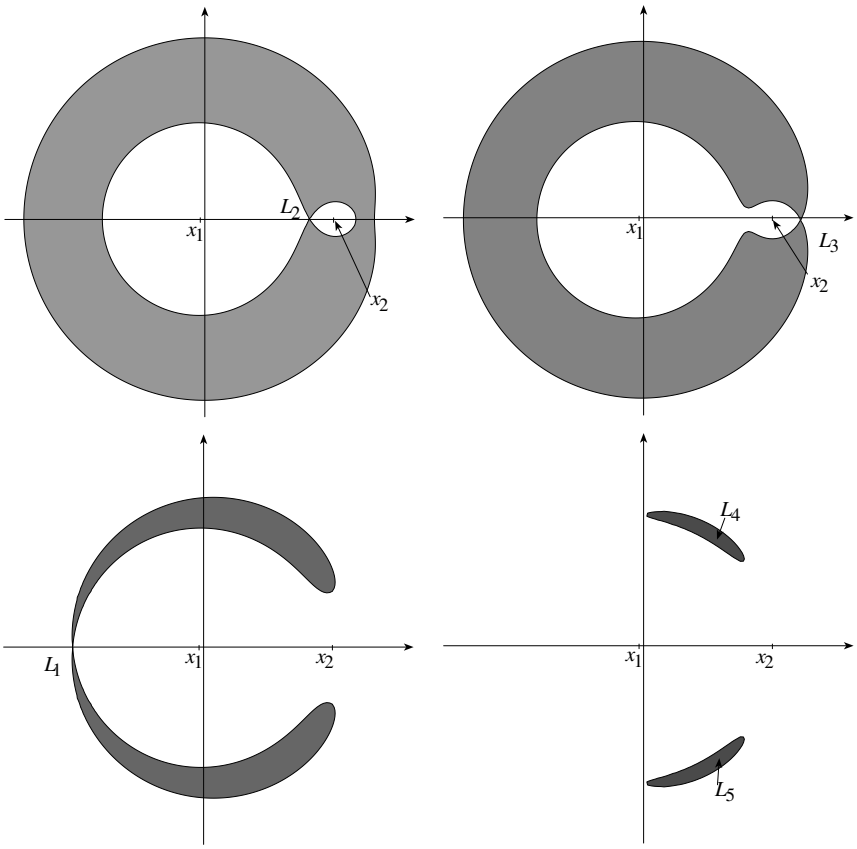
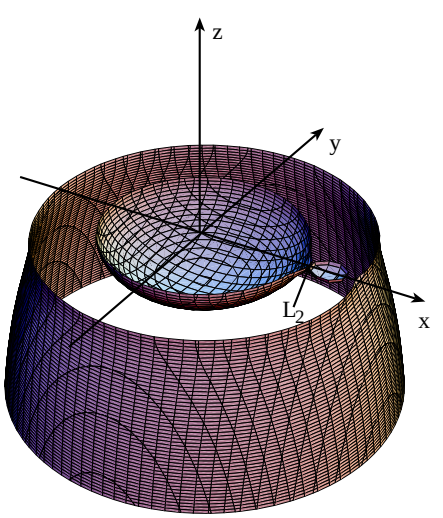


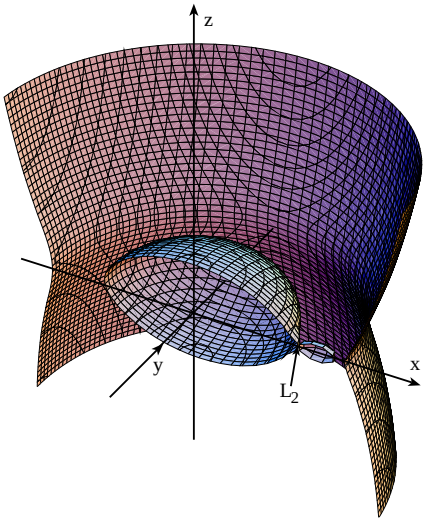
Figure 9.10: Critical Zero Relative Velocity Surface Contours of the Earth-Moon System Touching the Lagrange Stationary Points

m to theoretically escape the two-body system. The next critical stage is where the zero velocity surface just touches the L_1 point. The last forbidden regions to vanish are the banana shaped ones around the L_4 and L_5 libration points. Therefore, from an energy point of view, some people opt to label L_2 as the first libration point, since it requires the least amount of energy to reach. The remaining libration points are then labeled according to the ascending energy states.

Three-dimensional cross sections of these critical zero relative velocity surfaces of the Earth-Moon system are shown in Figures 9.11 through 9.14. Both $x - y$ and $x - z$ cross sectional cuts are shown. Due to symmetry, the missing halves of these surfaces are mirror images of the displayed halves. Figure 9.11 shows the critical surface which touches L_2. A steep cylindrical wall separates possible trajectories outside the two-body system and within them. At this critical energy state, the oval surface shape about the Moon just touches the

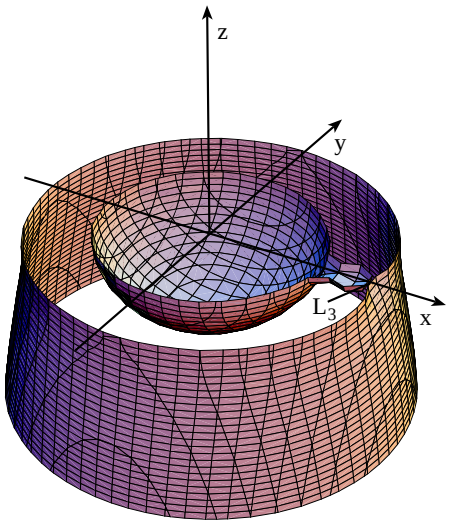


(i) Cross-Section Cut at $z = 0$

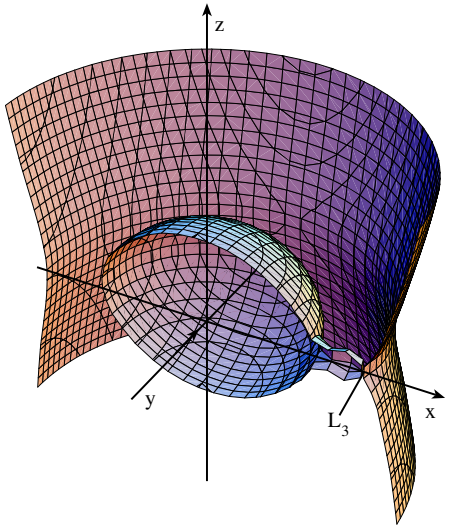


(ii) Cross-Section Cut at $y = 0$

Figure 9.11: Zero Relative Velocity Surface Touching L_2

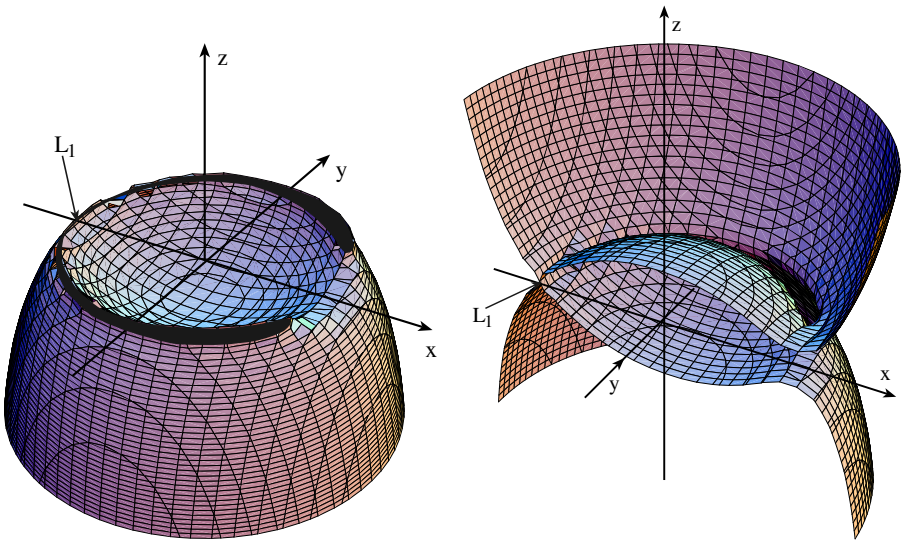


(i) Cross-Section Cut at $z = 0$



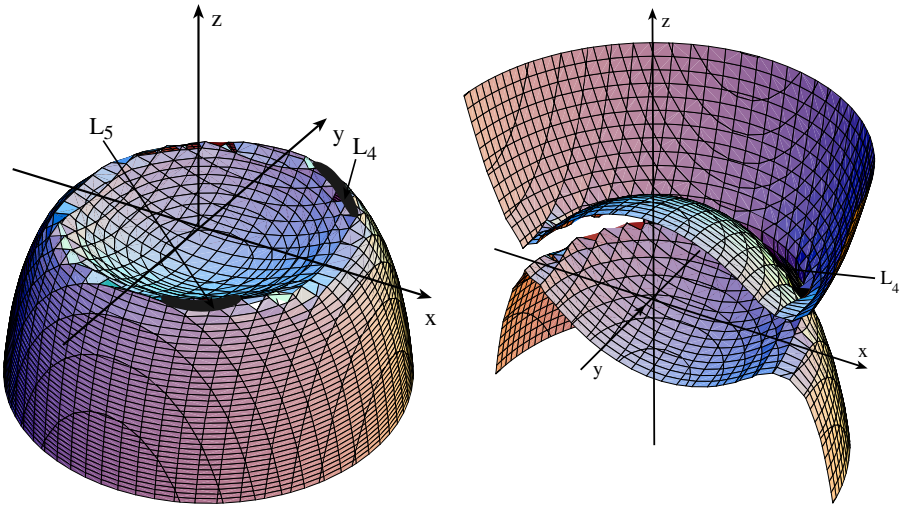
(ii) Cross-Section Cut at $y = 0$

Figure 9.12: Zero Relative Velocity Surface Touching L_3



(i) Cross-Section Cut at $z = 0$ **(ii)** Cross-Section Cut at $y = 0$

Figure 9.13: Zero Relative Velocity Surface Touching L_1



(i) Cross-Section Cut at $z = 0$ **(ii)** Cross-Section Cut at $y = 0$

Figure 9.14: Zero Relative Velocity Surface in the Neighborhood of L_4 and L_5

corresponding surface about Earth. Further, due to the centrifugal effect of moving in an orbiting two body system, the possible motions of body m are clearly more restricted along the two-body rotation axis than within the rotation plane (x, y). This is manifested through the flattening effect of the oval surface shapes about Earth and Moon. As is shown in Eqs. (9.95) and (9.96), for lower energy states (larger C) about Earth and Moon these surfaces will approach spherical shapes.

As the energy state of m increases, a corridor for feasible motion opens between Earth and Moon as is shown in Figure 9.12. Note however that m cannot escape the Earth-Moon system yet until the "bubble" around the Moon just touches the outside cylindrical surface. Increasing the relative energy state of m even more now opens up a corridor through to the open region on the far side of the Moon. Note that m does not yet have enough energy to escape Earth's gravitational influence without the assistance of the moon at this point. However, since the moon acts here as a gravitational boost, it is possible for the body m to escape Earth's influence.

Adding more energy, the body m can escape the Earth-Moon system through multiple directions. Only the regions around the L_4 and L_5 Lagrange stationary points are still unreachable with zero relative velocity. The vertical cylinder is now almost completely separated into a upper and lower sections. The two surfaces only connect around L_4 and L_5. Note that the energy increase between a surface touching the L_4 and L_5 points and a surface touching the L_1 point is relatively small. The energy increments between each surface are not uniform in Figures 9.11 through 9.14. Rather, particular energy states were chosen to illustrate interesting behaviors. Further, note that these surfaces only show the limits to all feasible trajectories. Given only these surfaces, no statements can be made about the various trajectories themselves however.

Example 9.6: Two Earth-Moon trajectories are illustrated as seen by the rotating reference frame $\mathcal{F}$. The first example illustrates an Apollo type mission discussed in Reference 2. At periselenium (point of closest approach to the moon), the mass m has the non-dimensional coordinates

$$(x, y) = (0.992761, 0)$$

which correspond to a miss distance from the Moon's surface of about 150 km. The non-dimensional relative velocity magnitude v at periselenium is 2.47 or about 2.531 km/s. The corresponding Earth-Moon trajectory is shown as seen in the rotating reference frame in Figure 9.15.

The famous hour-glass trajectory is thickened out in this figure since it drawn in the rotating reference frame. The critical zero-velocity surface which touches the L_3 Lagrange Libration point is superimposed in this illustration. Clearly this trajectory penetrates this surface. Thus the body m has enough energy to escape the Earth-Moon system given the proper initial position and velocity direction. This illustrates again that the zero velocity surface can only predict which regions a body will not be able to enter with a given energy level. They do no predict what paths a body will take. More specifically,

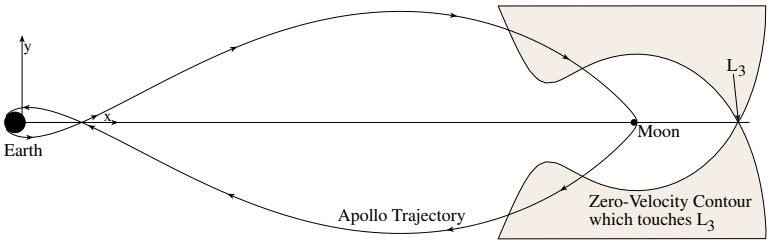


Figure 9.15: Apollo Type Earth-Moon Trajectory as seen by a Rotating Reference Frame

restricting energy below the value corresponding to a given closed zero velocity surface, we are *guarantee* the body cannot exit that closed region; we cannot guarantee a higher energy will escape that surface during any finite time interval.

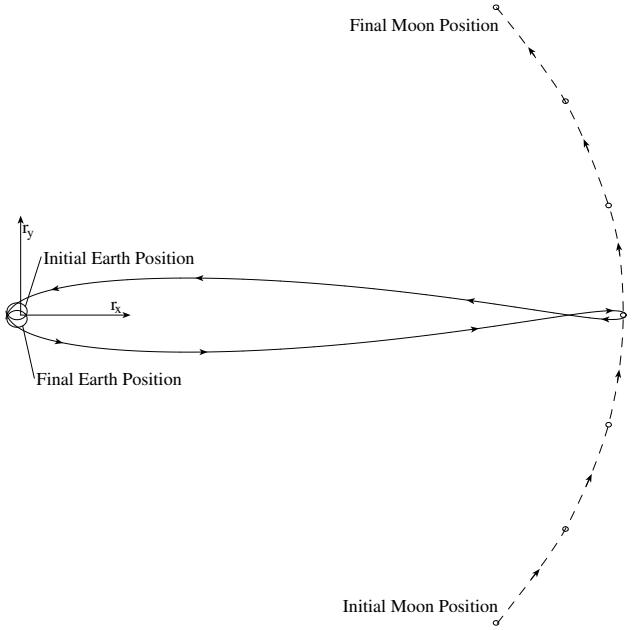


Figure 9.16: Apollo Type Earth-Moon Trajectory as seen by an Inertial Reference Frame

The same trajectory is shown as seen by a non-rotating, inertial frame in Figure 9.16. This view illustrates well how the traveling Moon sharply bends the spacecraft trajectory back towards Earth. Also, it is evident that the Moon travels a large distance during this trajectory, thus constantly changing the direction of its gravitational pull. While this figure provides a better illustration of the actual flight path shape, note that it is more difficult to assess if the spacecraft will impact with either Earth or Moon. Both bodies have

time varying positions and it isn't clear that the spacecraft trajectory does not intersect with either the Earth's or the Moon's surface. This analysis is best performed in Earth- or Moon-centered coordinate systems which translate with the Earth or Moon and study only the portion of the trajectory near closest approach.

As a comparison, another Earth-Moon trajectory is shown in Figure 9.17. At periselenium, the body m has a Moon surface miss distance of 884 km and the critical L_3 velocity of 1.872 km/s.

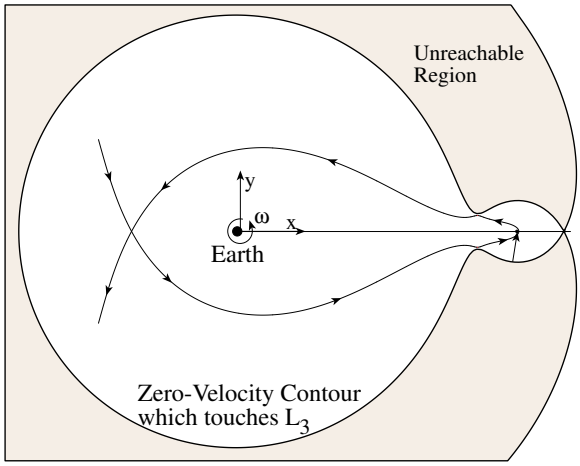


Figure 9.17: Subcritical Earth-Moon Trajectory as seen by Rotating Reference Frame

This trajectory has a close approach with the critical L_3 surface, meaning that at that point the relative velocity of m as seen by the rotating frame is close to zero. Even though the energy state of this trajectory is less than the previous one, it is clearly less desirable for mission planning. The closest approach to the Moon and especially to Earth are much larger. This would require additional maneuvers out of Earth and Moon parking orbits to reach the this trajectory. A quick numerical study shows that it is impossible to reach the Moon from a tight Earth parking orbit with a sub-critical L_3 energy state.

9.2.3 Lagrange Libration Point Stability

Of particular interest is whether motions near the Lagrange stationary points L_i are stable solutions of the relative equations of motion. The question is: If a body starts out at rest near a Lagrange libration point, will it remain in the vicinity or will it wander off over time? If the motions are stable, one could expect that it would take less fuel for a spacecraft to maintain its relative position there. Studying the zero velocity contours in Figures 9.9 and 9.10,

initial guesses can be made as to the stability of the five libration points. Due to the "saddle-point" nature of the zero-velocity contours touching L_1, L_2 and L_3, we can expect the first three Libration points to be unstable. Only the L_4 and L_5 Libration points of the equilateral triangle solution may by neutrally stable. To study the stability of a particular L_i point, we linearize the relative equations of motion about L_i and check whether any eigenvalues of the linearized plant matrix have positive, real components. To simplify the development, the non-dimensional equations of motion in Eq. (9.86) are written in vector form as

$$\overset{\circ\circ}{\boldsymbol{r}} + [\tilde{\boldsymbol{a}}]\,\overset{\circ}{\boldsymbol{r}} + [\tilde{\boldsymbol{a}}]^2\boldsymbol{r} = -\frac{1-\mu}{\rho_1^3}(\boldsymbol{r}-\boldsymbol{r}_1) - \frac{\mu}{\rho_2^3}(\boldsymbol{r}-\boldsymbol{r}_2) = \boldsymbol{f}(\boldsymbol{r}) \tag{9.97}$$

where $\boldsymbol{r} = (x, y, z)^T$, $\boldsymbol{r}_1 = (-\mu, 0, 0)^T$, $\boldsymbol{r}_2 = (1-\mu, 0, 0)^T$ and $\boldsymbol{\Omega} = (1, 0, 0)^T$ is a non-dimensionalized angular velocity vector. Let the departure motion $\delta\boldsymbol{r}$ about a point $\boldsymbol{r}_0$ be defined as

$$\delta\boldsymbol{r} = \boldsymbol{r} - \boldsymbol{r}_0 \tag{9.98}$$

Then the linearized departure motion about $\boldsymbol{r}_0$ is given by

$$\delta\overset{\circ\circ}{\boldsymbol{r}} + [\tilde{\boldsymbol{\Omega}}]\delta\overset{\circ}{\boldsymbol{r}} + [\tilde{\boldsymbol{\Omega}}]^2\delta\boldsymbol{r} = \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{r}}\right|_{\boldsymbol{r}_0} \delta\boldsymbol{r} \tag{9.99}$$

To evaluate the partial derivative of $\boldsymbol{f}$ with respect to $\boldsymbol{r}$, the following partial derivative is useful:

$$\frac{\partial \rho_i}{\partial \boldsymbol{r}} = \frac{1}{\rho_i}(\boldsymbol{r}-\boldsymbol{r}_i)^T \tag{9.100}$$

Linearizing the force vector $\boldsymbol{f}$ about $\boldsymbol{r}_0$ then yields

$$\begin{aligned}\boldsymbol{F} = \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{r}}\right|_{\boldsymbol{r}_0} &= \frac{1-\mu}{\rho_1^5}\left(3(\boldsymbol{r}-\boldsymbol{r}_1)(\boldsymbol{r}-\boldsymbol{r}_1)^T - \rho_1^2[I_{3\times 3}]\right) \\ &\quad + \frac{\mu}{\rho_2^3}\left(3(\boldsymbol{r}-\boldsymbol{r}_2)(\boldsymbol{r}-\boldsymbol{r}_2)^T - \rho_2^2[I_{3\times 3}]\right)\end{aligned} \tag{9.101}$$

By defining the state vector $\boldsymbol{X}$ as

$$\boldsymbol{X} = (\delta\boldsymbol{r}, \delta\overset{\circ}{\boldsymbol{r}})^T \tag{9.102}$$

we write the equations of motion in Eq. (9.99) in first order state space form

$$\overset{\circ}{\boldsymbol{X}} = [M(\boldsymbol{r}_0)]\boldsymbol{X} \tag{9.103}$$

with the plant matrix $[M]$ defined as

$$[M(\boldsymbol{r}_0)] = \begin{bmatrix} 0_{3\times 3} & I_{3\times 3} \\ F(\boldsymbol{r}_0) - [\tilde{\boldsymbol{\Omega}}]^2 & -2[\tilde{\boldsymbol{\Omega}}] \end{bmatrix} \tag{9.104}$$

By evaluating $[M]$ at the five Lagrange libration points and checking the corresponding eigenvalues for positive real parts, we are able to make some statements concerning the local stability of motion near these points.

Let us first investigate the stability of the collinear L_1, L_2 and L_3 Lagrange libration points. Let $\boldsymbol{r}_0 = (x_0, 0, 0)^T$ be a position vector of one of these stationary points. Since for each of these $y_0 = z_0 = 0$, the terms ρ_1 and ρ_2 evaluated at $\boldsymbol{r}_0$ are

$$\rho_1(\boldsymbol{r}_0) = |x_0 + \mu| \qquad \rho_2(\boldsymbol{r}_0) = |x_0 - 1 + \mu| \tag{9.105}$$

Evaluating the matrix $[F]$ at $\boldsymbol{r}_0$ we then find

$$[F(\boldsymbol{r}_0)] = \begin{bmatrix} 2E & 0 & 0 \\ 0 & -E & 0 \\ 0 & 0 & -E \end{bmatrix} \tag{9.106}$$

with the constant positive scalar E defined as

$$E(\boldsymbol{r}_0) = \frac{1-\mu}{\rho_1^3(\boldsymbol{r}_0)} + \frac{\mu}{\rho_2^3(\boldsymbol{r}_0)} \tag{9.107}$$

Substituting Eq. (9.106) into Eq. (9.104) and computing the six eigenvalues of $[M]$, we find

$$\lambda_{1,2,3,4}^2 = \frac{E - 2 \pm \sqrt{9E^2 - 8E}}{2} \tag{9.108}$$

$$\lambda_{5,6}^2 = -1 \tag{9.109}$$

The collinear stationary points are unstable if Eq. (9.108) is positive, since this results in one of the eigenvalues having a positive real root. For Eq. (9.108) to be positive, then

$$(2E + 1)(E - 1) > 0 \tag{9.110}$$

must be true. Since $E > 0$, this condition can simplified to

$$E > 1 \tag{9.111}$$

It has been shown that $E > 1$ holds for all possible values of μ between 0 and 0.5, for all collinear stationary points.[1] Some of the first four eigenvalues in Eq. (9.108) have positive real parts. Therefore all collinear Lagrange libration points must be considered unstable.

To study the stability of the L_4 and L_5 Lagrange libration points, we note that for the non-dimensional equilateral triangle the relative distances ρ_i are

$$\rho_1 = \rho_2 = 1 \tag{9.112}$$

The position vector $\boldsymbol{r}_0$ of L_4 and L_5 is given by

$$\boldsymbol{r}_0 = \begin{pmatrix} \frac{1}{2} - \mu \\ \pm\frac{\sqrt{3}}{2} \\ 0 \end{pmatrix} \tag{9.113}$$

where the negative sign corresponds to the L_5 case. Substituting Eqs. (9.112) and (9.113) into Eq. (9.101), the $[F]$ matrix is computed for these Lagrange points as

$$[F(\boldsymbol{r}_0)] = \frac{1-\mu}{4}\begin{bmatrix} -1 & \pm 3\sqrt{3} & 0 \\ \pm 3\sqrt{3} & 5 & 0 \\ 0 & 0 & -4 \end{bmatrix} + \frac{\mu}{4}\begin{bmatrix} -1 & \mp 3\sqrt{3} & 0 \\ \mp 3\sqrt{3} & 5 & 0 \\ 0 & 0 & -4 \end{bmatrix} \tag{9.114}$$

Using Eq. (9.114) in Eq. (9.104), we find that the eigenvalues of $[M(\boldsymbol{r}_0)]$ are the same for both L_4 and L_5.

$$\lambda^2_{1,2,3,4} = \frac{-1 \pm \sqrt{1 - 27\mu(1-\mu)}}{2} \tag{9.115}$$

$$\lambda^2_{5,6} = -1 \tag{9.116}$$

For the eigenvalues in Eq. (9.115) to be purely imaginary, it is necessary that

$$1 > 1 - 27\mu(1-\mu) > 0 \tag{9.117}$$

A quick numerical check shows that the upper bound is always satisfied. Since $0 < \mu \leq 0.5$, the maximum permissible μ such that the right inequality of Eq. (9.117) is still satisfied is

$$\mu_{max} = \frac{1}{2} - \frac{1}{6}\sqrt{\frac{23}{3}} \approx 0.0385209 \tag{9.118}$$

Since for the Earth-Moon system $\mu \approx 0.01230$, motions near these L_4 and L_5 points are neutrally stable in their close vicinity. Using the definition of μ, the condition in Eq. (9.117) can also be written as[2]

$$\frac{m_1}{m_2} + \frac{m_2}{m_1} \geq 25 \tag{9.119}$$

Defining α to be the direct ratio of m_1 and m_2, the minimum ratio necessary for Eq. (9.119) to be true is

$$\alpha_{min} = \frac{25 + 3\sqrt{69}}{2} \approx 24.9599 \tag{9.120}$$

This implies for motion near the corresponding L_4 and L_5 Lagrange libration points to be neutrally stable, the larger mass m_1 must be at least roughly 25 times larger than m_2.

Nature has provided us with proof that the equilateral triangle configuration of the circular restricted three-body problem is indeed neutrally stable. In the Sun-Jupiter system ($\mu \approx 0.001$) a group of asteroids called the Trojans have been found in 1906 at the corresponding L_4 and L_5 libration points; a group of five were detected at L_4 and a group of ten at L_5. Since then, over 1000 Trojan asteroids have been found at L_4 and L_5, providing nature's very significant empirical statement regarding the stability of motion near L_4 and L_5. All of these asteroids oscillate in an apparently neutrally stable manner in the vicinity of these stationary points.

Example 9.7: This example illustrates the neutral stability of the Earth-Moon L_4 Lagrange libration points. The initial position vector $\boldsymbol{r}(t_0)$ touches the zero relative velocity surface with $C = 2.9884$.

$$\boldsymbol{r}(t_0) = (0.4925060, .85, 0)^T$$

The resulting motion is illustrated to scale with Earth and Moon in Figure 9.18.

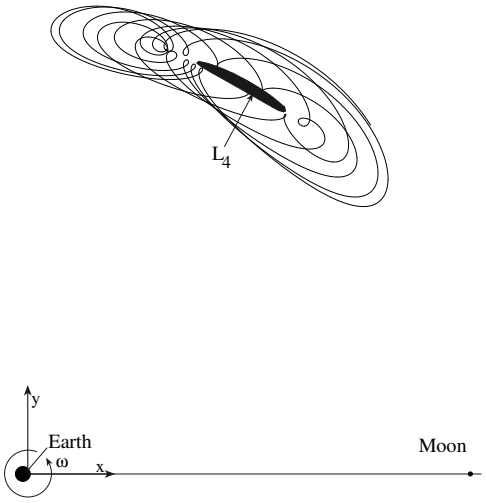


Figure 9.18: Unpowered Motion in the Vicinity of the Earth-Moon L_4 Lagrange Libration Point

While the body m does wander away from the immediate vicinity of L_4, it does "librate" in the neighborhood of L_4. Note that by starting out with zero relative velocity (i.e. touching the shown zero-velocity surface), the body does periodically return very near this surface again. While it can theoretically touch this zero velocity surface, it is not guaranteed to do so. Note that the illustrated zero velocity surface is actually a rather large neighborhood about L_4. If a tighter surface is chosen (i.e. $\boldsymbol{r}(t_0)$ closer to L_4), then the resulting oscillations about L_4 would occupy a much smaller region near L_4. We note that rigorous nonlinear guarantees that the oscillations stay in a bounded region near L_4 have not been established. However, numerical studies and the Trojan asteroids indicate neutral stability for a large family of initial conditions near L_4 and L_5.

9.3 Periodic Stationary Orbits

Clearly the stationary points of the circular restricted three-body problem are of great practical interest. After their discovery by Lagrange, much effort has been

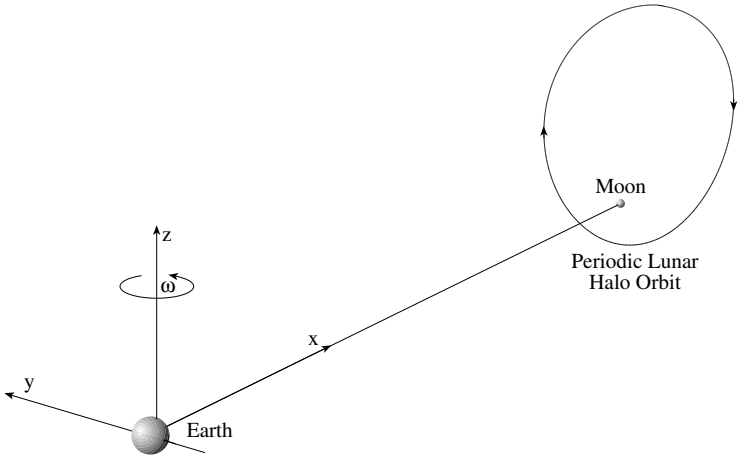
spent on trying to find *periodic stationary orbits* of this three-body problem. These orbits form closed three-dimensional curves and remain fixed as seen by the rotating reference frame $\mathcal{F}$. Thousands such orbits have been found for various values of μ, not all of them are necessarily of practical value. It turns out that these stationary periodic orbits can be grouped into families of orbits which help to explain the more general classes of orbits for this restricted three-body problem. An excellent treatise on this topic is Victor Szebehely's book entitled *Theory of Orbits, The Restricted Problem of Three Bodies.*[3] However, presenting these very intriguing orbits is beyond the scope of this book.

To appreciate the flavor of these nonlinear periodic motions, an interesting stationary periodic lunar orbit is shown in Figure 9.19. This type of orbit is commonly referred to as a "halo" orbit for obvious reasons. As seen from Earth, this orbit describes a ring or "halo" about the Moon. Where previously we were typically illustrating planar trajectories between the Earth and Moon, this lunar orbit illustrates a three-dimensional trajectory. As seen from Earth ($y-z$ plane view in Figure 9.19(ii)) the closed flight path has a shape similar to an ellipse with the moon in what would be the focal point. In the side view in Figure 9.19(iii) it is apparent that the space curve is essentially located on the far side of the moon between the moon and L_3. Also, the "orbit plane" is not flat but curved away from the Earth and Moon toward L_3. By always being in line of sight with Earth and hovering behind the moon, this lunar halo trajectory would be a great parking orbit for a lunar communications satellite. Currently, if any spacecraft travels to the far side of the Moon all communications cease. For future missions to the moon and establishing permanent bases and installations there, having a continuous communication capability between Earth and Moon would be essential. In particular, the far side of the moon is a superb location to place a deep-space radio telescope. However, since the Moon doesn't rotate as seen by Earth, communicating with this installation would be impossible without a communications satellite in lunar orbit.

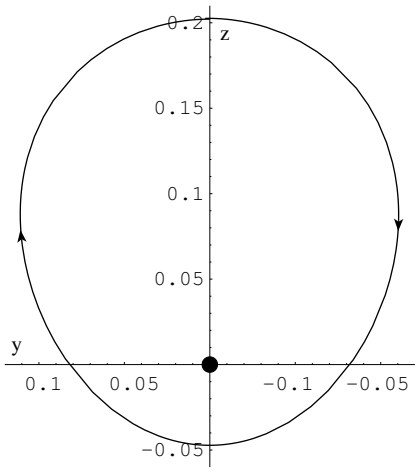
For a periodic orbit to be practically useful, its stability must also be investigated. Since the two massive bodies never precisely have circular orbits about their system mass center, and often there are other bodies such as planets or the sun which exert a small influence on the three-body system, it must be assumed that small external influences are always present. Also, we observe that launch errors result in imperfect initial conditions. Therefore placing a spacecraft in a near-periodic orbit or near a libration point, a periodic orbit correction must be expected to maintain the desired trajectory. The frequency of these corrections would depend on stability of the stationary orbit. Roy[1] outlines a systematic process for searching out periodic orbits in the three-body problem, and Breakwell[4] provides an approach for studying orbital stability.

9.4 The Disturbing Function

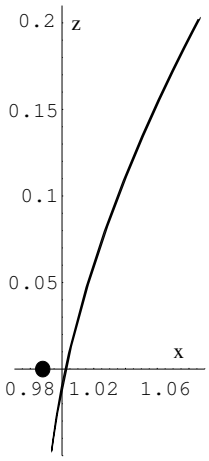
When one celestial body dominates the motion of other bodies, it is often convenient to write the position vectors of these other bodies relative to the dominant



(i) Orthogonal View of Lunar Stationary Halo Orbit



(ii) $y-z$ Plane View of Halo Orbit



(iii) $x-z$ Plane View of Halo Orbit

Figure 9.19: Stationary Lunar Halo Orbit

body. For example, in our solar system the sun is the dominant influence on the motion of the planets. The mutual attraction of the various planets only has a very minor effect on their flight paths. Another example is the lunar motion about Earth. The lunar orbit is mainly determined through the gravitational attraction of the Earth, while the sun's gravitational attraction only has a secondary effect. Writing the equations of motion relative to a dominant celestial body, we are able so what secondary effect the gravitational attraction the remaining bodies will have.

In this section, the motion of each body is not restricted as was the case with the previous cases. Also, we now consider a more general problem where n bodies are present. However, we do assume that one of the bodies plays a dominant role on the motion of a set of bodies. Without loss of generality, let us take m_1 to be the dominant body. From Eq. (9.3), its equations of motion are given by

$$\ddot{\boldsymbol{r}}_1 = G\sum_{j=2}^{n}\frac{m_j}{r_{1j}^3}\boldsymbol{r}_{1j} \tag{9.121}$$

where $\boldsymbol{r}_i$ are inertial position vectors of each body and the relative position vectors $\boldsymbol{r}_{ij}$ are defined as

$$\boldsymbol{r}_{ij} = \boldsymbol{r}_j - \boldsymbol{r}_i = \boldsymbol{r}_{1j} - \boldsymbol{r}_{1i} \tag{9.122}$$

The equations of motion of each remaining body is given by

$$\ddot{\boldsymbol{r}}_i = G\sum_{j=1,j\neq i}^{n}\frac{m_j}{r_{ij}^3}\boldsymbol{r}_{ij} \qquad \text{for } i = 2,\dots,n \tag{9.123}$$

Subtracting Eq. (9.123) from (9.121), the relative equations of motion of the $n-1$ bodies relative to m_1 are found to be

$$\ddot{\boldsymbol{r}}_{1i} = \ddot{\boldsymbol{r}}_i - \ddot{\boldsymbol{r}}_1 = G\left(\frac{m_1}{r_{i1}^3}\boldsymbol{r}_{i1} + \sum_{j=2,j\neq i}^{n}\frac{m_j}{r_{ij}^3}\boldsymbol{r}_{ij} - \sum_{j=2,j\neq i}^{n}\frac{m_j}{r_{1j}^3}\boldsymbol{r}_{1j} - \frac{m_i}{r_{1i}^3}\boldsymbol{r}_{1i}\right) \tag{9.124}$$

Using Eq. (9.122) and $\boldsymbol{r}_{ij} = -\boldsymbol{r}_{ji}$ these relative equations of motion are rewritten as

$$\ddot{\boldsymbol{r}}_{1i} + \frac{G(m_1+m_i)}{r_{1i}^3}\boldsymbol{r}_{1i} = G\sum_{j=2,j\neq i}^{n} m_j\left(\frac{\boldsymbol{r}_{1j}-\boldsymbol{r}_{1i}}{r_{ij}^3} - \frac{\boldsymbol{r}_{1j}}{r_{1j}^3}\right) \tag{9.125}$$

If only two bodies are present, note that the standard Keplerian two-body equations of motion are retrieved. The gravitational effect of bodies other than m_1 and m_i appear a disturbance terms in this formulation.

Example 9.8: Consider the Earth-Moon-Sun three-body system with Earth being mass $m_\oplus$, the moon $m_\leftmoon$ being the mass moving in Earth's proximity, and the sun $m_\odot$ being the external influence of this two-body system. Eq. (9.125) in this case reduces to

$$\ddot{\boldsymbol{r}}_{\oplus\leftmoon} + \frac{G(m_\oplus+m_\leftmoon)}{r_{\oplus\leftmoon}^3}\boldsymbol{r}_{\oplus\leftmoon} = Gm_\odot\left(\frac{\boldsymbol{r}_{\oplus\odot}-\boldsymbol{r}_{\oplus\leftmoon}}{r_{\leftmoon\odot}^3} - \frac{\boldsymbol{r}_{\oplus\odot}}{r_{\oplus\odot}^3}\right)$$

At first glance it may seem like the sun's gravitation influence would be very large in these relative equations of motion since $m_\oplus, m_\leftmoon \ll m_\odot$. However, due to the large relative distance involved between the sun and the Earth

and Moon we can make the approximation $r_{\leftmoon\odot} \approx r_{\oplus\odot}$. Therefore, the lunar relative equation of motion are approximated as

$$\ddot{\boldsymbol{r}}_{\oplus\leftmoon} + \frac{G(m_\oplus + m_\leftmoon)}{r^3_{\oplus\leftmoon}}\boldsymbol{r}_{\oplus\leftmoon} \approx -Gm_\odot \frac{\boldsymbol{r}_{\oplus\leftmoon}}{r^3_{\oplus\odot}}$$

For this Earth-Moon-Sun three-body system, the gravitational acceleration magnitudes of the Earth-Moon attraction and the solar attraction are related through

$$\frac{m_\odot r_{\oplus\leftmoon}}{r^3_{\oplus\odot}} \approx 0.00558\frac{(m_\oplus + m_\leftmoon)}{r^2_{\oplus\leftmoon}}$$

Therefore the solar gravitational attraction on the moon is over two orders of magnitude smaller than the relative Keplerian Earth-Moon motion attraction.

As is the case in Keplerian motion, it is possible to write the acceleration expression as the gradient of a potential function. The gravitational potential function V_i, which leads to a Keplerian $m_1 - m_i$ motion, is given by

$$V_i(\boldsymbol{r}_{1i}) = -\frac{G(m_1 + m_i)}{r_{1i}} \tag{9.126}$$

Note that Eq. (9.126) is identical to the potential function defined in Eq. (8.47). Computing the gradient of V_i with respect to the position vector $\boldsymbol{r}_{1i}$ we find

$$\nabla_{\boldsymbol{r}_{1i}} V_i = \frac{G(m_1 + m_i)}{r^3_{1i}}\boldsymbol{r}_{1i} \tag{9.127}$$

Eq. (9.127) provides the acceleration due to the gravitational attraction between m_1 and m_i. To compute the disturbance acceleration due the the remaining $(n-2)$ bodies, we define the scalar *disturbance function* R_i to be

$$R_i = G\sum_{j=2, j\neq i}^{n} m_j \left(\frac{1}{r_{ij}} - \frac{\boldsymbol{r}_{1i}\cdot\boldsymbol{r}_{1j}}{r^3_{1j}}\right) \tag{9.128}$$

The gradient of the disturbance function R_i produces the acceleration due the gravitational attraction with the remaining bodies.

$$\nabla_{\boldsymbol{r}_{1i}} R_i = G\sum_{j=2, j\neq i}^{n} \left(\frac{\boldsymbol{r}_{1j} - \boldsymbol{r}_{1i}}{r^3_{ij}} - \frac{\boldsymbol{r}_{1j}}{r^3_{1j}}\right) \tag{9.129}$$

Using Eqs. (9.127) and (9.129), the equations of motion of m_i relative to m_1 are written compactly as

$$\ddot{\boldsymbol{r}}_{1i} = -\nabla_{\boldsymbol{r}_{1i}}(V_i - R_i) \tag{9.130}$$

Example 9.9: The gravitational solar disturbance acceleration the relative motion of the Moon experiences in orbit about the Earth was very small in Example 9.8 because of the large distances involved between the sun and the Earth-Moon system. In this example, we treat the sun ($m_\odot$) as the dominant mass influencing the motion of the Earth ($m_\oplus$). The gravitational effect of the moon ($m_\leftmoon$) on the relative motion of Earth about the sun is now treated as a perturbation. The relative equations of motion of the Earth about the sun are written as

$$\ddot{\boldsymbol{r}}_{\odot\oplus} + \frac{G(m_\odot + m_\oplus)}{r^3_{\odot\oplus}}\boldsymbol{r}_{\odot\oplus} = Gm_\leftmoon\left(\frac{\boldsymbol{r}_{\odot\leftmoon} - \boldsymbol{r}_{\odot\oplus}}{r^3_{\leftmoon\oplus}} - \frac{\boldsymbol{r}_{\odot\leftmoon}}{r^3_{\odot\leftmoon}}\right)$$

The magnitude of the perturbative acceleration due to the moon has the upper bound

$$Gm_\leftmoon\left(\frac{\boldsymbol{r}_{\odot\leftmoon} - \boldsymbol{r}_{\odot\oplus}}{r^3_{\leftmoon\oplus}} - \frac{\boldsymbol{r}_{\odot\leftmoon}}{r^3_{\odot\leftmoon}}\right) \leq Gm_\leftmoon\left(\frac{2}{r^2_{\leftmoon\oplus}} - \frac{1}{r^2_{\odot\leftmoon}}\right) \approx \frac{2Gm_\leftmoon}{r^2_{\leftmoon\oplus}}$$

where the approximation is done since $r_{\leftmoon\oplus} \ll r_{\odot\leftmoon}$. Note that this upper bound on lunar influence is computed when the Moon lies perfectly between the Sun and Earth or on the far side of the Earth. Therefore the acceleration magnitudes of the two-body solution and the lunar gravitational attraction relate through

$$\frac{2m_\leftmoon}{r^2_{\leftmoon\oplus}} \approx 0.0112\frac{(m_\odot + m_\oplus)}{r^2_{\odot\oplus}}$$

During periods of strongest influence, the gravitational lunar attraction of the Earth only affects its orbit about the sun by about 1 percent.

Problems

9.1 Compute where the Moon would have to be located relative to the Earth for it to continuously be in a "full Moon" state as seen from Earth, and ignoring the Earth's shadow. Treat the Sun-Earth-Moon as a three-body system.

9.2 Assume that the bodies $m_1 = 2m_2 = 3m_3$ are in elliptic orbits about their system center of mass where m_1 has the mass of Earth.

a) Choose a coordinate system and compute the location of each mass if they are either in a triangular or collinear configuration.

b) Assuming that the initial velocity magnitude of each body is given by Eq. (9.56), numerically compute the resulting orbits if the initial velocity vectors form a 30° angle with the respective radial position vectors.

c) Repeat the previous task with the initial velocity vectors forming 90° angles with the corresponding radial position vectors.

9.3 Set up the three masses $m_1 = 2m_2 = 3m_3$ in an equilateral triangle configuration with a zero angular velocity about the system mass center.

a) Verify that the net sum force $\boldsymbol{F}_i$ acting on each body m_i indeed points from each mass directly to the system mass center.

b) Verify that the force magnitude F_i of each mass is equal to GM_i/r_i^2 where the equivalent masses M_i are defined in Eqs. (9.38) through (9.40).

9.4 For the Sun-Earth two-body system, compute the corresponding libration points in both dimensional and non-dimensional units.

9.5 Derive the non-dimensional equations of motion in Eq. (9.86) from Eq. (9.69). Show all algebra.

9.6 Treating the Sun-Earth-spacecraft system as a circular restricted three-body problem, what is the minimum relative energy v^2 necessary that a spacecraft must have to travel in theory from Earth to the sun?

9.7 For the Sun-Earth system, compute the corresponding zero-velocity surfaces that touch the L_1, L_2 and L_3 points.

9.8 ♣ Assume the Sun-Jupiter system two-body system is describing circular orbits about their center of mass. The radial distance between Jupiter and the sun is $778 \cdot 10^6$ km. The mass of Jupiter is $m_{♃} = 1.9 \cdot 10^{27}$ kg and the mass of the sun is $m_{\odot} = 1.989 \cdot 10^{30}$ kg.

a) Compute the five stationary Lagrange points as seen by the rotating coordinate system.

b) Investigate the stability of each point placing an object in the neighborhood of each libration point and numerically solving for the resulting motion. Make conclusions on the "level" of stability or instability.

9.9 Consider the Keplerian two-body problem where the inertially fixed mass m_1 is a very massive body as compared to the second mass m. The gravitational potential of m is assumed to be too small to affect the motion of m_1.

a) Write the energy (or vis-viva) equation in a form analogous to Eq. (9.76) or (9.89).

b) Investigate the zero-velocity surfaces for this two-body problem. What type of orbit must m have about m_1 for it to reach this surface?

9.10 Consider the Earth-Moon system. Compute the three locations where the centrifugal acceleration is canceled exactly by the combined gravitational attraction of Earth and Moon. Verify that these three locations are indeed the three collinear Lagrange stationary solutions of the circular restricted three-body problem.

Bibliography

[1] Roy, A. E., *Orbital Motion*, Adam Hilger Ltd, Bristol, England, 2nd ed., 1982.

[2] Battin, R. H., *An Introduction to the Mathematics and Methods of Astrodynamics*, AIAA Education Series, New York, 1987.

[3] Szebehely, V., *Theory of Orbits, The Restricted Problem of Three Bodies*, Academic Press, New York, 1967.

[4] Breakwell, J. V. and Brown, J. V., "The 'Halo' Family of 3-Dimensional Periodic Orbits in the Earth-Moon Restricted 3-Body Problem," *Celestial Mechanics*, Vol. 20, 1979, pp. 389–404.

Chapter Ten

Gravitational Potential Field Models

Keplerian motion is equivalent to the gravitational two body problem; i.e. the situation where the motion of a particle of unit mass is determined by a point mass gravitational field $V_0 = -Gm/r$ generated by a second body (a point mass or spherically symmetric finite body) of mass m. The scalar parameter r is the relative distance between the two center of masses and G is the universal gravity constant. This type of gravity field is also sometimes referred to as an inverse square gravity field, since the gravitational force drops of with the square of the relative distance. The study of the motion induced by such simple gravity fields has led to the well known analytical solution of the two body problem. The essential features of this solution were discovered empirically and geometrically by Kepler who published these results in 1609. Newton subsequently derived an analytical solution based upon calculus differential equations, and the universal law of gravitation.

However, modern celestial mechanics applications often require orbit precision which exceeds that obtained by the simplified Keplerian motion. For example, satellites in near Earth orbit are subject to a variety of gravitational attractions besides the point mass attraction of Earth. Since the Earth's shape is not perfectly spherical, but rather more nearly that of an oblate ellipsoid, there is more mass along the equator than there is along the polar regions. This flattening out of the Earth is the cause of various orbit perturbations and precessions. However, orbit precession does not necessarily have to be a mission design nuisance, it can act in our favor. It is possible to set up a satellite orbit inclination angle such that the orbit precesses at the same rate as the Earth is traveling about the Sun. These orbits are called Sun-synchronous orbits and can, for example, allow the spacecraft to remain in continuous sunlight. Besides the non-spherical shape of Earth, its non-homogeneous mass distribution further leads to small irregular variations in the Earth's gravity field. This chapter develops methods to express the gravity potential field about an arbitrary finite

body. A typical starting point is to write the actual gravity potential V as

$$V(\boldsymbol{r}) = V_0(\boldsymbol{r}) - R(\boldsymbol{r}) \tag{10.1}$$

where $V_0(\boldsymbol{r})$ is the reference potential, typically the dominant point mass potential and $R(\boldsymbol{r})$ is the potential function due to all other variations from the spherical homogenous idealization of the Earth's mass distribution. More generally, $R(\boldsymbol{r})$ could represent all conservative force potentials other then $V_0(\boldsymbol{r})$. This could include, for example, attractions by other solar system bodies such as the Sun, Moon, Jupiter, and so on, because it is also obviously possible for the ideal Keplerian gravity field to be significantly perturbed by the presence of additional celestial bodies. Which perturbations must be included depends on many things, but most strongly upon: (i) where is the spacecraft relative to the other solar bodies, (ii) what level of precision is sought. For example, satellites in higher Earth orbit are obviously perturbed by the gravitational attraction of the Moon and the Sun. The motion of Jupiter's moons is affected by their mutual presence. Beyond the two-body theory, unfortunately there exists no general analytical theory to describe the orbits of a multi-body system. Some special three-body solutions were discussed in the previous chapter. However, clearly one can frequently treat the Moon or the Earth as a single body and ignore the presence of other planets and moons in the solar system if one is close enough to a particular body. The last section of this chapter outlines a method to compute the gravitational *sphere of influence* of a celestial body. For example, this theory provides measures of the regions in which satellite motion is dominated by Earth's gravity field, and alternatively, the regions in which motion is dominated by either the gravity field of the Moon or the Sun.

10.1 Gravitational Potential of Finite Bodies

Assume that a celestial body has an arbitrary shape and composition as shown in Figure 10.1. The coordinate system $\mathcal{C} : \{\hat{\imath}_\xi, \hat{\imath}_\eta, \hat{\imath}_\zeta\}$ is fixed with the body. Note that its coordinate origin is not necessarily fixed to the center of mass at this point. We would like to determine the gravitational potential that a one would experience at an arbitrary point P outside the body. At its worst case, the true body may be of arbitrary shape such as is the case when attempting to orbit about a comet or asteroid. However, we will see that the mathematics greatly simplifies when we can approximate the body as being axially symmetric.

Any finite body can be considered to be the sum of an infinite number of infinitesimal mass components dm. Since each dm is infinitesimal, it produces an elementary point differential mass gravitational field. The gravitation field dV that is experienced at point P due to the differential mass dm is then given by

$$dV = -\frac{G\ dm}{s} \tag{10.2}$$

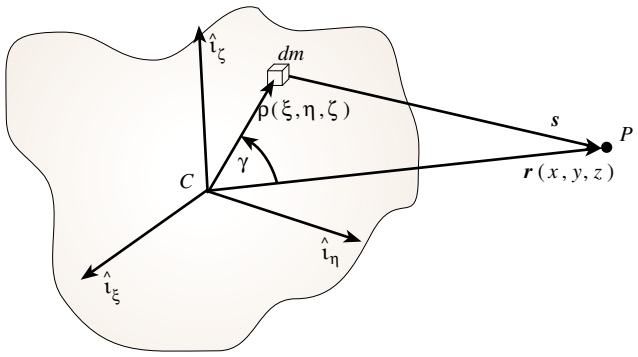


Figure 10.1: Gravity Potential of an Arbitrary Body using Cartesian Coordinates

where G is the universal gravitational constant and s is the magnitude of the relative position vector between dm and P. Note that this gravity potential definition yields a gravity potential per unit mass. Thus, taking the gradient of V, as is done in Eq. (2.5), will yield the acceleration a body will experience at point P. To obtain the gravitational potential energy expression between the mass dm and a mass at point P as discussed in chapter 2, we would use Eq. (2.6). Using the gradient of the gravitational potential function shown in Eq. (10.2) as the inertial acceleration implicitly assumes that the gravitation field of body $\mathcal{B}$ is not affected by the small bodies moving in its vicinity. While using the relative position vector $\boldsymbol{s}$ to express the infinitesimal potential field of dm, it is necessary to integrate this result to yield the total gravitation field. The integration is facilitated if we express all vectors in the body fixed frame $\mathcal{C}$. Thus the relative distance $\boldsymbol{s}$ is written as

$$\boldsymbol{s} = \boldsymbol{r} - \boldsymbol{\rho} \tag{10.3}$$

Note $s^2 = \boldsymbol{s}\cdot\boldsymbol{s} = r^2 + \rho^2 - 2\boldsymbol{\rho}\cdot\boldsymbol{r}$, leading to the law of cosines; the scalar relative distance s is then expressed as

$$s = r\left(1 + \left(\frac{\rho}{r}\right)^2 - 2\left(\frac{\rho}{r}\right)\cos\gamma\right)^{1/2} \tag{10.4}$$

Substituting Eq. (10.4) into Eq. (10.2), the gravitational potential of dm is expressed as

$$dV(\boldsymbol{r},\boldsymbol{\rho},\gamma,dm) = -\frac{G\,dm}{r\left(1 + \left(\frac{\rho}{r}\right)^2 - 2\left(\frac{\rho}{r}\right)\cos\gamma\right)^{1/2}} \tag{10.5}$$

Before we attempt to integrate the gravitational potential field of the entire body, we digress to the introduce a convenient definition of the Legendre polynomials $P_k(\nu)$. Expanding $(1 - 2\nu x + x^2)^{-1/2}$ using the binomial theorem, and

collecting on x^k, we are led to

$$(1 - 2\nu x + x^2)^{-1/2} = \sum_{k=0}^{\infty} P_k(\nu) x^k \tag{10.6}$$

Note that the infinite series will only converge if $|x| < 1$. Legendre polynomials are a classical set of orthogonal polynomials which can be computed recursively. From Eq. (10.6), it is easy to verify that the first four Legendre polynomials are given by

$$P_0(\nu) = 1 \tag{10.7a}$$
$$P_1(\nu) = \nu \tag{10.7b}$$
$$P_2(\nu) = (3\nu^2 - 1)/2 \tag{10.7c}$$
$$P_3(\nu) = (5\nu^3 - 3\nu)/2 \tag{10.7d}$$

The higher order Legendre polynomials can be obtained using the recursive formula

$$P_{n+1}(\nu) = \frac{2n+1}{n+1}\nu P_n(\nu) - \frac{n}{n+1} P_{n-1}(\nu) \tag{10.8}$$

These polynomials also satisfy the convenient zero mean condition

$$\int_{-1}^{1} P_k(\nu) d\nu = 0 \tag{10.9}$$

as well as the orthogonality condition

$$\int_{-1}^{1} P_j(\nu) P_k(\nu) d\nu = 0 \qquad \text{for } j \neq k \tag{10.10}$$

Using the Legendre polynomial identity in Eq. (10.6), the gravitational potential of dm is rewritten as the infinite sum

$$dV(\boldsymbol{r}, \boldsymbol{\rho}, \gamma, dm) = -\frac{G\, dm}{r} \sum_{k=0}^{\infty} \left(\frac{\rho}{r}\right)^k P_k(\cos\gamma) \tag{10.11}$$

Note that we are assuming here that ρ/r is less than one that the point of interest $\boldsymbol{r}$ is outside of the body $\mathcal{B}$. Finally, integrating over the entire body, we eliminate dependence on $\boldsymbol{\rho}$, γ and dm and obtain a general solution of the gravitation potential field of an arbitrary body $\mathcal{B}$.

$$V(\boldsymbol{r}) = -\frac{G\, m}{r} - \frac{G}{r} \sum_{k=1}^{\infty} \iiint_{\mathcal{B}} \left(\frac{\rho}{r}\right)^k P_k(\cos\gamma) dm \tag{10.12}$$

Note that this solution is valid for a body of arbitrary shape and arbitrary density variation. The only restriction applied so far is that the coordinate

system $\mathcal{C}$ is fixed in the body. An immediate benefit of using the orthogonal Legendre polynomials in this infinite series expression is that they allow us to break down the gravity field components as a series of successively less relevant (typically) contributions. Since $\rho/r < 1$, the contribution of the k-th element is multiplied by $(\rho/r)^k$ and goes to zero as k grows infinitely large. It is evident that as $r \to \infty$, the Gm/r term dominates Eq. (10.12). Therefore an arbitrary body's potential approaches the point mass potential as $r \to \infty$. This partitioning of the gravity potential field function is used extensively in the following approximation due to MacCullagh.

10.2 MacCullagh's Approximation

In the following discussion we only consider an approximation to the gravitational potential field model introduced by James MacCullagh (1809-1847), a professor of mathematics and natural philosophy at the Trinity College in Dublin, Ireland. This approximation involves retaining only the first three terms of the infinite series expression of the gravitational potential field expression. Substituting the first three Legendre polynomial definitions in Eq. (10.7) into Eq. (10.12) yields

$$V(\boldsymbol{r}) \approx \underbrace{-\frac{G\,m}{r}}_{\text{term 1}} \underbrace{-\frac{G}{r^2}\iiint_{\mathcal{B}} \rho\,\cos\gamma\;dm}_{\text{term 2}} \underbrace{-\frac{G}{2r^3}\iiint_{\mathcal{B}} \rho^2(3\cos\gamma^2 - 1)\;dm}_{\text{term 3}} \tag{10.13}$$

Note that term 1 is simply the point mass contribution of the gravitational potential field. This assumes that the body generating the gravitation field can be modeled as a point with mass m.

Next we investigate term 2. Since the orientation of the body fixed coordinate system $\mathcal{C}$ is arbitrary, we chose to let the ξ axis be aligned with the point P position vector $\boldsymbol{r}$. In this case $\xi = \rho\cos\gamma$ and term 2 is rewritten as

$$\text{term 2} = \frac{G}{r^2}\iiint_{\mathcal{B}} \xi\;dm \tag{10.14}$$

If the coordinate system origin of $\mathcal{C}$ is constrained to be the body center of mass (see Eq. (2.77)), then term 2 is identified to be proportional to the first mass moment of the body $\mathcal{B}$ and is thus zero.

$$\text{term 2} = 0 \tag{10.15}$$

A more subtle truth is that when leaving term 2 out of the gravity model which is fit to precisely measured orbits, the coordinate origin has implicitly been chosen at the mass center.

Before investigating term 3, the following body $\mathcal{B}$ inertia definitions are in-

troduced:

$$I_{\xi\xi} = \iiint_{\mathcal{B}} (\eta^2 + \zeta^2)\ dm \tag{10.16a}$$

$$I_{\eta\eta} = \iiint_{\mathcal{B}} (\xi^2 + \zeta^2)\ dm \tag{10.16b}$$

$$I_{\zeta\zeta} = \iiint_{\mathcal{B}} (\xi^2 + \eta^2)\ dm \tag{10.16c}$$

A convenient identity is that the sum of I_{ii} is the trace of the inertia matrix:

$$I_{\xi\xi} + I_{\eta\eta} + I_{\zeta\zeta} = 2\iiint_{\mathcal{B}} \rho^2\ dm \tag{10.17}$$

Further, the polar inertia of body $\mathcal{B}$ about the vector $\boldsymbol{r}$ is

$$I_{\boldsymbol{r}} \equiv \iiint_{\mathcal{B}} \rho^2 \sin^2\gamma\ dm \tag{10.18}$$

Using the trigonometric identity $\cos^2\gamma = 1 - \sin^2\gamma$, the term 3 expression is rewritten as

$$\text{term 3} = \frac{G}{2} r^3 \iiint_{\mathcal{B}} (2\rho^2\ dm - 3\rho^2 \sin^2\gamma\ dm) \tag{10.19}$$

Making use of the inertia definitions in Eqs. (10.16) and (10.18), term 3 is expressed as the elegant form

$$\text{term 3} = \frac{G}{2r^3}(I_{\xi\xi} + I_{\eta\eta} + I_{\zeta\zeta} - 3I_{\boldsymbol{r}}) \tag{10.20}$$

Thus, the gravitational potential field of an arbitrary body with a body fixed coordinate system centered at the body center of mass is approximated as

$$V(\boldsymbol{r}) \approx -\frac{G\ m}{r} - \frac{G}{2r^3}(I_{\xi\xi} + I_{\eta\eta} + I_{\zeta\zeta} - 3I_{\boldsymbol{r}}) + \ldots \tag{10.21}$$

Note that if the body $\mathcal{B}$ is spherically symmetric, then $I_{\xi\xi} = I_{\eta\eta} = I_{\zeta\zeta} = I_{\boldsymbol{r}}$ and term 3 vanishes. This result indicates that the gravitational potential field of a perfectly spherical object can always be modelled as that of a point mass. In fact, the expression of term 3 provides a measure of asymmetry of body $\mathcal{B}$ from the ideal spherically symmetric case.

In its current form, however, the gravitational potential function approximation is not very convenient since $I_{\boldsymbol{r}}$ depends on the location of the external reference point P. To avoid this, we write $I_{\boldsymbol{r}}$ as

$$I_{\boldsymbol{r}} = \iiint_{\mathcal{B}} \rho^2 \sin^2\gamma\ dm = \iiint_{\mathcal{B}} \rho^2(1 - \cos^2\gamma)\ dm \tag{10.22}$$

In terms of $\mathcal{C}$ coordinate frame components, let $\boldsymbol{r} = (x, y, z)^T$ and $\boldsymbol{\rho} = (\xi, \eta, \zeta)^T$. Note then that $\cos\gamma$ can be written as

$$\cos\gamma = \frac{\boldsymbol{\rho}\cdot\boldsymbol{r}}{\rho\ r} = \frac{\xi x + \eta y + \zeta z}{\rho r} \tag{10.23}$$

After substituting Eq. (10.23) into Eq. (10.22) and expanding we obtain

$$I_{\boldsymbol{r}} = \frac{1}{r^2}\iiint_{\mathcal{B}}\Big[x^2(\eta^2+\xi^2) + y^2(\xi^2+\zeta^2) + z^2(\xi^2+\eta^2) - 2xy\xi\eta - 2xz\xi\zeta - 2yz\eta\zeta\Big]dm \tag{10.24}$$

Noting that x, y and z can be taken outside of the body integral, and making use of the cross product of inertia definitions

$$I_{\xi\eta} = -\iiint_{\mathcal{B}}\xi\eta dm \tag{10.25a}$$

$$I_{\xi\zeta} = -\iiint_{\mathcal{B}}\xi\zeta dm \tag{10.25b}$$

$$I_{\eta\zeta} = -\iiint_{\mathcal{B}}\eta\zeta dm \tag{10.25c}$$

the polar inertia of body $\mathcal{B}$ about the axis $\boldsymbol{r}$ is written as

$$I_{\boldsymbol{r}} = \frac{1}{r^2}\Big[I_{\xi\xi}x^2 + I_{\eta\eta}y^2 + I_{\zeta\zeta}z^2 + 2(xyI_{\xi\eta} + xzI_{\xi\zeta} + yzI_{\eta\zeta}\Big] \tag{10.26}$$

The advantage of this expression of $I_{\boldsymbol{r}}$ is that it is written in terms of the inertias of $\mathcal{B}$ with only the (x, y, z) position of P left unspecified. Note that the expression in Eq. (10.22) involved computing the polar inertia for every new reference point P. To avoid having to keep track of the products of inertia, here we assume at this point that the coordinate system $\mathcal{C}$ is a principal coordinate system which diagonalizes the inertia matrix of the body $\mathcal{B}$. Substituting Eq. (10.26) into Eq. (10.21), the *MacCullagh Gravity Potential Approximation* is given by

$$V(x,y,z) = -\frac{G\ m}{r} - \frac{G}{2r^3}\left[I_{\xi\xi}\left(1-3\frac{x^2}{r^2}\right) + I_{\eta\eta}\left(1-3\frac{y^2}{r^2}\right) + I_{\zeta\zeta}\left(1-3\frac{z^2}{r^2}\right)\right] \tag{10.27}$$

The potential function $R(\boldsymbol{r})$ is then approximated as

$$R(\boldsymbol{r}) = \frac{G}{2r^3}\left[I_{\xi\xi}\left(1-3\frac{x^2}{r^2}\right) + I_{\eta\eta}\left(1-3\frac{y^2}{r^2}\right) + I_{\zeta\zeta}\left(1-3\frac{z^2}{r^2}\right)\right] \tag{10.28}$$

Note that the gravity field of the arbitrary body $\mathcal{B}$ is defined completely in terms of its mass m and its principal inertias $I_{\xi\xi}$, $I_{\eta\eta}$ and $I_{\zeta\zeta}$. Of course, in practice it

is rather difficult to obtain the precise inertias of a celestial body. However, it is possible to estimate these values by observing the natural motion of a small satellite about a larger celestial body.

To compute the acceleration $\boldsymbol{a}_p$ experienced by a small body in the vicinity of $\mathcal{B}$, we compute the negative gradient of V.

$$\boldsymbol{a}_P(x,y,z) = -\nabla V(x,y,z) = -\frac{\partial V}{\partial x}\hat{\imath}_\xi - \frac{\partial V}{\partial y}\hat{\imath}_\eta - \frac{\partial V}{\partial z}\hat{\imath}_\zeta \tag{10.29}$$

Making use of the fact that

$$\nabla\left(\frac{1}{r}\right) = -\frac{1}{r^2}\hat{\imath}_r \tag{10.30}$$

where $\hat{\imath}_r$ is the unit directional vector $\boldsymbol{r}$, the acceleration $\boldsymbol{a}$ is written as

$$\boldsymbol{a}_P = -\frac{Gm}{r^2}\hat{\imath}_r - \frac{3G}{2r^4}\left(I_{\xi\xi} + I_{\eta\eta} + I_{\zeta\zeta} - 5\left(I_{\xi\xi}\frac{x^2}{r^2} + I_{\eta\eta}\frac{y^2}{r^2} + I_{\zeta\zeta}\frac{z^2}{r^2}\right)\right)\hat{\imath}_r - \frac{3G}{r^5}\left(I_{\xi\xi}x\hat{\imath}_\xi + I_{\eta\eta}y\hat{\imath}_\eta + I_{\zeta\zeta}z\hat{\imath}_\zeta\right) \tag{10.31}$$

Substituting Eq. (10.26), we are able to write the acceleration in terms of principal inertias and polar inertia I_r.

$$\boldsymbol{a}_P = -\frac{Gm}{r^2}\hat{\imath}_r - \frac{3G}{2r^4}\left(I_{\xi\xi} + I_{\eta\eta} + I_{\zeta\zeta} - 5I_r\right)\hat{\imath}_r - \frac{3G}{r^5}\left(I_{\xi\xi}x\hat{\imath}_\xi + I_{\eta\eta}y\hat{\imath}_\eta + I_{\zeta\zeta}z\hat{\imath}_\zeta\right) \tag{10.32}$$

Note that the last expression of either Eqs. (10.31) or (10.32) is the non-radial perturbing acceleration. However, if $I_{\xi\xi} = I_{\eta\eta} = I_{\zeta\zeta}$, then the second and third terms combine to yield zero. This reiterates that the gravitational potential field of a spherical body of mass m is identical to that of a particle of mass m.

10.3 Spherical Harmonic Gravity Potential

While MacCullagh's approximation defines the gravity potential field in terms of the body inertias, an alternate approach is needed to model the general gravity field in terms of a spherical harmonic series. Instead of writing the infinitesimal body mass position components as Cartesian coordinates, the corresponding position vector is now expressed in terms of spherical coordinates as $\boldsymbol{\rho} = \boldsymbol{\rho}(\rho, \lambda, \beta)$ as shown in Figure 10.2. Analogously, the position vector of P is written as $\boldsymbol{r} = \boldsymbol{r}(r, \theta, \phi)$. The angle γ is the angle between the $\boldsymbol{\rho}$ and $\boldsymbol{r}$ position vectors as in the previous development.

Recall the general gravity potential field expression in Eq. (10.12) for an arbitrary body $\mathcal{B}$.

$$V(\boldsymbol{r}) = -\frac{G\ m}{r} - \frac{G}{r}\sum_{k=1}^{\infty}\iiint_{\mathcal{B}}\left(\frac{\rho}{r}\right)^k P_k(\cos\gamma)\ dm \tag{10.33}$$

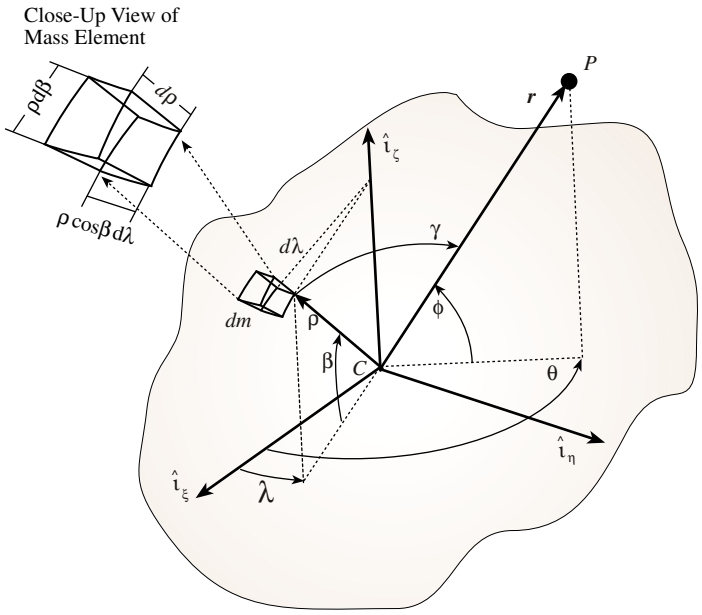


Figure 10.2: Gravity Potential of an Arbitrary Body using Spherical Coordinates

Using the body-fixed spherical coordinates ρ, λ and β, the differential mass element dm is expressed as

$$dm = D(\rho, \lambda, \beta)\rho^2 \cos\beta \; d\rho \; d\beta \; d\lambda \tag{10.34}$$

with $D = D(\rho, \lambda, \beta)$ being the local density of $\mathcal{B}$. Since the gravity potential is expressed as an infinite expansion using Legendre polynomials depending on $\cos\gamma$, it is convenient to express $\cos\gamma$ as a function of the angular coordinates of the $\boldsymbol{\rho}$ and $\boldsymbol{r}$ position vectors. Using the spherical trigonometric law of cosines we write

$$\cos\gamma = \sin\phi\sin\beta + \cos\phi\cos\beta\cos(\theta - \lambda) \tag{10.35}$$

To further facilitate the development of the gravitational spherical harmonic series, we make use of the definition of the *associated Legendre Functions.*[1]

$$P_k^j(\nu) = (1-\nu^2)^{\frac{1}{2}j} \frac{d^j}{d\nu^j}\left(P_k(\nu)\right) \tag{10.36}$$

The parameter j is referred to as the order of the associated Legendre function, while k is referred to as the degree. Note that zeroth order associated Legendre functions are simply the Legendre polynomials.

$$P_k^0(\nu) \equiv P_k(\nu) \tag{10.37}$$

Further, since $P_k(\nu)$ is a polynomial expression of degree k, then

$$P_k^j = 0 \qquad \forall\ j > k$$

must be true. The associated Legendre functions for Legendre polynomials up to third degree as explicitly given as:

$$P_1^1(\nu) = \sqrt{1-\nu^2} \qquad P_2^1(\nu) = 3\nu\sqrt{1-\nu^2} \qquad P_3^1(\nu) = \frac{3}{2}\sqrt{1-\nu^2}(5\nu^2-1) \tag{10.38a}$$

$$P_2^2(\nu) = 3(1-\nu^2) \qquad P_3^2(\nu) = 15\nu(1-\nu^2) \tag{10.38b}$$

$$P_3^3(\nu) = 15(1-\nu^2)^{\frac{3}{2}} \tag{10.38c}$$

For our present study we set $\nu = \sin\alpha$. The corresponding associated Legendre functions up to second degree are

$$P_1^1(\sin\alpha) = \cos\alpha \tag{10.39a}$$

$$P_2^1(\sin\alpha) = 3\sin\alpha\cos\alpha \tag{10.39b}$$

$$P_2^2(\sin\alpha) = 3\cos^2\alpha \tag{10.39c}$$

Notice that the first Legendre polynomial can now be written in terms of associated Legendre functions by making use of the spherical trigonometric identity in Eq. (10.35).

$$P_1(\cos\gamma) = \cos\gamma = P_1(\sin\phi)P_1(\sin\beta) + P_1^1(\sin\phi)P_1^1(\sin\beta)\cos(\theta-\lambda) \tag{10.40}$$

To write the second Legendre polynomial in terms of associated Legendre functions, we first express $\cos^2\gamma$ as

$$\begin{aligned}\cos^2\gamma = {}& \frac{1}{2}\Big(3\sin^2\phi\sin^2\beta - \sin^2\phi - \sin^2\beta + 1\Big) \\ & + \frac{1}{2}\cos^2\phi\cos^2\beta\cos\left(2(\theta-\lambda)\right) + 2\sin\phi\cos\phi\sin\beta\cos\beta\cos(\theta-\lambda)\end{aligned} \tag{10.41}$$

Using Eqs. (10.39) and (10.41), the second Legendre polynomial is expressed as

$$\begin{aligned}P_2(\cos\gamma) = \frac{1}{2}(3\cos^2\gamma - 1) = {}& P_2(\sin\phi)P_2(\sin\beta) \\ & + \frac{1}{3}P_2^1(\sin\phi)P_2^1(\sin\beta)\cos(\theta-\lambda) \\ & + \frac{1}{12}P_2^2(\sin\phi)P_2^2(\sin\beta)\cos(2(\theta-\lambda))\end{aligned} \tag{10.42}$$

Carrying these expansion to higher degrees leads to the *general addition theorem for Legendre polynomials.*[1]

$$\begin{aligned}P_k(\cos\gamma) = {}& P_k(\sin\phi)P_k(\sin\beta) \\ & + 2\sum_{j=1}^{k}\frac{(k-j)!}{(k+j)!}P_k^j(\sin\phi)P_k^j(\sin\beta)\cos(j(\theta-\lambda))\end{aligned} \tag{10.43}$$

Substituting Eq. (10.43) into Eq. (10.33), and making use of the trigonometric identity

$$\cos(j(\theta-\lambda)) = \cos j\theta\ \cos j\lambda + \sin j\theta\ \sin j\lambda$$

the gravitational potential field is expressed in terms of spherical coordinates as

$$V(r,\phi,\theta) = -\frac{G\ m}{r} - \sum_{k=1}^{\infty} \frac{1}{r^{k+1}} \Big(A_k P_k(\sin\phi) + \sum_{j=1}^{k} P_k^j(\sin\phi) \Big(B_k^j \cos j\theta + C_k^j \sin j\theta \Big) \Big) \quad (10.44)$$

with the coefficients A_k (zonal harmonics) and B_k^j and C_k^j (sectorial harmonics) defined as the following $(k+2)$-th degree mass moments

$$A_k = G \iiint_{\mathcal{B}} \rho^{k+2} D(\rho,\lambda,\beta) P_k(\sin\beta) \cos\beta\ d\rho\ d\beta\ d\lambda \quad (10.45a)$$

$$B_k^j = 2G \frac{(k-j)!}{(k+j)!} \iiint_{\mathcal{B}} \rho^{k+2} D(\rho,\lambda,\beta) P_k^j(\sin\beta) \cos j\lambda \cos\beta\ d\rho\ d\beta\ d\lambda \quad (10.45b)$$

$$C_k^j = 2G \frac{(k-j)!}{(k+j)!} \iiint_{\mathcal{B}} \rho^{k+2} D(\rho,\lambda,\beta) P_k^j(\sin\beta) \sin j\lambda \cos\beta\ d\rho\ d\beta\ d\lambda \quad (10.45c)$$

Note that the gravity potential field expression in Eq. (10.44) separates the position coordinates which depend on the body $\mathcal{B}$ mass distribution (i.e. A_k, B_k^j and C_k^j) and the those which depend on the location of P (i.e. ϕ and θ).

Consider the body integrals of Eq. (10.45). If these integrals vanish, obviously the gravitational potential field expression in Eq. (10.44) reduces to $-Gm/r$. This occurs if the density is a function of ρ only; or, in other words, when the body $\mathcal{B}$ is spherically symmetric with $D = D(\rho)$. To establish this result, let us first investigate the integrals for the common case where the body is assumed to have an axis of symmetry. We chose this axis to be $\hat{\imath}_\zeta$. This means that the density $D(\rho,\lambda,\beta) = D(\rho,\beta)$ does not depend on the longitude λ. Evaluating the λ integral first, B_k^j and C_k^j are expressed as

$$B_k^j = 2G \frac{(k-j)!)}{(k+j)!} \int_0^R \int_{-\pi/2}^{\pi/2} \rho^{k+2} D(\rho,\beta) P_k^j(\sin\beta) \cos\beta\, d\beta\, d\rho \left(\int_0^{2\pi} \cos(j\lambda) d\lambda \right) \quad (10.46)$$

$$C_k^j = 2G \frac{(k-j)!)}{(k+j)!} \int_0^R \int_{-\pi/2}^{\pi/2} \rho^{k+2} D(\rho,\beta) P_k^j(\sin\beta) \cos\beta\, d\beta\, d\rho \left(\int_0^{2\pi} \sin(j\lambda) d\lambda \right) \quad (10.47)$$

where R is the outer most radius of the body $\mathcal{B}$. Since both λ integrals are zero, we come to the conclusion that

$$B_k^j = C_k^j = 0 \qquad \forall\ D = D(\rho,\beta) \quad (10.48)$$

for any body of revolution (i.e. cylinders, ellipsoids, etc.). Thus B_k^j and C_k^j are only non-zero if the body density depends on the longitude λ.

Investigating the integral A_k, we assume at this point that the body $\mathcal{B}$ is spherically symmetric with $D = D(\rho)$. Performing the longitude integration first, we are able to write A_k as

$$A_k = 2\pi G \left(\int_0^R \rho^{k+2} D(\rho) d\rho \right) \left(\int_{-\pi/2}^{\pi/2} P_k(\sin\beta) \cos\beta d\beta \right) \tag{10.49}$$

Using the variable $\nu = \sin\beta$, this integral is rewritten as

$$A_k = 2\pi G \left(\int_0^R \rho^{k+2} D(\rho) d\rho \right) \left(\int_{-1}^{1} P_k(\nu) d\nu \right) \tag{10.50}$$

Using the Legendre polynomial zero mean condition in Eq. (10.9), the integrals A_k are found to be zero for the spherically symmetric body case.

$$A_k = B_k^j = C_k^j = 0 \qquad \forall\ D = D(\rho) \tag{10.51}$$

Example 10.1: Imagine a planetary body which is hollow. In this example we investigate what the gravity potential field will look like if a person is inside this planetary body. The gravitational potential field of differential mass element dm is given in Eq. (10.11) for the case where $r > \rho$ and the point of interest $\boldsymbol{r}$ is entirely outside of the planetary body. If the point of interest is inside a hollow body and $r < \rho$, then we rewrite the differential potential field expression in Eq. (10.5) as

$$dV(\boldsymbol{r}, \boldsymbol{\rho}, \gamma, dm) = -\frac{G\ dm}{\rho \left(1 + \left(\frac{r}{\rho}\right)^2 - 2\left(\frac{r}{\rho}\right)\cos\gamma \right)^{1/2}} \tag{10.52}$$

Since $r/\rho < 1$, we are able to use the Legendre polynomial identity in Eq. (10.6) to express the differential potential field of dm as an infinite series.

$$dV(\boldsymbol{r}, \boldsymbol{\rho}, \gamma, dm) = -\frac{G\ dm}{\rho} \sum_{k=0}^{\infty} \left(\frac{r}{\rho}\right)^k P_k(\cos\gamma) \tag{10.53}$$

Integrating over the entire body $\mathcal{B}$, the gravitational potential field $V(\boldsymbol{r})$ is given by

$$V(\boldsymbol{r}) = \underbrace{-\iiint_{\mathcal{B}} \frac{G}{\rho} dm}_{V_0} - G \sum_{k=1}^{\infty} \iiint_{\mathcal{B}} \frac{1}{\rho} \left(\frac{r}{\rho}\right)^k P_k(\cos\gamma) dm \tag{10.54}$$

where the point of interest $\boldsymbol{r}$ interior to the hollow body. For the standard case where $r > \rho$ the term V_0 becomes the standard inverse-square gravitational attraction of a point mass. Note that for the present case where $r < \rho$ this term becomes a constant of the body $\mathcal{B}$. The function V_0 only depends on

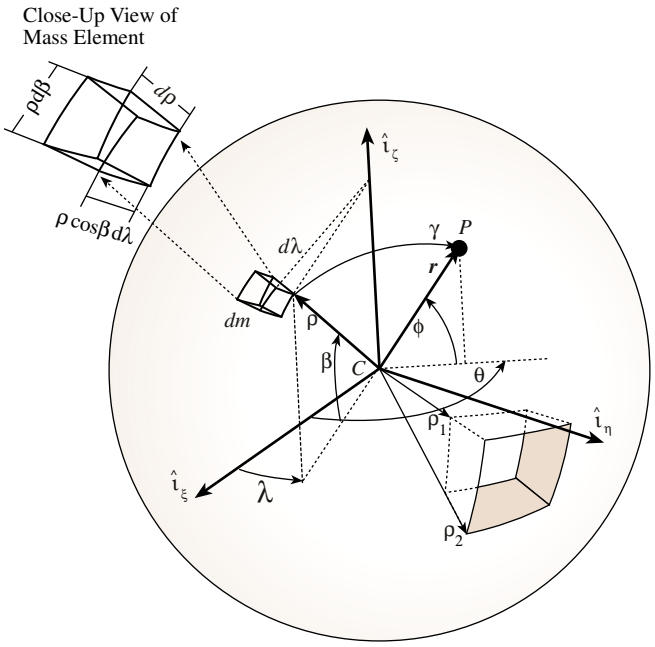


Figure 10.3: Illustration of a Spherically Symmetric Shell-Shaped Planetoid

the body element position vector $\boldsymbol{\rho}$ and not on the position vector $\boldsymbol{r}$. Since $V_0 = V_0(\boldsymbol{\rho})$ only, the gradient of V_0 with respect to $\boldsymbol{r}$ will always be zero for this case and V_0 will not contribute any gravitational acceleration for any point $\boldsymbol{r}$ interior of the hollow body $\mathcal{B}$.

Let us use the Legendre polynomial identity of $P_k(\cos\gamma)$ in Eq. (10.43). This allows us to write the general potential field expression of an interior point of a hollow body using spherical harmonics.

$$V(r,\phi,\theta) = V_0 - \sum_{k=1}^{\infty} r^k \Big(A_k P_k(\sin\phi) + \sum_{j=1}^{k} P_k^j(\sin\phi)\Big(B_k^j \cos j\theta + C_k^j \sin j\theta\Big)\Big) \quad (10.55)$$

with the body specific components A_k, B_k^j and C_k^j begin defined as

$$A_k = G\iiint_{\mathcal{B}} \frac{1}{\rho^{k-1}} D(\rho,\lambda,\beta) P_k(\sin\beta)\cos\beta \; d\rho \; d\beta \; d\lambda \quad (10.56a)$$

$$B_k^j = G_{kj}\iiint_{\mathcal{B}} \frac{1}{\rho^{k-1}} D(\rho,\lambda,\beta) P_k^j(\sin\beta)\cos j\lambda\cos\beta \; d\rho \; d\beta \; d\lambda \quad (10.56b)$$

$$C_k^j = G_{kj}\iiint_{\mathcal{B}} \frac{1}{\rho^{k-1}} D(\rho,\lambda,\beta) P_k^j(\sin\beta)\sin j\lambda\cos\beta \; d\rho \; d\beta \; d\lambda \quad (10.56c)$$

with G_{kj} being defined as

$$G_{kj} = 2G\frac{(k-j)!}{(k+j)!} \tag{10.57}$$

Equivalent arguments regarding A_k, B_k^j and C_k^j can be made as were done for the case where $r > \rho$. If the density function satisfies $D = D(\rho, \beta)$, then B_k^j and C_k^j will always be zero. This density function condition applies to any body of revolution. Ellipsoids and sphere would be examples of such bodies. Further, if the density function satisfies $D = D(\rho)$, then the gravitational parameter A_k will also be zero. An example of this would be a spherical shell as shown in Figure 10.3. Note that in this case the gravitational potential field $V(\boldsymbol{r})$ experienced at any point $\boldsymbol{r}$ in the interior of this shell will be the constant V_0. Since the gravitational acceleration is the gradient of the potential field function with respect to $\boldsymbol{r}$, the gravitational acceleration will be zero for any interior point. Assuming the spherically symmetric shell shown in Figure 10.3 has a homogenous density D, the constant gravitational interior potential V_0 is

$$V_0 = -GD2\pi\left(\rho_2^2 - \rho_1^2\right) \tag{10.58}$$

A spacecraft inside this shell planetoid will be attracted to the planetoid mass. However, the various gravitational forces will cancel each other at *any* interior point due to the symmetry of the planetoid body.

If the origin of the coordinate system $\mathcal{C}$ is the body mass center, it can be shown that A_1 will always be zero. To do so, we substitute the dm definition in Eq. (10.34) into Eq. (10.45a) with k set to 1.

$$A_1 = G\iiint_{\mathcal{B}} \rho \sin\beta \; dm \tag{10.59}$$

Using analogous arguments that lead to Eq. (10.14) being zero, it is clear that A_1 is always zero if the coordinate system origin is at the mass center.

For a body with rotational symmetry, the gravitational potential field function V is then expressed as the sum of the point mass contribution and the zonal harmonics.

$$V(r,\phi) = -\frac{G\ m}{r} - \sum_{k=2}^{\infty} \frac{1}{r^{k+1}} A_k P_k(\sin\phi) \tag{10.60}$$

Note that the rotation of this body about the symmetry axis does not affect the gravity field. The physical reason is that the mass distribution is not changed when a symmetric body is rotated about it's symmetry axis. As a consequence, if rotational symmetry is assumed, Earth rotation does not change the gravity field and this greatly simplifies the equations of motion.

The conventional notation in the orbital mechanics literature for the zonal harmonics is

$$J_k = -\frac{A_k}{r_{eq}^k} \tag{10.61}$$

with r_{eq} being the equatorial radius of body $\mathcal{B}$. The gravity potential V now takes on its most famous for:

$$V(r,\phi) = -\frac{G\ m}{r}\left[1 - \sum_{k=2}^{\infty}\left(\frac{r_{eq}}{r}\right)^k J_k P_k(\sin\phi)\right] \tag{10.62}$$

Note that as the point P moves away from the body $\mathcal{B}$ (i.e. $r \to \infty$), the effect of the zonal harmonics diminish to zero quickly.

Practically speaking, it is essentially impossible to obtain values for J_k by integration of its analytical expression, because we do not have an accurate knowledge of the Earth's mass distribution $D(\rho,\lambda,\beta)$. Instead, these coefficients are typically obtained by observing the motion of a satellite about the body and then extracting these harmonics through an estimation method. For the Earth, the first six zonal harmonics are given by[1, 2]

$$\begin{aligned} J_2 &= 1082.63\ 10^{-6} \\ J_3 &= -2.52\ 10^{-6} \\ J_4 &= -1.61\ 10^{-6} \\ J_5 &= -0.15\ 10^{-6} \\ J_6 &= \ \ 0.57\ 10^{-6} \end{aligned}$$

The J_2 harmonic, also referred to as the oblateness perturbation, is clearly the dominant harmonic. It causes a highly noticeable precession of the near-Earth satellite orbits.

Setting $\mu = G\ m$, the gravitational perturbation function $R(\boldsymbol{r})$ for the J_2 through J_6 gravitational perturbations is then given by

$$\begin{aligned} R(\boldsymbol{r}) = &-\frac{J_2}{2}\frac{\mu}{r}\left(\frac{r_{eq}}{r}\right)^2\left(3\sin^2\phi - 1\right) \\ &-\frac{J_3}{2}\frac{\mu}{r}\left(\frac{r_{eq}}{r}\right)^3\left(5\sin^3\phi - 3\sin\phi\right) \\ &-\frac{J_4}{8}\frac{\mu}{r}\left(\frac{r_{eq}}{r}\right)^4\left(35\sin^4\phi - 30\sin^2\phi + 3\right) \\ &-\frac{J_5}{8}\frac{\mu}{r}\left(\frac{r_{eq}}{r}\right)^5\left(63\sin^5\phi - 70\sin^3\phi + 15\sin\phi\right) \\ &-\frac{J_6}{16}\frac{\mu}{r}\left(\frac{r_{eq}}{r}\right)^6\left(231\sin^6\phi - 315\sin^4\phi + 105\sin^2\phi - 5\right) \end{aligned} \tag{10.63}$$

Computing the gradient of $R(\boldsymbol{r})$ and making use of $z/r = \sin\phi$, the perturbing acceleration $\boldsymbol{a}_{J_i}$ due to J_i is given in terms of inertial Cartesian coordinates as:[2]

$$\boldsymbol{a}_{J_2} = -\frac{3}{2}J_2\left(\frac{\mu}{r^2}\right)\left(\frac{r_{eq}}{r}\right)^2\begin{pmatrix}\left(1 - 5\left(\frac{z}{r}\right)^2\right)\frac{x}{r} \\ \left(1 - 5\left(\frac{z}{r}\right)^2\right)\frac{y}{r} \\ \left(3 - 5\left(\frac{z}{r}\right)^2\right)\frac{z}{r}\end{pmatrix} \tag{10.64}$$

$$\boldsymbol{a}_{J_3} = -\frac{1}{2}J_3\left(\frac{\mu}{r^2}\right)\left(\frac{r_{eq}}{r}\right)^3\begin{pmatrix} 5\left(7\left(\frac{z}{r}\right)^3 - 3\left(\frac{z}{r}\right)\right)\frac{x}{r} \\ 5\left(7\left(\frac{z}{r}\right)^3 - 3\left(\frac{z}{r}\right)\right)\frac{y}{r} \\ 3\left(10\left(\frac{z}{r}\right)^2 - \frac{35}{3}\left(\frac{z}{r}\right)^4 - 1\right)\end{pmatrix} \tag{10.65}$$

$$\boldsymbol{a}_{J_4} = -\frac{5}{8}J_4\left(\frac{\mu}{r^2}\right)\left(\frac{r_{eq}}{r}\right)^4\begin{pmatrix} \left(3 - 42\left(\frac{z}{r}\right)^2 + 63\left(\frac{z}{r}\right)^4\right)\frac{x}{r} \\ \left(3 - 42\left(\frac{z}{r}\right)^2 + 63\left(\frac{z}{r}\right)^4\right)\frac{y}{r} \\ -\left(15 - 70\left(\frac{z}{r}\right)^2 + 63\left(\frac{z}{r}\right)^4\right)\frac{z}{r}\end{pmatrix} \tag{10.66}$$

$$\boldsymbol{a}_{J_5} = -\frac{J_5}{8}\left(\frac{\mu}{r^2}\right)\left(\frac{r_{eq}}{r}\right)^5\begin{pmatrix} 3\left(35\left(\frac{z}{r}\right) - 210\left(\frac{z}{r}\right)^3 + 231\left(\frac{z}{r}\right)^5\right)\frac{x}{r} \\ 3\left(35\left(\frac{z}{r}\right) - 210\left(\frac{z}{r}\right)^3 + 231\left(\frac{z}{r}\right)^5\right)\frac{y}{r} \\ \left(15 - 315\left(\frac{z}{r}\right)^2 + 945\left(\frac{z}{r}\right)^4 - 693\left(\frac{z}{r}\right)^6\right)\end{pmatrix} \tag{10.67}$$

$$\boldsymbol{a}_{J_6} = \frac{J_6}{16}\left(\frac{\mu}{r^2}\right)\left(\frac{r_{eq}}{r}\right)^6\begin{pmatrix} \left(35 - 945\left(\frac{z}{r}\right)^2 + 3465\left(\frac{z}{r}\right)^4 - 3003\left(\frac{z}{r}\right)^6\right)\frac{x}{r} \\ \left(35 - 945\left(\frac{z}{r}\right)^2 + 3465\left(\frac{z}{r}\right)^4 - 3003\left(\frac{z}{r}\right)^6\right)\frac{y}{r} \\ \left(3003\left(\frac{z}{r}\right)^6 - 4851\left(\frac{z}{r}\right)^4 + 2205\left(\frac{z}{r}\right)^2 - 315\right)\frac{z}{r}\end{pmatrix} \tag{10.68}$$

Example 10.2: The J_i induced gravitational acceleration vectors in Eq. (10.64) – (10.68) are developed with components taken in a Cartesian coordinate frame $\{\hat{\imath}_x, \hat{\imath}_y, \hat{\imath}_z\}$. In this example we take the gradient of the gravitational perturbation function $R(\boldsymbol{r})$ in Eq. (10.63) using the spherical coordinates (r, ϕ, θ). Let $\mathcal{S}$ be a coordinate frame defined through $\{\hat{\imath}_r, \hat{\imath}_\theta, \hat{\imath}_\phi\}$ as shown in Figure 10.4. The acceleration experienced due to the gravitational potential

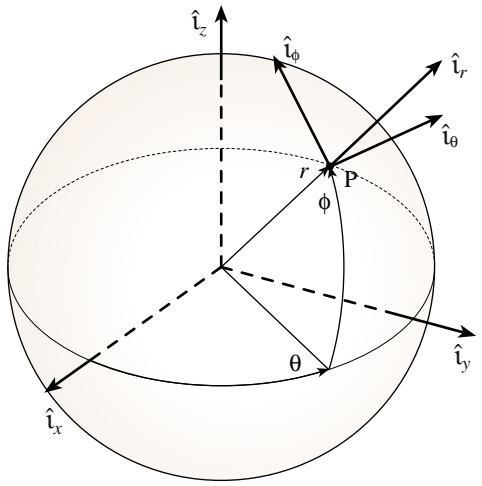


Figure 10.4: Illustration of Spherical Coordinates and the Coordinate Frame Unit Direction Vectors

R is given by

$$\boldsymbol{a} = \nabla R = \frac{\partial R}{\partial r}\hat{\imath}_r + \frac{1}{r}\frac{\partial R}{\partial \phi}\hat{\imath}_\phi + \frac{1}{r\cos\phi}\frac{\partial R}{\partial \theta}\hat{\imath}_\theta$$

Note that the pertubance potential R does not depend on the longitude angle θ, thus their will be no acceleration component in the $\hat{\imath}_\theta$ direction. The perturbation acceleration vectors $\boldsymbol{a}_{J_i}$ due to J_i are expressed in terms of the spherical coordinate frame unit direction vectors as:

$$\begin{aligned}
\boldsymbol{a}_{J_2} &= -\frac{3}{4}J_2\kappa_2\Big[(3\cos(2\phi)-1)\hat{\imath}_r + 2\sin(2\phi)\hat{\imath}_\phi\Big] \\
\boldsymbol{a}_{J_3} &= -\frac{1}{8}J_3\kappa_3\Big[(20\sin(3\phi)-12\sin\phi)\hat{\imath}_r \\
&\qquad - (15\cos(3\phi)-3\cos\phi)\hat{\imath}_\phi\Big] \\
\boldsymbol{a}_{J_4} &= -\frac{5}{64}J_4\kappa_4\Big[(-35\cos(4\phi)+20\cos(2\phi)-9)\hat{\imath}_r \\
&\qquad - (28\sin(4\phi)-8\sin(2\phi))\hat{\imath}_\phi\Big] \\
\boldsymbol{a}_{J_5} &= -\frac{3}{128}J_5\kappa_5\Big[2(-63\sin(5\phi)+35\sin(3\phi)-30\sin\phi)\hat{\imath}_r \\
&\qquad + (105\cos(5\phi)-35\cos(3\phi)+10\cos\phi)\hat{\imath}_\phi\Big] \\
\boldsymbol{a}_{J_6} &= -\frac{7}{512}J_6\kappa_6\Big[(221\cos(6\phi)-126\cos(4\phi)+105\cos(2\phi)-50)\hat{\imath}_r \\
&\qquad + (198\sin(6\phi)-72\sin(4\phi)+30\sin(2\phi))\hat{\imath}_\phi\Big]
\end{aligned}$$

with κ_i being defined as

$$\kappa_i = \left(\frac{\mu}{r^2}\right)\left(\frac{r_{eq}}{r}\right)^i$$

10.4 Multi-Body Gravitational Acceleration

In chapter 9 the equations of motion for the three body problem were discussed. It was shown that if a body has a small small mass compared to the other two bodies, then it would essentially abide by Keplerian motion in the vicinity of either remaining body. For example, the motion of a satellite in near Earth orbit is dominated by Earth's gravitational attraction. However, the Moon does cause a small perturbation of the orbits. The acceleration experienced by the spacecraft due to the Moon can be viewed as a result of a disturbing potential function. This section will provide the general equations of motion for a multi-body system of point masses and derive such a multi-body perturbative potential function.

In the following development, let m_1 be the primary mass about which a second mass m_2 is orbiting. The remaining masses m_i (with $2 < i \leq N$) are

assumed to be close or massive enough to have an effect on the (m_1, m_2) two-body solution. Using Eq. (9.3), the equations of motion for m_1, m_2 and m_i are

$$\ddot{\boldsymbol{r}}_1 = G\frac{m_2}{r_{12}^3}\boldsymbol{r}_{12} + G\sum_{j=3}^{N}\frac{m_j}{r_{1j}^3}\boldsymbol{r}_{1j} \tag{10.69}$$

$$\ddot{\boldsymbol{r}}_2 = G\frac{m_1}{r_{21}^3}\boldsymbol{r}_{21} + G\sum_{j=3}^{N}\frac{m_j}{r_{2j}^3}\boldsymbol{r}_{2j} \tag{10.70}$$

$$\ddot{\boldsymbol{r}}_i = G\sum_{j=1}^{N}\frac{m_j}{r_{ij}}\boldsymbol{r}_{ij} \qquad \text{summed for } i \neq j \tag{10.71}$$

with the relative position vector $\boldsymbol{r}_{ij}$ being defined as

$$\boldsymbol{r}_{ij} = \boldsymbol{r}_j - \boldsymbol{r}_i \tag{10.72}$$

Using the position vector property $\boldsymbol{r}_{ij} = -\boldsymbol{r}_{ji}$, the equations of motion of m_2 relative to m_1 are

$$\ddot{\boldsymbol{r}}_{12} = G\left[-\frac{m_1+m_2}{r_{12}^3}\boldsymbol{r}_{12} + \sum_{j=3}^{N} m_j\left(\frac{\boldsymbol{r}_{2j}}{r_{2j}^3} - \frac{\boldsymbol{r}_{1j}}{r_{1j}^3}\right)\right] \tag{10.73}$$

Expressing all position vectors relative to m_1, the equations of motion of m_2 are given by

$$\ddot{\boldsymbol{r}}_{12} + \frac{G(m_1+m_2)}{r_{12}^3}\boldsymbol{r}_{12} = \underbrace{G\sum_{j=3}^{N} m_j\left(\frac{\boldsymbol{r}_{1j}-\boldsymbol{r}_{12}}{r_{2j}^3} - \frac{\boldsymbol{r}_{1j}}{r_{1j}^3}\right)}_{\boldsymbol{a}_d} \tag{10.74}$$

Note that the left hand side of Eq. (10.74) is the result of the standard Keplerian motion between masses m_1 and m_2. The right hand side forms the perturbative acceleration $\boldsymbol{a}_d$ away from this solution. No assumption has been made here as to whether this acceleration is small or large. Further, note that the mass index labeling has been setup arbitrarily here. By switching the indices, the equations of motion provided in Eq. (10.74) are applicable to any of the N bodies.

Assume that the N-body system consists of the Earth (m_1), the Moon (m_2) and the Sun (m_3), and we are interesting in the motion of the Moon relative to the Earth. As seen in Table 8.2, the Sun is at least 10^6 times more massive than either the Earth or the Moon. At first glance then, it would appear as if the perturbative acceleration $\boldsymbol{a}_d$ would be very large in this case. However, since the relative distances r_{23} and r_{13} are almost identical, then the effect of the Sun's gravitational attraction on the relative motion of the Moon to the Earth is very small. In essence, the Earth-Moon system are approximately free-falling together around the Sun.

The disturbing potential function generated by mass m_j on the motion of m_2 is given by[1, 3]

$$R_j(\boldsymbol{r}_{12}) = Gm_j \left(\frac{1}{r_{2j}} - \frac{\boldsymbol{r}_{12} \cdot \boldsymbol{r}_{1j}}{r_{1j}^3} \right) \tag{10.75}$$

To verify that this potential function does lead to the previously derived perturbative acceleration $\boldsymbol{a}_d$, we need to compute the gradient of the scalar distance r_{2j}.

$$\nabla r_{2j} = \begin{pmatrix} \frac{\partial r_{2j}}{\partial x_2} \\ \frac{\partial r_{2j}}{\partial y_2} \\ \frac{\partial r_{2j}}{\partial z_2} \end{pmatrix} = -\frac{\boldsymbol{r}_{2j}}{r_{2j}} = -\frac{\boldsymbol{r}_{1j} - \boldsymbol{r}_{12}}{r_{2j}} \tag{10.76}$$

Making use of Eq. (10.76), we find that indeed

$$\boldsymbol{a}_d = \sum_{j=3}^{N} \nabla R_j(\boldsymbol{r}_{12}) \tag{10.77}$$

10.5 Spheres of Gravitational Influence

Consider the classical two-body problem where one relatively small object of mass m, such as a man-made spacecraft, is orbiting about a relatively massive object of mass m_1 such as the Earth. Depending on the relative energy of m to m_1, the relative orbit of m will be one of the three conic solutions discussed in chapter 8 (i.e. either be elliptic, parabolic or hyperbolic). If additional celestial objects are present, then the precise motion of m relative to m_1 is no longer the classical two-body solution, but rather a trajectory which is perturbed from this two-body solution. For example, consider a satellite in Earth orbit with the Moon acting as the additional celestial body. The gravitational influence of the Moon will cause slight perturbations to the nominal two-body solution of the satellite orbit. However, the closer the satellite orbit is to the Earth, the smaller the gravitational perturbation would be. Thus, for a very low Earth orbiting satellite, the lunar gravitational effect could be ignored, since the satellite motion is overwhelmingly dominated by the Earth's gravitational attraction. As the satellite travels between the Earth and Moon, however, the lunar effect must be included when determining the satellite orbit. This concept is evident mathematically in Eq. (10.74) where the equations of motion for a multi-body system are defined with the additional gravitational effects written as a disturbance to the classical two-body solution.

Treating the gravitational acceleration of additional bodies as a perturbations leads to the concept of defining a *sphere of influence* about a particular celestial body. These spheres are regions around celestial objects where the particular object's gravitational attraction will largely determine the trajectory of any other small mass within its vicinity. The gravitational attraction due to any

remaining objects would typically be very small within the sphere of influence. Thus, as an initial approximation of the multi-body problem, the gravitational effect of the remaining bodies is ignored while a small object resides within such a sphere of influence. This idea is illustrated in Figure 10.5 where a conceptual trajectory is shown of a satellite traveling from Earth to Mars. The spheres of gravitational influence of Earth and Mars are illustrated as transparent spheres around each planet. Note that the Sun also possesses a sphere of influence. However, the Sun's region of attraction is so large that it encompasses the entire solar system. Thus, if any object is outside the sphere of influence of a planet, it is by default assumed to be under the dominant gravitational influence of the Sun. The spacecraft motion shown in Figure 10.5 begins close to Earth. The craft has enough energy relative to the Earth such that it is on a hyperbolic escape orbit. This trajectory description is reasonably accurate up to the point where the craft departs the Earth's sphere of influence. From here on its motion is dominated by the gravitational influence of the Sun. Assuming the spacecraft doesn't have enough energy relative to the Sun to be on an escape trajectory, its orbit can be described by a heliocentric ellipse. At a later point of time the craft enters the sphere of influence of Mars, whose gravitational attraction will determine its orbit from here on. Since the velocity "far from Mars" is non-zero, we can be certain the motion relative to Mars will be hyperbolic.

Note that with this approximation, the spacecraft motion within a particular sphere of influence is solely described through two-body conic intersections. Since the entire trajectory of the spacecraft through the multi-body gravitational field is approximated as a series of conic solutions (i.e. locally elliptic, parabolic or hyperbolic relative orbits), this method of determining the trajectory is referred to as the *method of patched conics*. As the name describes, the various conic solutions are patched together to find an approximate solution to the multi-body problem.

To express the sphere of influence concept in mathematical terms, we rewrite the multi-body equations of motion in Eqs. (10.74). Let m_1 and m_2 be celestial bodies, while m is a spacecraft with $m \ll m_1$ and $m \ll m_2$. The position vector of m relative to m_i is given by $\boldsymbol{r}_i$, while the position vector from body i to body j is given by $\boldsymbol{r}_{ij}$ as illustrated in Figure 10.6. Defining $\mu_i = Gm_i$, the equations of motion of m relative to either m_1 or m_2 are given by

$$\ddot{\boldsymbol{r}}_1 = \underbrace{-\frac{\mu_1}{r_1^3}\boldsymbol{r}_1}_{\boldsymbol{a}_1} \underbrace{-\mu_2\left(\frac{\boldsymbol{r}_2}{r_2^3} + \frac{\boldsymbol{r}_{12}}{r_{12}^3}\right)}_{\boldsymbol{a}_{d_1}} \tag{10.78}$$

$$\ddot{\boldsymbol{r}}_2 = \underbrace{-\frac{\mu_2}{r_2^3}\boldsymbol{r}_2}_{\boldsymbol{a}_2} \underbrace{-\mu_1\left(\frac{\boldsymbol{r}_1}{r_1^3} - \frac{\boldsymbol{r}_{12}}{r_{12}^3}\right)}_{\boldsymbol{a}_{d_2}} \tag{10.79}$$

In each case, the gravitational attraction of the other celestial body is expressed as a disturbance acceleration $\boldsymbol{a}_{d_i}$ on the two-body solution about m_i. The sphere of influence is defined as the vicinity around m_i where the disturbance vector

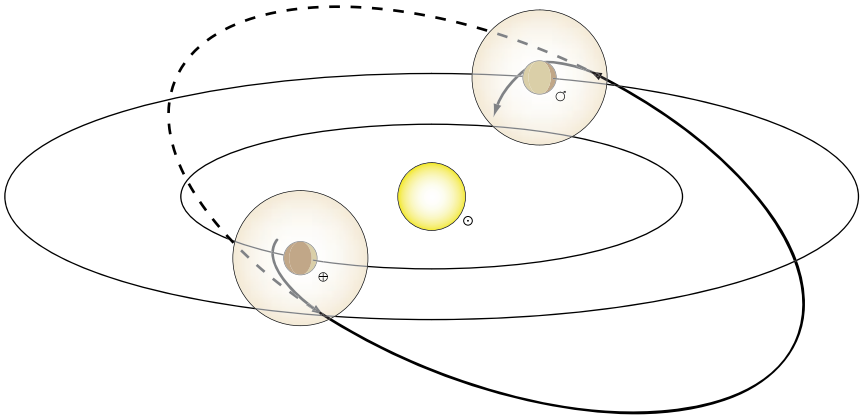


Figure 10.5: Approximating a Trajectory Among Multiple Bodies Through Spheres of Influences

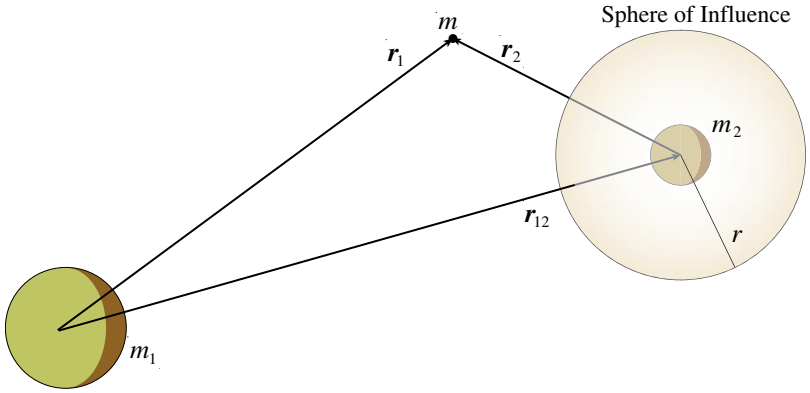


Figure 10.6: Gravitational Spheres of Influence

$\boldsymbol{a}_{d_i}$ is of the same magnitude as the two-body acceleration vector $\boldsymbol{a}_i$. Assuming that the mass m_2 is smaller than the mass m_1, this surface about m_2 can be approximated as a sphere. The radius r of this approximate spherical surface is determined through the formula[1, 3]

$$r = \left(\frac{m_2}{m_1}\right)^{2/5} r_{12} \tag{10.80}$$

Using the planetary mass coefficients in Table 8.2, the sphere of influence radii of the solar system planets are computed in Table 10.1.

Table 10.1: Spheres of Influence Radii of the Solar System Planets Relative to the Sun Gravitational Influence

Planet	Average Orbit Radius [km]	Approx. Sphere of Influence [km]
Mercury ☿	57,910,000	112,500
Venus ♀	108,200,000	616,400
Earth ⊕	149,600,000	916,600
Mars ♂	227,940,000	577,400
Jupiter ♃	778,330,000	48,223,000
Saturn ♄	1,429,400,000	54,679,000
Uranus ♅	2,870,990,000	51,792,000
Neptune ♆	4,504,300,000	86,975,000
Pluto ♇	5,913,520,000	15,146,000

Note that the spheres of influence around a planet grows larger as the planets mean orbit radius increases, and thus the local gravitational influence of the Sun diminishes. This is how the Saturn sphere of influence radius $r_{♄}$ is larger than the Jupiter sphere of influence radius $r_{♃}$. This is true even though Jupiter is much more massive than Saturn.

Problems

10.1 Use the binomial expansion theory to verify the Legendre polynomial identity given in Eq. (10.6).

10.2 Verify the derivation of gravitational acceleration expression given in Eq. (10.32) by starting with Eq. (10.29) and showing all steps in between.

10.3 Verify that $I_{\xi\xi} = I_{\eta\eta} = I_{\zeta\zeta}$ does reduce the gravitation acceleration to be

$$\boldsymbol{a}_p = -\frac{G}{r^2}\hat{\imath}_r$$

10.4 ♣ Consider a planet spinning at a constant rate ω about its polar axis $\hat{\imath}_\zeta$. Assume that the planet is an ellipsoid of revolution with the inertia's $I_{\xi\xi} = I_{\eta\eta}$ and $I_{\zeta\zeta}$. The motion of a particle on the planet's surface is studied in this problem.

a) Show that for this case the MacCullagh Gravity Potential Approximation is given by

$$V(\boldsymbol{r}) \approx -\frac{\mu}{r} + \frac{G}{2}\frac{(I_{\zeta\zeta} - I_{\xi\xi})}{r^3}\left(3\sin^2\phi - 1\right)$$

b) Write the equations of motion of a particle located at $\boldsymbol{r} = (r, \phi, \theta)$ on the planet's surface using the unit direction vectors $\{\hat{\imath}_r, \hat{\imath}_\theta, \hat{\imath}_\phi\}$ shown in Figure 10.4. If the planet's surface is frictionless, in what direction would the particle slide?

c) Establish that the modified potential function $V' = V - \frac{1}{2}r^2\omega^2\cos^2\phi$ provides the measured gravitational acceleration of this particle on the planet's surface.

d) Assume that $I_{\zeta\zeta} < I_{\xi\xi}$ (Note: this is not typically the case). Consider the approximation $r \approx r_e(1+\eta)$ of the radial distance of the particle with respect to the planet's center with $\eta << 1$, where r_e is the planet's equatorial radius. Assuming a frictionless planet surface, establish the condition for equilibrium such that the particle will remain fixed as seen by the rotating planet reference frame.

Bibliography

[1] Battin, R. H., *An Introduction to the Mathematics and Methods of Astrodynamics*, AIAA Education Series, New York, 1987.

[2] Bate, R., Mueller, D. D., and White, J. E., *Fundamentals of Astrodynamics*, Dover Publications, Inc., New York, NY, 1971.

[3] Roy, A. E., *Orbital Motion*, Adam Hilger Ltd, Bristol, England, 2nd ed., 1982.

Chapter Eleven

Perturbation Methods

While on the macroscopic scale unpowered spacecraft and celestial bodies typically prescribe elliptic, parabolic or hyperbolic trajectories, on the small scale every body is suspect to minor disturbative accelerations. Precise orbit calculations require us to account for these perturbations. For example, perturbations could be due to the gravitational attraction of other celestial bodies, the non-spherical shape of planets, atmospheric drag or solar radiation pressure. While these affects are usually relatively small compared to the dominant point mass gravitational attraction, they do have an important impact when studying the *long-term* behavior of these orbits.

Most perturbations methods have in common that instead of directly numerically integrating the orbits themselves, only deviations from a two-body solution are studied. This allows us to separate the large effect of the dominant point mass gravitational field from the small effect of the disturbance and enables the use of analytical approximations. In the mid *19-th* century, both the English astronomer John Couch Adams (1819–1892) and the French astronomer Urbain-Jean-Joseph Le Verrier (1811–1877) independently used the method of variations of parameters when studying the irregularities of the motion of Uranus.[1] Their amazingly precise observations and calculations predicted the existence of a then unknown planet Neptune which was causing the observed deviations. In 1846 both were able to detect Neptune within one degree of the predicted position. After Le Verrier became director of the Paris Observatory in 1854, he began to study the motion of the planet Mercury. He found that this planet's orbit had similar irregularities as he found in Uranus' trajectory. He thus predicted the existence of new planet closer to the sun, which he name Vulcan. However, after his death the wobbles in the motion of Mercury were later explained using Einstein's general theory on relativity without the need to introduce a new planet.

This chapter will first discuss Encke's method to introduce the ideas of perturbation induced departure motion. Following this a general procedure called the *variation of parameters* is discussed which is mathematically more challenging, but provides very valuable algebraic insight into the effects of these

disturbances on the various evolutions of orbits. Finally, the state transition matrix is introduced. This matrix is often used in dynamics and control theory to predict the departure motion of an object relative to a nominal motion trajectory if small perturbations are present.

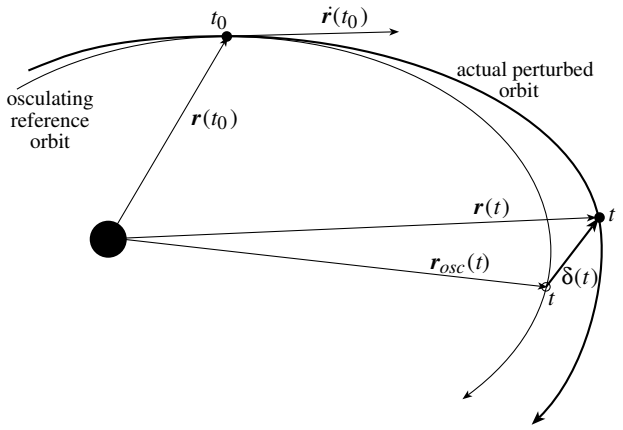


Figure 11.1: Illustration of Encke's Method

11.1 Encke's Method

The relative equations of motion of a mass m_2 relative to m_1 was found in Eq. (8.42) to be

$$\ddot{\boldsymbol{r}} = -\frac{\mu}{r^3}\boldsymbol{r} + \boldsymbol{a}_d \tag{11.1}$$

where $\boldsymbol{r} = \boldsymbol{r}_2 - \boldsymbol{r}_1$ and $\boldsymbol{a}_d$ is the perturbative acceleration defined in Eq. (8.43). To directly numerically integrate these equations of motion is commonly referred to as *Cowell's Method*. This is the preferred method, as compared to analytical approximations, if the perturbative acceleration vector $\boldsymbol{a}_d$ is changing generally and is of the same order of magnitude as the dominant gravitational acceleration. However, in many applications $\boldsymbol{a}_d$ is orders of magnitude smaller than the dominant gravitational force. For example, for a satellite in a low Earth orbit, the dominant Earth oblateness effect is three orders of magnitude smaller than the spherical Earth gravity field acceleration. Other effects such as solar radiation drag, atmospheric drag and the gravitational pull of the moon are even smaller. Using Cowell's method to solve the equations of motion accurately captures the small deviations from the two-body Keplerian solution; however, this numerical solution takes no advantage of the nearness of the motion to the analytically solvable two-body case.

The guiding principle of Encke's method is to use the known, closed-form Keplerian solution to compute the dominant trajectory, and then numerically

solve a second differential equation for the deviations $\boldsymbol{\delta}$ from the two-body solution. This concept is illustrated in Figure 11.1. At time t_0 the reference Keplerian orbit is established using the instantaneous $\boldsymbol{r}(t_0)$ and $\dot{\boldsymbol{r}}(t_0)$ vectors. This reference orbit is referred to as the osculating orbit since at t_0 it "kisses" or osculates the actual orbit at t_0.

$$\boldsymbol{r}(t_0) = \boldsymbol{r}_{osc}(t_0) \qquad \dot{\boldsymbol{r}}(t_0) = \dot{\boldsymbol{r}}_{osc}(t_0) \tag{11.2}$$

Therefore, the two orbits only differ in their mutual curvatures at time t_0 due to the different acceleration expressions. The positions and velocity vectors of each orbit will be identical at t_0. The true trajectory is given by the differential equation in Eq. (11.1). The osculating orbit is determined through

$$\ddot{\boldsymbol{r}}_{osc} = -\frac{\mu}{r_{osc}^3}\boldsymbol{r}_{osc} \tag{11.3}$$

Defining the orbit deviation $\boldsymbol{\delta}$ to be

$$\boldsymbol{\delta} = \boldsymbol{r} - \boldsymbol{r}_{osc} \tag{11.4}$$

the *relative equations of motion* of the actual orbit compared to the osculating orbit are given by

$$\ddot{\boldsymbol{\delta}} = \ddot{\boldsymbol{r}} - \ddot{\boldsymbol{r}}_{osc} = \mu\left(\frac{\boldsymbol{r}_{osc}}{r_{osc}^3} - \frac{\boldsymbol{r}}{r^3}\right) + \boldsymbol{a}_d \tag{11.5}$$

Adding and subtracting $\boldsymbol{r}/r_{osc}^3$ terms and using Eq. (11.4), the $\ddot{\boldsymbol{\delta}}$ expression is rewritten as

$$\ddot{\boldsymbol{\delta}} = -\frac{\mu}{r_{osc}^3}\boldsymbol{\delta} - \frac{\mu}{r_{osc}^3}\left(\frac{r_{osc}^3}{r^3} - 1\right)\boldsymbol{r} + \boldsymbol{a}_d \tag{11.6}$$

Defining the scalar function f to be

$$f(\boldsymbol{r}, \boldsymbol{r}_{osc}) = \frac{r_{osc}^3}{r^3} - 1 \tag{11.7}$$

the deviation equations of motion are written in their final form as

$$\ddot{\boldsymbol{\delta}} = -\frac{\mu}{r_{osc}^3}\left(\boldsymbol{\delta} + f(\boldsymbol{r}, \boldsymbol{r}_{osc})\boldsymbol{r}\right) + \boldsymbol{a}_d \tag{11.8}$$

Unfortunately computing the function f directly using the algebraic expression in Eq. (11.7) is numerically challenging since two essentially identical numbers are subtracted from another. To compute this term without having to resort to higher precision arithmetic, it is possible to rewrite the function $f()$ into a more convenient form. Let us define the scalar parameter q as

$$q \equiv \frac{\boldsymbol{\delta}\cdot\boldsymbol{\delta} - 2\boldsymbol{\delta}\cdot\boldsymbol{r}}{r^2} \tag{11.9}$$

where it is easy to show from Eq. (11.9) that

$$\frac{r_{osc}^2}{r^2} = 1 + q \tag{11.10}$$

Substituting Eq. (11.10) into the $f()$ expression in Eq. (11.7) and multiplying and dividing the result by $(1 + (1 + q)^{3/2})$, we are able to express $f()$ directly in terms of q without subtracting any near-equal numbers in the following form

$$f(q) = q\frac{3 + 3q + q^2}{1 + (1 + q)^{3/2}} \tag{11.11}$$

Let's summarize Encke's method. At time t_0 the osculating orbit is setup invoking the conditions in Eq. (11.2) and the conditions

$$\boldsymbol{\delta}(t_0) = \dot{\boldsymbol{\delta}}(t_0) = 0 \tag{11.12}$$

From here on, osculating position and velocity vectors $\boldsymbol{r}_{osc}$ and $\dot{\boldsymbol{r}}_{osc}$ are computed using the 2-body solution such as the F and G functions. The deviation position and velocity vectors $\boldsymbol{\delta}$ and $\dot{\boldsymbol{\delta}}$ are computed by numerically integrating Eq. (11.8) twice. The actual orbit position and velocity vectors $\boldsymbol{r}(t)$ and $\dot{\boldsymbol{r}}(t)$ are then computed through

$$\boldsymbol{r}(t) = \boldsymbol{r}_{osc}(t) + \boldsymbol{\delta}(t) \tag{11.13}$$

$$\dot{\boldsymbol{r}}(t) = \dot{\boldsymbol{r}}_{osc}(t) + \dot{\boldsymbol{\delta}}(t) \tag{11.14}$$

The result is a natural splitting in the computation of the 2-body and disturbance components. For systems with a relatively small $\boldsymbol{a}_d$ vector, this allows larger integration step sizes to be used than if Cowell's method were employed.

If the deviation vector $\boldsymbol{\delta}$ grows too large at time t_1, then the osculating orbit conditions are reset and a new 2-body reference orbit is found which osculates with the current $\boldsymbol{r}(t)$ and $\dot{\boldsymbol{r}}(t)$. This process of resetting the reference osculating orbit is called *orbit rectification*. Given a chosen small tolerance ϵ, a common method to determine whether or not to rectify the osculating orbit is to see if

$$\left|\frac{\delta}{r}\right| + \left|\frac{\dot{\delta}}{\dot{r}}\right| \geq \epsilon \tag{11.15}$$

The newly rectified osculating orbit will again "kiss" or osculate the actual orbit at time t_1.

11.2 Variation of Parameters

The method of variation of parameters can be viewed as the continuous limit of rectification of an osculating orbits at each instant of time. Given the instantaneous inertial position and velocity vectors $\boldsymbol{r}(t)$ and $\dot{\boldsymbol{r}}(t)$ of the perturbed problem, we can always compute a corresponding set of six instantaneous orbit

elements e_i as is discussed previously in Chapter 8. However, for this perturbed problem the six orbit elements will no longer remain constants, but become time varying parameters. These instantaneous orbit elements, whose corresponding keplerian 2-body orbit kisses the current perturbed orbit, are referred to as the *osculating* orbit elements.

11.2.1 General Methodology

Assuming the scalar parameters e_i are integration constants of an *un-perturbed* motion, then the method of variation of parameters seeks a corresponding set of differential equations for $\dot{e}_i$ such that the perturbed motion description instantaneously has the same algebraic form as the unperturbed motion. We would like the only difference between the solution of the perturbed and un-perturbed problem to be that the elements $e_i(t)$ are time-varying. For example, the parameters e_i could be the initial conditions of a general dynamical problem or the six orbit elements of a Keplerian two-body solution.

Let $\boldsymbol{x}$ be a N-dimensional position vector. Since mechanical dynamical systems are second order, the solution will have have $2N$ integration constants e_i. Let the $2N \times 1$ vector $\boldsymbol{e}$ be defined as

$$\boldsymbol{e} = (e_1, \cdots, e_{2N})^T \tag{11.16}$$

In general, the *un-perturbed* solution of a dynamical system can be written as

$$\boldsymbol{x}(t) = \boldsymbol{f}(t, \boldsymbol{e}) \tag{11.17}$$

$$\dot{\boldsymbol{x}}(t) = \frac{d\boldsymbol{f}}{dt}(t, \boldsymbol{e}) = \frac{\partial \boldsymbol{f}}{\partial t} \tag{11.18}$$

$$\ddot{\boldsymbol{x}}(t) = \frac{d^2\boldsymbol{f}}{dt^2}(t, \boldsymbol{e}) = \frac{\partial^2 \boldsymbol{f}}{\partial t^2} \tag{11.19}$$

The last steps in Eqs. (11.18)and (11.19) hold since the elements e_i are constants for the un-perturbed problem. Using variation of parameters, we seek a solution of the *perturbed* motion which has the same algebraic form, so we require

$$\boldsymbol{x}(t) = \boldsymbol{f}(t, \boldsymbol{e}(t)) \tag{11.20}$$

$$\frac{d\boldsymbol{x}(t)}{dt} = \frac{\partial \boldsymbol{f}}{\partial t}(t, \boldsymbol{e}(t)) \tag{11.21}$$

This is analogous to the osculating conditions used to compute Encke's method in Eq. (11.12). These conditions force the perturbed and the osculating unperturbed solutions to have the *same* position and velocity vector, the only difference will be in the acceleration expression where the perturbative acceleration $\boldsymbol{a}_d$ appears.

$$\frac{d^2\boldsymbol{x}}{dt^2} = \frac{\partial^2 \boldsymbol{f}}{\partial t^2} + \boldsymbol{a}_d \tag{11.22}$$

To enforce the osculation condition and the equations of motion, the elements e_i are now treated as time-varying parameters. With $\boldsymbol{e} = \boldsymbol{e}(t)$, the chain rule of differentiation provides

$$\frac{d\boldsymbol{x}(t)}{dt} = \frac{\partial \boldsymbol{f}}{\partial t} + \left[\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{e}}\right] \frac{d\boldsymbol{e}}{dt} \tag{11.23}$$

Comparing Eqs. (11.21) and (11.23), it is evident that for the perturbed and unperturbed solutions to have the same velocity expression, the following condition must be true:

$$\left[\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{e}}\right] \frac{d\boldsymbol{e}}{dt} = \mathbf{0}_{Nx1} \tag{11.24}$$

Eq. (11.24) provides N "osculation" constraints that the differential equations $\dot{e}_i$ must satisfy. The second set of N constraints required to determine $\dot{\boldsymbol{e}}$ is found through the acceleration expression. Taking the derivative of Eq. (11.21) we find

$$\frac{d^2\boldsymbol{x}}{dt^2} = \frac{\partial^2 \boldsymbol{f}}{\partial t^2} + \left[\frac{\partial^2 \boldsymbol{f}}{\partial t \partial \boldsymbol{e}}\right] \frac{d\boldsymbol{e}}{dt} \tag{11.25}$$

Comparing Eqs. (11.22) and (11.25) we find the remaining N constraints on $\dot{\boldsymbol{e}}$ to be

$$\left[\frac{\partial^2 \boldsymbol{f}}{\partial t \partial \boldsymbol{e}}\right] \frac{d\boldsymbol{e}}{dt} = \boldsymbol{a}_d \tag{11.26}$$

Combining Eqs. (11.24) and (11.26), the coupled osculating conditions on $\dot{\boldsymbol{e}}$ are written compactly as

$$\begin{bmatrix} \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{e}} \\ \frac{\partial^2 \boldsymbol{f}}{\partial t \partial \boldsymbol{e}} \end{bmatrix} \frac{d\boldsymbol{e}}{dt} = [L]\dot{\boldsymbol{e}} = \begin{bmatrix} \mathbf{0}_{Nx1} \\ \boldsymbol{a}_d \end{bmatrix} \tag{11.27}$$

where the $2N \times 2N$ matrix $[L]$ is referred to as the *Lagrangian matrix*. Note that this $[L]$ matrix, as defined in Eq. (11.27), is generally fully-populated and may depend explicitly on time. The desired differential equations $\dot{\boldsymbol{e}}$ are now found by inverting the matrix in Eq. (11.27). For many applications such as the perturbed two-body problem, a compact analytical inverse of this matrix is possible. Lagrange developed a very elegant process called the Lagrange Brackets that facilitates this process as is shown in the next section.

Example 11.1: Let us illustrate the basic concepts of variation of parameters by attempting to solve the forced linear oscillator

$$\ddot{x} = -\omega^2 x + a_d(t, x, \dot{x}, \dots)$$

with the initial conditions $x(0) = x_0$ and $\dot{x}(0) = \dot{x}_0$. For the un-perturbed case, a well-known closed-form solution exists for this oscillator problem. Let

us choose to write the un-perturbed motion $x(t)$ and associated velocity $\dot{x}(t)$ as

$$x(t) = x_0 \cos\omega t + \frac{\dot{x}_0}{\omega} \sin\omega t$$
$$\dot{x}(t) = -\omega x_0 \sin\omega t + \dot{x}_0 \cos\omega t$$

For this scalar system there are only two integration constants e_i, namely the initial conditions x_0 and $\dot{x}_0$. Therefore we choose to set

$$e_1 = x_0 \qquad e_2 = \dot{x}_0$$

Using the method of variation of parameters, we seek for the solution of the perturbed system to be of the same algebraic form as shown above for the un-perturbed system, with the exception that x_0 and $\dot{x}_0$ are now treated as time-varying parameters to compensate for the perturbations. In essence, we are trying to find differential equations for $e_1(t)$ and $e_2(t)$ such that

$$x(t) = e_1(t) \cos\omega t + \frac{e_2(t)}{\omega} \sin\omega t$$
$$\dot{x}(t) = -\omega e_1(t) \sin\omega t + e_2(t) \cos\omega t$$

holds for the perturbed system. To write the motion $x(t)$ in the form of Eq. (11.17) we define f to be

$$f = e_1 \cos\omega t + \frac{e_2}{\omega} \sin\omega t$$

Computing the necessary partial derivatives of f as required in Eq. (11.27), the osculating conditions on $\dot{\boldsymbol{e}}$ are given by

$$\begin{bmatrix} \cos\omega t & \frac{1}{\omega}\sin\omega t \\ -\omega\sin\omega t & \cos\omega t \end{bmatrix} \begin{pmatrix} \dot{e}_1 \\ \dot{e}_2 \end{pmatrix} = \begin{pmatrix} 0 \\ a_d \end{pmatrix}$$

Inverting the 2×2 matrix, the desired differential equations for x_0 and $\dot{x}_0$ are

$$\dot{e}_1 = \frac{dx_0}{dt} = -\left(\frac{1}{\omega}\sin\omega t\right) a_d$$
$$\dot{e}_2 = \frac{d\dot{x}_0}{dt} = (\cos\omega t)\, a_d$$

Given a perturbative acceleration a_d, these differential equations show how the initial conditions x_0 and $\dot{x}_0$ have to be adjusted for the algebraic form of the un-perturbed solution still to hold.

11.2.2 Lagrangian Brackets

Lagrange developed a convenient method to calculate the various elements of the Lagrange matrix $[L]$ called the *Lagrangian brackets*. His method leads to a matrix $[L]$ which is sparsely populated and thus easy to invert. We will develop these brackets here using a Keplerian orbit as the example. Assume all perturbations experienced are conservative and can therefore be modeled

through a scalar disturbance potential function R, where $R = R(\boldsymbol{r})$ only. The potential energy V per unit mass of the system is then given by

$$V(\boldsymbol{r}) = -\frac{\mu}{r} - R(\boldsymbol{r}) \tag{11.28}$$

where r is the instantaneous orbit radius. The equations of motion for this system are given by

$$\frac{d\boldsymbol{r}}{dt} = \boldsymbol{v} \tag{11.29}$$

$$\frac{d\boldsymbol{v}}{dt} = -\frac{\partial V}{\partial \boldsymbol{r}} = -\frac{\mu}{r^3}\boldsymbol{r} + \left[\frac{\partial R}{\partial \boldsymbol{r}}\right]^T \tag{11.30}$$

Note that the partial derivative of the disturbance potential function R with respect to the position vector $\boldsymbol{r}$ is the same as the previously defined disturbance acceleration $\boldsymbol{a}_d$.

$$\left[\frac{\partial R}{\partial \boldsymbol{r}}\right]^T = \boldsymbol{a}_d \tag{11.31}$$

We have seen that for Keplerian motion, the orbit shape and orientation is parameterized by six orbit elements. Therefore, let the 6×1 vector $\boldsymbol{e}$ contain the six orbit elements. Which set of six elements is chosen is of no consequence in the general development. If the disturbance potential function R were zero, then these six orbit elements would remain constant. Using the F and G solution, for example, we are able to write the instantaneous position and velocity vectors as functions of the constant orbit elements and time.

$$\boldsymbol{r} = \boldsymbol{r}(\boldsymbol{e}, t) \qquad \boldsymbol{v} = \boldsymbol{v}(\boldsymbol{e}, t) \tag{11.32}$$

However, in the presence of the the disturbance potential function R, the chosen six orbit elements will vary with time. Using variation of parameters, we again seek a solution to the perturbed problem whose instantaneous position and velocity vectors are equal to the unperturbed Keplerian solution given in Eq. (11.32). To match up the velocity vectors, the condition in Eq. (11.21) requires that

$$\boldsymbol{v} = \frac{\partial \boldsymbol{r}}{\partial t} \tag{11.33}$$

Taking the derivative of the position vector given in Eq. (11.32) and allowing $\boldsymbol{e}$ to be time varying, we find that

$$\frac{d\boldsymbol{r}}{dt} = \boldsymbol{v} = \frac{d\boldsymbol{r}}{d\boldsymbol{e}}\frac{d\boldsymbol{e}}{dt} + \frac{\partial \boldsymbol{r}}{\partial t} \tag{11.34}$$

Making use of Eq. (11.33), we again find the first osculating condition given in Eq. (11.24).

$$\frac{d\boldsymbol{r}}{d\boldsymbol{e}}\frac{d\boldsymbol{e}}{dt} = \boldsymbol{0}_{3\times 1} \tag{11.35}$$

Taking the derivative of the velocity expression in Eq. (11.32) we find

$$\frac{d\boldsymbol{v}}{dt} = \frac{\partial \boldsymbol{v}}{\partial t} + \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{e}} \frac{d\boldsymbol{e}}{dt} \tag{11.36}$$

Since $\partial \boldsymbol{v}/\partial t$ is the Keplerian component of the acceleration, comparing Eq. (11.36) to Eq. (11.30) we find the second osculating condition in Eq. (11.26) to be

$$\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{e}} \frac{d\boldsymbol{e}}{dt} = \left[\frac{\partial R}{\partial \boldsymbol{r}}\right]^T \tag{11.37}$$

We could combine the conditions on the orbit element rate vector $\dot{\boldsymbol{e}}$ in Eqs. (11.35) and (11.37) as is done in Eq. (11.27). However, since the perturbation potential R is often expressed in terms of the orbit elements, the more convenient compact form is possible. Following a pattern introduced by Lagrange, we pre-multiply Eq. (11.37) by $(\partial \boldsymbol{r}/\partial \boldsymbol{e})^T$ and pre-multiply Eq. (11.35) by $-(\partial \boldsymbol{v}/\partial \boldsymbol{e})^T$ and then subtract the latter from the first. After simplifying the algebraic expression we find

$$\underbrace{\left[\left(\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{e}}\right)^T \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{e}} - \left(\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{e}}\right)^T \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{e}}\right]}_{[L]} \frac{d\boldsymbol{e}}{dt} = \left[\frac{\partial R}{\partial \boldsymbol{r}} \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{e}}\right]^T = \left[\frac{\partial R}{\partial \boldsymbol{e}}\right]^T \tag{11.38}$$

where $[L]$ is a new Lagrangian coefficient matrix. Thus the Lagrangian variational equations, which express the orbit element drift $\dot{\boldsymbol{e}}$ due to the disturbance potential R, are written as

$$\dot{\boldsymbol{e}} = [L]^{-1} \left[\frac{\partial R}{\partial \boldsymbol{e}}\right]^T \tag{11.39}$$

The individual entries L_{ij} of this matrix are called the *Lagrangian brackets* and are computed through

$$L_{ij} = [e_i, e_j] = \left(\frac{\partial \boldsymbol{r}}{\partial e_i}\right)^T \frac{\partial \boldsymbol{v}}{\partial e_j} - \left(\frac{\partial \boldsymbol{v}}{\partial e_i}\right)^T \frac{\partial \boldsymbol{r}}{\partial e_j} \tag{11.40}$$

Using the state vector $\boldsymbol{s}$

$$\boldsymbol{s}(t, \boldsymbol{e}) = \begin{pmatrix} \boldsymbol{r}(t, \boldsymbol{e}) \\ \boldsymbol{v}(t, \boldsymbol{e}) \end{pmatrix} \tag{11.41}$$

the matrix $[L]$ is written in the compact form

$$[L] = \frac{\partial \boldsymbol{s}}{\partial \boldsymbol{e}}^T [J] \frac{\partial \boldsymbol{s}}{\partial \boldsymbol{e}} \tag{11.42}$$

where the matrix $[J]$ is defined as

$$[J] = \begin{bmatrix} 0_{3\times 3} & I_{3\times 3} \\ -I_{3\times 3} & 0_{3\times 3} \end{bmatrix} \tag{11.43}$$

Since $[J]$ is symplectic, it satisfies the property

$$[J][J] = -[I_{6\times 6}] \tag{11.44}$$

The Lagrangian bracket operator L_{ij} satisfies the the following three properties:

$$[e_i, e_j] = -[e_j, e_i] \tag{11.45a}$$
$$[e_i, e_i] = 0 \tag{11.45b}$$
$$\frac{\partial}{\partial t}[e_i, e_j] = 0 \tag{11.45c}$$

The latter property is the most amazing truth; the L_{ij} elements are constants of the un-perturbed motion. In terms of the Lagrangian matrix $[L]$, these properties are summarized as

$$[L]^T = -[L] \tag{11.46a}$$
$$\frac{\partial}{\partial t}[L] = 0 \tag{11.46b}$$

The skew-symmetry of $[L]$ is immediately verified by studying the definition of the Lagrangian bracket operator in Eq. (11.40). To verify that $[L]$ never explicitly depends on the time variable t, we take the partial derivative of the Lagrangian bracket definition with respect to time.

$$\begin{aligned}\frac{\partial}{\partial t}[e_i, e_j] = &\frac{\partial}{\partial t}\left(\frac{\partial \boldsymbol{r}}{\partial e_i}\right)^T\left(\frac{\partial \boldsymbol{v}}{\partial e_j}\right) + \left(\frac{\partial \boldsymbol{r}}{\partial e_i}\right)^T\frac{\partial}{\partial t}\left(\frac{\partial \boldsymbol{v}}{\partial e_j}\right)\\ &- \frac{\partial}{\partial t}\left(\frac{\partial \boldsymbol{v}}{\partial e_i}\right)^T\left(\frac{\partial \boldsymbol{r}}{\partial e_j}\right) - \left(\frac{\partial \boldsymbol{v}}{\partial e_i}\right)^T\frac{\partial}{\partial t}\left(\frac{\partial \boldsymbol{r}}{\partial e_j}\right)\end{aligned} \tag{11.47}$$

Switching the order of the partial differentiation and making use of the osculating condition in Eq. (11.33), we find

$$\begin{aligned}\frac{\partial}{\partial t}[e_i, e_j] = &\left(\frac{\partial \boldsymbol{v}}{\partial e_i}\right)^T\left(\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{e}_j}\right) + \left(\frac{\partial \boldsymbol{r}}{\partial e_i}\right)^T\frac{\partial}{\partial e_j}\left(\frac{\partial \boldsymbol{v}}{\partial t}\right)\\ &- \frac{\partial}{\partial e_i}\left(\frac{\partial \boldsymbol{v}}{\partial t}\right)^T\left(\frac{\partial \boldsymbol{r}}{\partial e_j}\right) - \left(\frac{\partial \boldsymbol{v}}{\partial e_i}\right)^T\left(\frac{\partial \boldsymbol{v}}{\partial e_j}\right)\end{aligned} \tag{11.48}$$

The first and fourth term clearly cancel each other. To see that the remaining two terms do cancel, note that the expression $\partial \boldsymbol{v}/\partial t$ is the unperturbed acceleration and therefore

$$\frac{\partial \boldsymbol{v}}{\partial t} = -\left(\frac{\partial V}{\partial \boldsymbol{r}}\right)^T \tag{11.49}$$

must hold. Using this condition, we are able to reduce Eq. (11.48) to

$$\frac{\partial}{\partial t}[e_i, e_j] = -\frac{\partial}{\partial e_j}\frac{\partial V}{\partial \boldsymbol{r}}\frac{\partial \boldsymbol{r}}{\partial e_i} + \frac{\partial}{\partial e_i}\frac{\partial V}{\partial \boldsymbol{r}}\frac{\partial \boldsymbol{r}}{\partial e_j} = -\frac{\partial^2 V}{\partial e_j \partial e_i} + \frac{\partial^2 V}{\partial e_i \partial e_j} = 0 \tag{11.50}$$

which proves that the Lagrangian bracket never depends explicitly on the time variable. Note that the Lagrangian bracket operator generates information required to produce the necessary elements rates $\dot{\boldsymbol{e}}$ such that the current two-body orbit, corresponding to the current elements $\boldsymbol{e}(t)$, osculates or "kisses" the actual non-Keplerian orbit. Since $[L]$ does not explicitly depend on time, only implicitly through $\boldsymbol{e}(t)$, it does not matter where along the *osculating orbit* we evaluate the bracket expressions. Therefore we will be able to evaluate the Lagrangian brackets anywhere it is convenient along the instantaneous two-body (osculating) orbit. In particular, we will find it convenient later on to evaluate the bracket expressions at periapses to reduce the amount of algebra involved. This is similar in spirit to the development in Chapter 8 where we chose to evaluate the total energy constant α at periapses.

Example 11.2: Let us examine the motion of vertical spring mass system under the influence of a constant gravitational field as illustrated in Figure 11.2. If the mass m is at a height $\hat{y}$, then the spring has zero potential energy. The height above ground is measured by the variable y. The velocity of the mass is given by $v = \dot{y}$. The potential energy of the system is then given by

$$V = mgy + \frac{k}{2}(y - \hat{y})^2$$

The equations of motion are then given by

$$\ddot{y} = -\frac{1}{m}\frac{\partial V}{\partial y} = -g - \frac{k}{m}(y - \hat{y})$$

where g is the local gravitational acceleration and k is the spring stiffness coefficient.

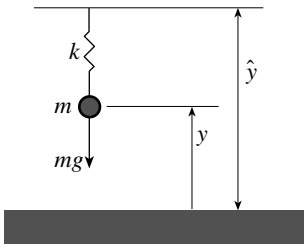


Figure 11.2: Oscillating Point Mass Illustration

Assume we choose to treat the effect of the spring as a disturbance. Therefore we set the mass-less disturbance potential R equal to

$$R = \frac{k}{2m}(y - \hat{y})^2$$

Without the spring present, the equations of motion of the mass m are easily solved to the well-known form

$$y(t) = y_0 + v_0 t - \frac{g}{2}t^2$$
$$v(t) = v_0 - gt$$

where y_0 and v_0 are the initial position and velocity. For the unperturbed problem (i.e. spring-less), these two quantities are the constants of integration and do not vary with time.

We now ask the question, how would we need to vary y_0 and v_0 such that the perturbed (i.e. spring effect included) problem has the same position and velocity expression as the unperturbed problem. Choosing y_0 and v_0 to be are constants elements of the unperturbed problem, we set

$$e_1 = y_0 \qquad e_2 = v_0$$

To find the desired variation $\dot{e}_i$, we must first find the various Lagrangian brackets. The unperturbed solutions above for $y(t)$ and $v(t)$ provides the necessary relationships between the unperturbed position and velocity quantities and the chosen elements e_i. Computing the required partial derivatives we find

$$\frac{\partial y}{\partial e_1} = 1 \qquad \frac{\partial y}{\partial e_2} = t$$
$$\frac{\partial v}{\partial e_1} = 0 \qquad \frac{\partial v}{\partial e_2} = 1$$

Since the skew-symmetric Lagrangian bracket for this problem is only a 2×2 matrix, we only need to compute the one off-diagonal term.

$$[e_1, e_2] = \frac{\partial y}{\partial e_1}\frac{\partial v}{\partial e_2} - \frac{\partial v}{\partial e_1}\frac{\partial y}{\partial e_2} = 1$$

The resulting osculating condition on $\dot{e}_i$ (i.e. $\dot{y}_0$ and $\dot{v}_0$) are

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{pmatrix} \dot{e}_1 \\ \dot{e}_2 \end{pmatrix} = \begin{pmatrix} \frac{\partial R}{\partial e_1} \\ \frac{\partial R}{\partial e_2} \end{pmatrix}$$

For a general conservative disturbance potential function R, the variational equations for y_0 and v_0 are then given by

$$\dot{e}_1 = \dot{y}_0 = -\frac{\partial R}{\partial e_1}$$
$$\dot{e}_2 = \dot{v}_0 = \frac{\partial R}{\partial e_2}$$

To compute the variational equations for this particular example, we write the disturbance potential R as

$$R = \frac{k}{2m}(e_1 + e_2 t - \frac{g}{2}t^2 - \hat{y})^2$$

Taking the required partial derivatives with respect to e_1 and e_2, the variational equations for "spring disturbance" are

$$\dot{e}_1 = \dot{y}_0 = -\frac{k}{m}(y - \hat{y})t = -\frac{k}{m}\left(e_1 + e_2 t - \frac{g}{2}t^2 - \hat{y}\right)t$$
$$\dot{e}_2 = \dot{v}_0 = \frac{k}{m}(y - \hat{y}) = \frac{k}{m}\left(e_1 + e_2 t - \frac{g}{2}t^2 - \hat{y}\right)$$

These variational equations are readily verified by taking two time derivatives of the unperturbed position expression and treating y_0 and v_0 now as time

varying quantities. The acceleration expression should be the same as is given in the exact equations of motion of the system.

11.2.3 Lagrange's Planetary Equations

Given the Lagrangian bracket operator, we now proceed to develop the variational equations for a chosen set of orbit elements $\boldsymbol{e}$. The resulting set of first order differential equations is commonly known as *Lagrange's planetary equations*. We chose the classical orbit element set

$$\boldsymbol{e} = (\Omega, i, \omega, a, e, M_0)^T \tag{11.51}$$

as our Keplerian integration constants where Ω, i and ω are the (3-1-3) Euler angle set orientating the orbit plane and line of periapses, a and e are the semi-major axis and eccentricity respectively, and M_0 is the initial mean anomaly. The following development will be analogous for other sets of orbit elements. Let n be the mean angular motion defined in Eq. (8.104), then M_0 can be expressed in terms of the time of periapses passage τ as

$$M_0 = -n\tau \tag{11.52}$$

Kepler's equation is then given by

$$M(t) = M_0 + nt = n(t - \tau) = E - e\sin E \tag{11.53}$$

To orient the orbit plane and line of periapses, we define the orbit reference frame $\mathcal{O}$ as

$$\mathcal{O} : \{\hat{\boldsymbol{\imath}}_e, \hat{\boldsymbol{\imath}}_p, \hat{\boldsymbol{\imath}}_h\} \tag{11.54}$$

where $\hat{\boldsymbol{\imath}}_e$ points toward periapses, $\hat{\boldsymbol{\imath}}_h$ is orbit plane normal and $\hat{\boldsymbol{\imath}}_p$ is perpendicular to the previous two unit vectors as discussed in Chapter 8. Before evaluating the Lagrangian brackets, we first need to find analytical expressions for the inertial position and velocity vectors $\boldsymbol{r}(t)$ and $\boldsymbol{v}(t)$ in terms of the chosen orbit element vector $\boldsymbol{e}$. Using Eqs. (8.16) and (8.17) we are able to express the position vector $\boldsymbol{r}$ in the orbit frame $\mathcal{O}$ as

$${}^{\mathcal{O}}\boldsymbol{r} = \begin{pmatrix} a(\cos E - e) \\ b\sin E \\ 0 \end{pmatrix}^{\mathcal{O}} \tag{11.55}$$

where $b = a\sqrt{1-e^2}$ is the semi-minor axis. Taking the $\mathcal{O}$ frame derivative and making use of the $\dot{E}$ expression in Eq. (8.101), the velocity vector $\boldsymbol{v}$ is expressed in the $\mathcal{O}$ frame as

$${}^{\mathcal{O}}\boldsymbol{v} = \begin{pmatrix} -a\sin E \\ b\cos E \\ 0 \end{pmatrix}^{\mathcal{O}} \frac{n}{1 - e\cos E} \tag{11.56}$$

Let $[NO]$ be the direction cosine matrix that maps orbit frame vector components into inertial frame $\mathcal{N}$ vector components. The position vector $\boldsymbol{r}$ is then expressed in inertial components as

$$^{\mathcal{N}}\boldsymbol{r} = [NO]\ ^{\mathcal{O}}\boldsymbol{r} \tag{11.57}$$

Note that for the two-body solution the orbit plane orientation is constant and therefore $[NO]$ is also constant. The inertial velocity vector $\boldsymbol{v}$ is then simply expressed as

$$^{\mathcal{N}}\boldsymbol{v} = [NO]\ ^{\mathcal{O}}\boldsymbol{v} \tag{11.58}$$

To express the position and velocity vectors in terms of the given orbit elements, we parameterize the direction cosine matrix $[NO]$ in terms of the (3-1-3) Euler angles Ω, i and ω as

$$[NO] = [\hat{\boldsymbol{i}}_e \hat{\boldsymbol{i}}_p \hat{\boldsymbol{i}}_h] = \begin{bmatrix} c\omega c\Omega - s\omega cis\Omega & -s\omega c\Omega - c\omega cis\Omega & sis\Omega \\ c\omega s\Omega + s\omega cic\Omega & -s\omega s\Omega + c\omega cic\Omega & -sic\Omega \\ s\omega si & c\omega si & ci \end{bmatrix} \tag{11.59}$$

where the short-hand notation $si = \sin i$ and $ci = \cos i$ is used again. Thus the position and velocity vectors are expressed in terms of Ω, i, ω, a, e and M_0. We note that dependence on the initial mean anomaly M_0 is implicit in the eccentric anomaly E which must satisfy Kepler's equation in Eq. (11.53).

Expressions for $\boldsymbol{r}$ and $\boldsymbol{v}$ in the form given in Eqs. (11.57) and (11.58) will be very convenient in the following development. Note that only $^{\mathcal{O}}\boldsymbol{r}$ and $^{\mathcal{O}}\boldsymbol{v}$ depend on the orbit elements a and e, and implicitly on M_0. Similarly, the direction cosine matrix $[NO]$ only depends on the three Euler angles and not on the other orbit elements. This separation of dependencies will greatly simplify the resulting algebra when computing the various partial derivatives required with the Lagrangian brackets. For example, the partial derivative of $\boldsymbol{r}$ with respect to Ω is simply given by

$$\frac{\partial \boldsymbol{r}}{\partial \Omega} = \frac{\partial}{\partial \Omega}[NO]\ ^{\mathcal{O}}\boldsymbol{r} \tag{11.60}$$

To find the partials of $[NO]$, it is convenient to first find the partials of the orbit frame $\mathcal{O}$ unit vectors. Expressing the unit vector components as

$$\hat{\boldsymbol{i}}_e = (\hat{\imath}_{e1}, \hat{\imath}_{e2}, \hat{\imath}_{e3})^T \tag{11.61a}$$

$$\hat{\boldsymbol{i}}_p = (\hat{\imath}_{p1}, \hat{\imath}_{p2}, \hat{\imath}_{p3})^T \tag{11.61b}$$

$$\hat{\boldsymbol{i}}_h = (\hat{\imath}_{h1}, \hat{\imath}_{h2}, \hat{\imath}_{h3})^T \tag{11.61c}$$

we are able to express the various partial derivatives of the $\mathcal{O}$ frame unit vectors

with respect to the three (3-1-3) Euler angles as

$$\frac{\partial \hat{\boldsymbol{i}}_e}{\partial \Omega} = \begin{pmatrix} -\hat{\imath}_{e2} \\ \hat{\imath}_{e1} \\ 0 \end{pmatrix} \qquad \frac{\partial \hat{\boldsymbol{i}}_p}{\partial \Omega} = \begin{pmatrix} -\hat{\imath}_{p2} \\ \hat{\imath}_{p1} \\ 0 \end{pmatrix} \qquad \frac{\partial \hat{\boldsymbol{i}}_h}{\partial \Omega} = \begin{pmatrix} -\hat{\imath}_{h2} \\ \hat{\imath}_{h1} \\ 0 \end{pmatrix} \tag{11.62}$$

$$\frac{\partial \hat{\boldsymbol{i}}_e}{\partial i} = \sin\omega \hat{\boldsymbol{i}}_h \qquad \frac{\partial \hat{\boldsymbol{i}}_p}{\partial i} = \cos\omega \hat{\boldsymbol{i}}_h \qquad \frac{\partial \hat{\boldsymbol{i}}_h}{\partial i} = \begin{pmatrix} \cos i \sin\Omega \\ -\cos i \cos\Omega \\ -\sin i \end{pmatrix} \tag{11.63}$$

$$\frac{\partial \hat{\boldsymbol{i}}_e}{\partial \omega} = \hat{\boldsymbol{i}}_p \qquad \frac{\partial \hat{\boldsymbol{i}}_p}{\partial \omega} = -\hat{\boldsymbol{i}}_e \qquad \frac{\partial \hat{\boldsymbol{i}}_h}{\partial \omega} = 0 \tag{11.64}$$

As was discussed earlier, since the Lagrangian brackets do not explicitly depend on time, we are able to compute the brackets (i.e. the partial derivatives) on any convenient location of the osculating orbit. Choosing to evaluate the partial derivatives at periapses, at this point we have $E = 0$ and $t = \tau$. The position and velocity vectors in Eqs. (11.55) and (11.56) evaluated at periapses reduce to

$${}^{\mathcal{O}}\boldsymbol{r}\big|_{periapses} = {}^{\mathcal{O}}\begin{pmatrix} q \\ 0 \\ 0 \end{pmatrix} \qquad {}^{\mathcal{O}}\boldsymbol{v}\big|_{periapses} = {}^{\mathcal{O}}\begin{pmatrix} 0 \\ \frac{nab}{q} \\ 0 \end{pmatrix} \tag{11.65}$$

where $q = a(1-e)$. The partial derivatives of $\boldsymbol{r}$ and $\boldsymbol{v}$ with respect to the (3-1-3) Euler angles, evaluated at periapses, are then given by

$$\frac{\partial \boldsymbol{r}}{\partial \Omega} = q \begin{pmatrix} -\hat{\imath}_{e2} \\ \hat{\imath}_{e1} \\ 0 \end{pmatrix} \qquad \frac{\partial \boldsymbol{r}}{\partial i} = q \sin\omega \hat{\boldsymbol{i}}_h \qquad \frac{\partial \boldsymbol{r}}{\partial \omega} = q\hat{\boldsymbol{i}}_p \tag{11.66}$$

$$\frac{\partial \boldsymbol{v}}{\partial \Omega} = q \begin{pmatrix} -\hat{\imath}_{p2} \\ \hat{\imath}_{p1} \\ 0 \end{pmatrix} \qquad \frac{\partial \boldsymbol{v}}{\partial i} = \frac{nab\cos\omega}{q}\hat{\boldsymbol{i}}_h \qquad \frac{\partial \boldsymbol{v}}{\partial \omega} = -\frac{nab}{q}\hat{\boldsymbol{i}}_e \tag{11.67}$$

Since only ${}^{\mathcal{O}}\boldsymbol{r}$ and ${}^{\mathcal{O}}\boldsymbol{v}$ depend on the semi-major axis a, finding the partial derivatives of $\boldsymbol{r}$ and $\boldsymbol{v}$ with respect to a only involves

$$\frac{\partial \boldsymbol{r}}{\partial a} = [NO]\frac{\partial {}^{\mathcal{O}}\boldsymbol{r}}{\partial a} \qquad \frac{\partial \boldsymbol{v}}{\partial a} = [NO]\frac{\partial {}^{\mathcal{O}}\boldsymbol{v}}{\partial a} \tag{11.68}$$

Let us first focus on the position vector partial derivative. Noting that the eccentric anomaly E implicitly depends on the orbit elements, we find that

$$\frac{\partial {}^{\mathcal{O}}\boldsymbol{r}}{\partial a} = {}^{\mathcal{O}}\begin{pmatrix} \cos E - e - a\sin E\frac{\partial E}{\partial a} \\ \sqrt{1-e^2}\sin E + b\cos E\frac{\partial e}{\partial a} \\ 0 \end{pmatrix} \tag{11.69}$$

To express $\partial E/\partial a$, we take the partial derivative of Kepler's equation of Eq. (11.53) with respect to a.

$$\frac{\partial E}{\partial a} - e\cos E\frac{\partial E}{\partial a} = \frac{\partial n}{\partial a}t = -\frac{3n}{2a}t \tag{11.70}$$

Making use of Eq. (8.15) we find that

$$\frac{\partial E}{\partial a} = -\frac{3nt}{2r} \tag{11.71}$$

Substituting Eq. (11.71) into Eq. (11.69) we express $\partial^{\mathcal{O}}\boldsymbol{r}/\partial a$ as

$$\frac{\partial^{\mathcal{O}}\boldsymbol{r}}{\partial a} = {}^{\mathcal{O}}\begin{pmatrix} \cos E - e + a\sin E\frac{3nt}{2r} \\ \sqrt{1-e^2}\sin E - b\cos E\frac{3nt}{2r} \\ 0 \end{pmatrix} \tag{11.72}$$

Since the Lagrangian bracket does not explicitly depend on time, we again choose to evaluate this partial derivative at periapses where $r = q$, $E = 0$ and $t = \tau$.

$$\left.\frac{\partial^{\mathcal{O}}\boldsymbol{r}}{\partial a}\right|_{periapses} = {}^{\mathcal{O}}\begin{pmatrix} q/a \\ \frac{3bM_0}{2q} \\ 0 \end{pmatrix} \tag{11.73}$$

Substituting Eq. (11.73) into (11.68) we obtain the required partial derivative of the position vector with respect to a.

$$\frac{\partial \boldsymbol{r}}{\partial a} = \frac{q}{a}\hat{\boldsymbol{\imath}}_e + \frac{3bM_0}{2q}\hat{\boldsymbol{\imath}}_p \tag{11.74}$$

Evaluating the partial derivative of the velocity vector with respect to a follows the same logic, but is algebraically more complicated. Using Eq. (11.71), taking the partial derivative of the velocity vector $\boldsymbol{v}$ with respect to a we find

$$\frac{\partial^{\mathcal{O}}\boldsymbol{v}}{\partial a} = {}^{\mathcal{O}}\begin{pmatrix} -\frac{an}{r}\sin E + \frac{3a^2nt}{2r^2}\left(n\cos E + \sin E - \frac{ane\sin^2 E}{r}\right) \\ -\frac{bn}{2r}\cos E + \frac{3abn^2t\sin E}{2r^2}\left(1 + \frac{a\cos E}{r}\right) \\ 0 \end{pmatrix} \tag{11.75}$$

Evaluating this partial derivative at periapses we get

$$\left.\frac{\partial^{\mathcal{O}}\boldsymbol{v}}{\partial a}\right|_{periapses} = -{}^{\mathcal{O}}\begin{pmatrix} \frac{3a^2nM_0}{2q^2} \\ \frac{bn}{2q} \\ 0 \end{pmatrix} \tag{11.76}$$

After substituting Eq. (11.76) into Eq. (11.68) we are able to write the partial derivative of $\boldsymbol{v}$ with respect to a as

$$\frac{\partial \boldsymbol{v}}{\partial a} = -\frac{3a^2nM_0}{2a^2}\hat{\boldsymbol{\imath}}_e - \frac{bn}{2q}\hat{\boldsymbol{\imath}}_p \tag{11.77}$$

The partial derivatives of $\boldsymbol{r}$ and $\boldsymbol{v}$ with respect to e and M_0 are found in a similar manner. By finding the partial derivative of Kepler's equation with

respect to either e or M_0, we find the required partial derivatives of the eccentric anomaly E for any point on the osculating orbit to be

$$\frac{\partial E}{\partial e} = \frac{a \sin E}{r} \qquad \frac{\partial E}{\partial M_0} = \frac{a}{r} \tag{11.78}$$

Evaluating these partial derivatives at periapses we find

$$\left.\frac{\partial E}{\partial e}\right|_{periapses} = 0 \qquad \left.\frac{\partial E}{\partial M_0}\right|_{periapses} = \frac{a}{q} \tag{11.79}$$

Using these partial derivatives and following the same path as is used in the development of the partial derivatives of $\boldsymbol{r}$ and $\boldsymbol{v}$ with respect to a, we find the remaining partial derivatives to be:

$$\frac{\partial \boldsymbol{r}}{\partial e} = -a\hat{\boldsymbol{\imath}}_e \qquad \frac{\partial \boldsymbol{v}}{\partial e} = \frac{na^3}{bq}\hat{\boldsymbol{\imath}}_p \tag{11.80}$$

$$\frac{\partial \boldsymbol{r}}{\partial M_0} = \frac{ab}{q}\hat{\boldsymbol{\imath}}_p \qquad \frac{\partial \boldsymbol{v}}{\partial M_0} = -\frac{na^3}{q^2}\hat{\boldsymbol{\imath}}_e \tag{11.81}$$

With the various partial derivative of $\boldsymbol{r}$ and $\boldsymbol{v}$ with respect to the orbit elements evaluated, finding the Lagrangian brackets now is a relatively straight forward matter. We will only carry out the algebra here for the Lagrangian bracket $[a, \omega]$ for illustrative purposes. Using Eq. (11.40), the bracket $[a, \omega]$ is defined as

$$[a, \omega] = \frac{\partial \boldsymbol{r}}{\partial a} \cdot \frac{\partial \boldsymbol{v}}{\partial \omega} - \frac{\partial \boldsymbol{v}}{\partial a} \cdot \frac{\partial \boldsymbol{r}}{\partial \omega} \tag{11.82}$$

Using the partial derivative expressions that we have just found, and noting that the orbit frame $\mathcal{O}$ unit vectors are mutually orthogonal, the bracket $[a, \omega]$ is reduced to

$$\begin{aligned}[a, \omega] &= \left(\frac{q}{a}\hat{\boldsymbol{\imath}}_e + \frac{3bM_0}{2q}\hat{\boldsymbol{\imath}}_p\right) \cdot \left(-\frac{nab}{q}\hat{\boldsymbol{\imath}}_e\right) + \left(\frac{3a^2 nM_0}{2q^{2'}}\hat{\boldsymbol{\imath}}_e + \frac{bn}{2q}\hat{\boldsymbol{\imath}}_p\right) \cdot q\hat{\boldsymbol{\imath}}_p \\ &= -nb + \frac{nb}{2} = -\frac{nb}{2}\end{aligned}$$

Due to the skew-symmetry of the Lagrangian matrix, only 15 distinct brackets need to be evaluated and are shown below:[1]

$$[i, \Omega] = nab \sin i \qquad [\omega, i] = 0 \qquad [a, \omega] = -\frac{nb}{2} \qquad [e, a] = 0 \quad [M_0, e] = 0$$

$$[\omega, \Omega] = 0 \qquad [a, i] = 0 \qquad [e, \omega] = \frac{na^3 e}{b} \quad [M_0, a] = \frac{na}{2}$$

$$[a, \Omega] = -\frac{nb}{2}\cos i \qquad [e, i] = 0 \quad [M_0, \omega] = 0$$

$$[e, \Omega] = \frac{na^3 e}{b}\cos i \quad [M_0, i] = 0$$

$$[M_0, \Omega] = 0$$

Using these Lagrangian brackets, the osculating constraints on $\dot{\boldsymbol{e}}$ for a conservative disturbance potential R are

$$\begin{bmatrix} 0 & -nab\sin i & 0 & \frac{nb}{2}\cos i & -\frac{na^3e}{b}\cos i & 0 \\ nab\sin i & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{nb}{2} & -\frac{na^3e}{b} & 0 \\ -\frac{nb}{2}\cos i & 0 & -\frac{nb}{2} & 0 & 0 & -\frac{na}{2} \\ \frac{na^3e}{b}\cos i & 0 & \frac{na^3e}{b} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{na}{2} & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{d\Omega}{dt} \\ \frac{di}{dt} \\ \frac{d\omega}{dt} \\ \frac{da}{dt} \\ \frac{de}{dt} \\ \frac{dM_0}{dt} \end{bmatrix} = \begin{bmatrix} \frac{\partial R}{\partial \Omega} \\ \frac{\partial R}{\partial i} \\ \frac{\partial R}{\partial \omega} \\ \frac{\partial R}{\partial a} \\ \frac{\partial R}{\partial e} \\ \frac{\partial R}{\partial M_0} \end{bmatrix} \tag{11.83}$$

Since many of the Lagrangian brackets are zero, it is relatively easy to invert the Lagrangian matrix $[L]$ in Eq. (11.83) and solve for the desired orbit parameter rate vector $\dot{\boldsymbol{e}}$. For example, due to the sparse nature of $[L]$, solving for $\dot{\Omega}$ and $\dot{a}$ is trivial. Solving for the six unknown e_i, we obtain the classical form of *Lagrange's planetary equations*:

$$\frac{d\Omega}{dt} = \frac{1}{nab\sin i}\frac{\partial R}{\partial i} \tag{11.84a}$$

$$\frac{di}{dt} = -\frac{1}{nab\sin i}\frac{\partial R}{\partial \Omega} + \frac{\cos i}{nab\sin i}\frac{\partial R}{\partial \omega} \tag{11.84b}$$

$$\frac{d\omega}{dt} = -\frac{\cos i}{nab\sin i}\frac{\partial R}{\partial i} + \frac{b}{na^3e}\frac{\partial R}{\partial e} \tag{11.84c}$$

$$\frac{da}{dt} = \frac{2}{na}\frac{\partial R}{\partial M_0} \tag{11.84d}$$

$$\frac{de}{dt} = -\frac{b}{na^3e}\frac{\partial R}{\partial \omega} + \frac{b^2}{na^4e}\frac{\partial R}{\partial M_0} \tag{11.84e}$$

$$\frac{dM_0}{dt} = -\frac{2}{na}\frac{\partial R}{\partial a} - \frac{b^2}{na^4e}\frac{\partial R}{\partial e} \tag{11.84f}$$

Note that these variational equations, and thus also the corresponding Lagrangian bracket matrix $[L]$, are singular whenever either the eccentricity is 0 (i.e. circular orbit) or the inclination angle i is zero degrees (i.e. equatorial orbit). Further, these differential equations also loose their validity whenever the orbit energies rise to the point where the corresponding trajectories are parabolic or hyperbolic. The reason for this is the underlying assumption in the preceding development that the position vector $\boldsymbol{r}(t)$ is the solution of an elliptic orbit. However, by choosing a different set of orbit elements it is possible to avoid these singularities.[1]

Example 11.3: Let us use Lagrange's planetary equations to compute how the J_2 gravitational perturbation will affect the orbit elements. The disturbance potential function $R(\boldsymbol{r})$ for the J_2 oblateness component of the spherical gravitational harmonics is defined in Eq. (10.63) and is given by

$$R(\boldsymbol{r}) = -\frac{J_2}{2}\frac{\mu}{r}\left(\frac{r_{eq}}{r}\right)^2\left(3\sin^2\phi - 1\right)$$

Using Eq. (8.130) we note that

$$\sin\phi = \hat{\boldsymbol{\imath}}_r \cdot \hat{\boldsymbol{\imath}}_z = \sin(\omega + f)\sin i$$

This allows us to write the disturbance potential function R in terms of the chosen orbit elements as

$$R(\boldsymbol{e}) = -\frac{J_2}{2}\frac{\mu}{r}\left(\frac{r_{eq}}{r}\right)^2 \left(3\sin^2(\omega+f)\sin^2 i - 1\right) \tag{11.85}$$

Expressing the orbit radius using Eq. (8.6) and performing the partial derivatives required in Lagrange's planetary equations, we find the J_2 gravitational perturbation to cause the following instantaneous rates in the orbit elements.[2]

$$\frac{d\Omega}{dt} = -3J_2 n\frac{a^2}{br}\left(\frac{r_{eq}}{r}\right)^2 \sin^2\theta\cos i \tag{11.86a}$$

$$\frac{di}{dt} = -\frac{3}{4}J_2 n\frac{a^2}{br}\left(\frac{r_{eq}}{r}\right)^2 \sin(2\theta)\sin(2i) \tag{11.86b}$$

$$\begin{aligned}\frac{d\omega}{dt} = \frac{3}{4}J_2 n\frac{a^2}{bep}\left(\frac{r_{eq}}{r}\right)^2 \Big[&\cos f\Big(2(1+e^2) \\ &- \big((3+e^2) - (3+5e^2)\cos(2i)\big)\sin^2\theta\Big) \\ &+ 2e\Big(2\cos(2i) + \cos(2\theta) - \cos(2(\theta-i)) \\ &+ 2\sin^2(\theta+i)\Big)\Big]\end{aligned} \tag{11.86c}$$

$$\begin{aligned}\frac{da}{dt} = -3J_2 n\frac{a^4}{br^2}\left(\frac{r_{eq}}{r}\right)^2 \Big[&e\sin f(1 - 3\sin^2\theta\sin^2 i) \\ &+ \frac{p}{r}\sin(2\theta)\sin^2 i\Big]\end{aligned} \tag{11.86d}$$

$$\begin{aligned}\frac{de}{dt} = -\frac{3}{2}J_2 n\frac{a^2}{br}\left(\frac{r_{eq}}{r}\right)^2 \Big[&\frac{p}{r}\sin f(1 - 3\sin^2\theta\sin^2 i) \\ &+ (e + \cos f(2 + e\cos f)\sin(2\theta)\sin^2 i)\Big]\end{aligned} \tag{11.86e}$$

$$\begin{aligned}\frac{dM_0}{dt} = -\frac{3}{16}J_2\frac{n}{e}\left(\frac{r_{eq}}{r}\right)^2 \cos f\Big[&2 + 6\cos(2i) + 6\cos(2\theta) \\ &- 3\cos(2(\theta-i)) - 3\cos(2(\theta+i))\Big]\end{aligned} \tag{11.86f}$$

Taking the partial derivatives, note that

$$\frac{\partial f}{\partial M_0} = \frac{\partial f}{\partial E}\cdot\frac{\partial E}{\partial M_0} = \frac{ab}{r^2}$$

Also, the true anomaly $\theta = \omega + f$ is used here to simplify the rate expressions. Note that the orbit elements rates provides in Eq. (11.86) are instantaneous, or sometimes also referred to as the osculating, orbit elements rates. In some application it is convenient to deal with orbit averaged rates, or also called the mean element rates. The J_2 perturbation causes three types of rates; (1) short period oscillations (2) long period oscillations and (3) secular drift. The short and long period oscillations are periodic deviation from the element mean values during an orbit. For long term orbit study the secular drift is of great interest. Using asymptotic expansion theory, it is possible to extract

the secular rates and express the mean or orbit average effect of J_2 on the orbit elements $\overline{\boldsymbol{e}}$ as:[3]

$$\frac{d\overline{\Omega}}{dt} = -\frac{3}{2}J_2 n\left(\frac{r_{eq}}{p}\right)^2 \cos i \tag{11.87a}$$

$$\frac{d\overline{i}}{dt} = 0 \tag{11.87b}$$

$$\frac{d\overline{\omega}}{dt} = \frac{3}{4}J_2 n\left(\frac{r_{eq}}{p}\right)^2 (5\cos^2 i - 1) \tag{11.87c}$$

$$\frac{d\overline{a}}{dt} = 0 \tag{11.87d}$$

$$\frac{d\overline{e}}{dt} = 0 \tag{11.87e}$$

$$\frac{d\overline{M}_0}{dt} = \frac{3}{4}J_2 n\left(\frac{r_{eq}}{p}\right)^2 \sqrt{1-e^2}(3\cos^2 i - 1) \tag{11.87f}$$

The rigorous derivation of these rates is beyond the scope of this chapter. The mapping between mean and osculating orbit elements is found as part of Brouwer's artificial satellite theory in Reference 3. Note that there exists a critical inclination angle $i_{crit} = 63.4249$ degrees where no mean regression of the argument of perigee occurs.

11.2.4 Poisson Brackets

The Lagrangian coefficient matrix $[L]$ is written in compact form in Eq. (11.42). The Poisson matrix $[P]$ is closely related to the Lagrangian matrix $[L]$. The 6×6 matrix $[P]$ is defined as

$$[P] = \frac{\partial \boldsymbol{e}}{\partial \boldsymbol{s}}[J]\frac{\partial \boldsymbol{e}}{\partial \boldsymbol{s}}^T \tag{11.88}$$

with $\boldsymbol{s}(t, \boldsymbol{e})$ being defined in Eq. (11.41) and the symplectic matrix $[J]$ being defined in Eq. (11.43). The elements of $[P]$ are called the Poisson bracket and defined as

$$P_{ij} = (e_i, e_j) = \frac{\partial e_i}{\partial \boldsymbol{r}}\left(\frac{\partial e_j}{\partial \boldsymbol{v}}\right)^T - \frac{\partial e_i}{\partial \boldsymbol{v}}\left(\frac{\partial e_j}{\partial \boldsymbol{r}}\right)^T \tag{11.89}$$

The Poisson bracket (e_i, e_j) satisfies the same three conditions as does the Lagrangian bracket $[e_i, e_j]$.

$$(e_i, e_j) = -(e_j, e_i) \tag{11.90a}$$

$$(e_i, e_i) = 0 \tag{11.90b}$$

$$\frac{\partial}{\partial t}(e_i, e_j) = 0 \tag{11.90c}$$

To find the relationship between $[L]$ and $[P]$, let us evaluate their matrix product:

$$[L][P] = \left(\frac{\partial \boldsymbol{s}}{\partial \boldsymbol{e}}\right)^T [J] \underbrace{\frac{\partial \boldsymbol{s}}{\partial \boldsymbol{e}} \frac{\partial \boldsymbol{e}}{\partial \boldsymbol{s}}}_{[I_{6\times 6}]} [J] \left(\frac{\partial \boldsymbol{e}}{\partial \boldsymbol{s}}\right)^T = \left(\frac{\partial \boldsymbol{s}}{\partial \boldsymbol{e}}\right)^T \underbrace{[J][J]}_{-[I_{6\times 6}]} \left(\frac{\partial \boldsymbol{e}}{\partial \boldsymbol{s}}\right)^T = -[I_{6\times 6}] \tag{11.91}$$

Thus the Poisson matrix $[P]$ is the negative inverse of the Lagrangian matrix $[L]$.

$$[P] = -[L]^{-1} \tag{11.92}$$

Using the skew-symmetry property in Eq. (11.90a), this is also written as

$$[P]^T = [L]^{-1} \tag{11.93}$$

Recall the Lagrangian variational equations in Eq. (11.39) where the inverse of the Lagrangian matrix $[L]$ appeared. Using Eq. (11.93), this equation can now be written as

$$\frac{d\boldsymbol{e}}{dt} = [P]^T \left[\frac{\partial R}{\partial \boldsymbol{e}}\right]^T \tag{11.94}$$

Using the Poisson matrix $[P]$ or the Poisson brackets (e_i, e_j), we are able to avoid the matrix inversion of $[L]$ when solving for $\dot{\boldsymbol{e}}$. After evaluating all the Poisson brackets (e_i, e_j) of $[P]$ for a given orbit element set vector $\boldsymbol{e}$, Eq. (11.94) would lead directly to Lagrange's planetary equations shown in Eq. (11.84).

Instead of re-deriving these equations, we will use the Poisson brackets to derive the $\dot{\boldsymbol{e}}$ expressions when the disturbance is provided by a disturbance acceleration $\boldsymbol{a}_d$. Substituting the $[P]$ definition in Eq. (11.88) into Eq. (11.94) yields

$$\frac{d\boldsymbol{e}}{dt} = -\frac{\partial \boldsymbol{e}}{\partial \boldsymbol{s}} [J] \left[\frac{\partial R}{\partial \boldsymbol{s}} \frac{\partial \boldsymbol{e}}{\partial \boldsymbol{s}}\right]^T \tag{11.95}$$

After substituting the definitions of $\boldsymbol{s}$ and $[J]$, this is written as

$$\frac{d\boldsymbol{e}}{dt} = \frac{\partial \boldsymbol{e}}{\partial \boldsymbol{v}} \left[\frac{\partial R}{\partial \boldsymbol{r}}\right]^T - \frac{\partial \boldsymbol{e}}{\partial \boldsymbol{r}} \left[\frac{\partial R}{\partial \boldsymbol{v}}\right]^T \tag{11.96}$$

Finally, making use of Eq. (11.31) and the fact that $R = R(\boldsymbol{r})$, we arrive at the compact orbit element rate equation

$$\frac{d\boldsymbol{e}}{dt} = \frac{\partial \boldsymbol{e}}{\partial \boldsymbol{v}} \boldsymbol{a}_d \tag{11.97}$$

Note that no coordinate frame has been chosen yet for the vectors $\boldsymbol{v}$ and $\boldsymbol{a}_d$ in Eq. (11.97). This equation is very useful in that it holds for all possible frames. The following subsections develop the variational equations for several classical orbit elements while also maintaining a general vector description without choosing a particular coordinate frame.

Variation of the Semi-Major Axis

To find the semi-major axis rate $\dot{a}$ due to a disturbance acceleration $\boldsymbol{a}_d$, we must first express the semi-major axis in terms of the velocity vector $\boldsymbol{v}$. To do so, we make us of the vis-viva equation:

$$v^2 = \boldsymbol{v}^T\boldsymbol{v} = \frac{2\mu}{r} - \frac{\mu}{a}$$

Taking the partial derivative of this equation leads to

$$\frac{\partial a}{\partial \boldsymbol{v}} = \frac{2a^2}{\mu}\boldsymbol{v}^T \tag{11.98}$$

Substituting this $\partial a/\partial \boldsymbol{v}$ expression into Eq. (11.97) leads to the variation equation of the semi-major axis a.

$$\frac{da}{dt} = \frac{\partial a}{\partial \boldsymbol{v}}\boldsymbol{a}_d = \frac{2a^2}{\mu}\boldsymbol{v}^T\boldsymbol{a}_d \tag{11.99}$$

Note that Eq. (11.99) holds for any coordinate frame assigned to the vectors $\boldsymbol{v}$ and $\boldsymbol{a}_d$.

Variation of the Eccentricity

To find the variation of the eccentricity e, the variation of the angular momentum magnitude h is required. To express h in terms of the velocity vector $\boldsymbol{v}$, we make use of the fundamental definition $\boldsymbol{h} = \boldsymbol{r} \times \boldsymbol{v}$ of the angular momentum vector. The scalar h is then related to $\boldsymbol{v}$ through

$$h^2 = (\boldsymbol{r} \times \boldsymbol{v}) \cdot (\boldsymbol{r} \times \boldsymbol{v}) = r^2\boldsymbol{v}^T\boldsymbol{v} - (\boldsymbol{r}^T\boldsymbol{v})^2 \tag{11.100}$$

where some standard vector cross product identities were used. Taking the partial derivative of Eq. (11.100) leads to

$$\frac{\partial h}{\partial \boldsymbol{v}} = \frac{1}{h}\left(r^2\boldsymbol{v}^T - (\boldsymbol{r}^T\boldsymbol{v})\boldsymbol{r}^T\right) \tag{11.101}$$

Since the partial derivatives of both h and a with respect to $\boldsymbol{v}$ are know at this point, it is convenient to solve for the eccentricity variation by using Eqs. (8.9) and (8.65).

$$h^2 = \mu p = \mu a(1 - e^2) \tag{11.102}$$

Taking the partial derivative of Eq. (11.102) with respect to $\boldsymbol{v}$ leads to

$$2h\frac{\partial h}{\partial v} = \mu\frac{\partial a}{\partial \boldsymbol{v}}(1 - e^2) - 2\mu ae\frac{\partial e}{\partial \boldsymbol{v}} \tag{11.103}$$

Substituting Eqs. (11.99) and (11.101) into Eq. (11.103) leads to

$$\frac{\partial e}{\partial \boldsymbol{v}} = \frac{1}{\mu ae}\left((pa - r^2)\boldsymbol{v}^T + (\boldsymbol{r}^T\boldsymbol{v})\boldsymbol{r}^T\right) \tag{11.104}$$

The variational equation for the eccentricity is then given by

$$\frac{de}{dt} = \frac{\partial e}{\partial \boldsymbol{v}}\boldsymbol{a}_d = \frac{1}{\mu a e}\left((pa - r^2)\boldsymbol{v}^T\boldsymbol{a}_d + (\boldsymbol{r}^T\boldsymbol{v})\boldsymbol{r}^T\boldsymbol{a}_d\right) \tag{11.105}$$

Variation of the Longitude of the Ascending Node and the Inclination Angle

To find the variations of the longitude of the ascending node Ω and inclination angle i, we first define an inertial coordinate frame $\mathcal{N} : \{\hat{\boldsymbol{\imath}}_x, \hat{\boldsymbol{\imath}}_y, \hat{\boldsymbol{\imath}}_z\}$ and the orbit frame $\mathcal{O} : \{\hat{\boldsymbol{\imath}}_n, \hat{\boldsymbol{\imath}}_m, \hat{\boldsymbol{\imath}}_z\}$. The angles Ω and i are the first two (3-1-3) Euler angles which determine the orbit plane orientation. Using Eq. (3.35), the frames $\mathcal{O}$ and $\mathcal{N}$ are then related through

$$\begin{Bmatrix} \hat{\boldsymbol{\imath}}_n \\ \hat{\boldsymbol{\imath}}_m \\ \hat{\boldsymbol{\imath}}_h \end{Bmatrix} = \begin{bmatrix} \cos\Omega & \sin\Omega & 0 \\ -\sin\Omega\cos i & \cos\Omega\cos i & \sin i \\ \sin\Omega\sin i & -\cos\Omega\sin i & \cos i \end{bmatrix} \begin{Bmatrix} \hat{\boldsymbol{\imath}}_x \\ \hat{\boldsymbol{\imath}}_y \\ \hat{\boldsymbol{\imath}}_z \end{Bmatrix} \tag{11.106}$$

The angular momentum vector $\boldsymbol{h}$ can now be expressed in terms of both Ω and i as

$$\boldsymbol{h} = h\hat{\boldsymbol{\imath}}_h = h(\sin\Omega\sin i\ \hat{\boldsymbol{\imath}}_x - \cos\Omega\sin i\ \hat{\boldsymbol{\imath}}_y + \cos i\ \hat{\boldsymbol{\imath}}_z) \tag{11.107}$$

Taking the partial derivative of $\boldsymbol{h}$ with respect to $\boldsymbol{v}$, and making use of the reference frame unit direction vector identities in Eq. (11.106), we find

$$\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{v}} = h\sin i\ \hat{\boldsymbol{\imath}}_h\frac{\partial \Omega}{\partial \boldsymbol{v}} - h\hat{\boldsymbol{\imath}}_m\frac{\partial i}{\partial \boldsymbol{v}} + \hat{\boldsymbol{\imath}}_h\frac{\partial h}{\partial \boldsymbol{v}} \tag{11.108}$$

To obtain $\partial \boldsymbol{h}/\partial \boldsymbol{v}$, we make use of the fundamental definition of $\boldsymbol{h}$.

$$\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{v}} = \frac{\partial(\boldsymbol{r}\times\boldsymbol{v})}{\partial v} = (\boldsymbol{r}\times) \tag{11.109}$$

Substituting Eq. (11.109) into Eq. (11.108) and taking the transpose, we find

$$-(\boldsymbol{r}\times) = h\sin i\left(\frac{\partial \Omega}{\partial \boldsymbol{v}}\right)^T\hat{\boldsymbol{\imath}}_n^T - h\left(\frac{\partial i}{\partial \boldsymbol{v}}\right)^T\hat{\boldsymbol{\imath}}_m^T + \left(\frac{\partial h}{\partial \boldsymbol{v}}\right)^T\hat{\boldsymbol{\imath}}_h^T \tag{11.110}$$

By multiplying Eq. (11.110) by $\hat{\boldsymbol{\imath}}_n$, we are able to isolate the Ω derivative.

$$\frac{\partial \Omega}{\partial \boldsymbol{v}} = \frac{1}{h\sin i}(\hat{\boldsymbol{\imath}}_n\times\boldsymbol{r})^T \tag{11.111}$$

Since the position vector $\boldsymbol{r}$ can be expressed as

$$\boldsymbol{r} = r(\cos\theta\hat{\boldsymbol{\imath}}_n + \sin\theta\hat{\boldsymbol{\imath}}_m) \tag{11.112}$$

where $\theta = \omega + f$, then the partial derivative of Ω is simplified to

$$\frac{\partial \Omega}{\partial \boldsymbol{v}} = \frac{r\sin\theta}{h\sin i}\hat{\boldsymbol{\imath}}_h^T \tag{11.113}$$

To find the partial derivative of the inclination angle with respect to $\boldsymbol{v}$, Eq. (11.110) is multiplied by $\hat{\boldsymbol{\imath}}_m$ and $\boldsymbol{r}$ is replaced by the expression in Eq. (11.112).

$$\frac{\partial i}{\partial \boldsymbol{v}} = \frac{r\cos\theta}{h}\hat{\boldsymbol{\imath}}_h^T \tag{11.114}$$

Substituting Eqs. (11.113) and (11.114) into Eq. (11.97), the variational equations for the longitude of the ascending node and the inclination angle are found to be

$$\frac{d\Omega}{dt} = \frac{\partial \Omega}{\partial \boldsymbol{v}}\boldsymbol{a}_d = \frac{r\sin\theta}{h\sin i}\,\hat{\boldsymbol{\imath}}_h^T\boldsymbol{a}_d \tag{11.115}$$

$$\frac{di}{dt} = \frac{\partial i}{\partial \boldsymbol{v}}\boldsymbol{a}_d = \frac{r\cos\theta}{h}\,\hat{\boldsymbol{\imath}}_h^T\boldsymbol{a}_d \tag{11.116}$$

Variation of the Anomalies

Finding the various anomaly variations is algebraically the most challenging task of all the orbit element variations discussed in this section. First, the variation of the true anomaly will be derived. This result will then be used to derive the variations of the eccentric anomaly E and mean anomaly M.

The orbit equation in Eq. (8.6) provides an equation which relates the true anomaly f to h and e. This is attractive since both the h and e partial derivatives with respect to $\boldsymbol{v}$ have already been derived. Taking the partial derivative of Eq. (8.6) leads to

$$re\sin f\frac{\partial f}{\partial \boldsymbol{v}} = r\cos f\frac{\partial e}{\partial \boldsymbol{v}} - \frac{2h}{\mu}\frac{\partial h}{\partial \boldsymbol{v}} \tag{11.117}$$

We could substitute the previously found partial derivative for both e and h at this point and attempt to solve for $\partial f/\partial \boldsymbol{v}$. However, this path involves a lot of algebra to reduce the answer to a simple form. Battin presents an elegant solution to this problem in Reference 1 which avoids some of the algebra involved in reducing the answer. From the position and velocity vector expressions in Eqs. (8.132) and (8.133), the following identity is found:

$$\frac{\mu}{h}re\sin f = \boldsymbol{r}^T\boldsymbol{v} \tag{11.118}$$

This equation provides another relationship between the true anomaly f and the parameters e and h. Taking the partial derivative of Eq. (11.118) we find

$$re\cos f\frac{\partial f}{\partial \boldsymbol{v}} = -r\sin f\frac{\partial e}{\partial \boldsymbol{v}} + \frac{\boldsymbol{r}^T\boldsymbol{v}}{\mu}\frac{\partial h}{\partial \boldsymbol{v}} + \frac{h}{\mu}\boldsymbol{r}^T \tag{11.119}$$

The reason that Reference 1 uses this second equation along with Eq. (11.117) is now clear. After multiplying Eq. (11.117) by $\sin f$ and Eq. (11.119) by $\cos f$ and adding the two, the following simplified equation for $\partial f/\partial \boldsymbol{v}$ is found.

$$reh\frac{\partial h}{\partial \boldsymbol{v}} = p\cos f\boldsymbol{r}^T - (p+r)\sin f\frac{\partial h}{\partial \boldsymbol{v}} \tag{11.120}$$

Note that the $\partial e/\partial \boldsymbol{v}$ term has been eliminated by this method and that the $\partial f/\partial \boldsymbol{v}$ term is no longer pre-multiplied by $\sin f$. Substituting the partial derivative expression for h in Eq. (11.100), the partial derivative of f is expressed as

$$\frac{\partial f}{\partial \boldsymbol{v}} = \frac{1}{he}\underbrace{\left(\frac{p}{r}\cos f + \frac{p+r}{p}e\sin^2 f\right)}_{\kappa}\boldsymbol{r}^T - \frac{r}{h^2 e}(p+r)\sin f\boldsymbol{v}^T \tag{11.121}$$

The term κ can be further simplified by making use of the orbit equation in Eq. (8.6).

$$\begin{aligned}\kappa &= \frac{p}{r}\cos f + \left(1+\frac{r}{p}\right)\sin^2 f e \\ &= \cos f + e\cos^2 f + e + \frac{re}{p} - \frac{re}{p}\cos^2 f - e\cos^2 f \\ &= e + \cos f\left(1 - \frac{e\cos f}{1+e\cos f}\right) + \frac{re}{p} \\ &= e + \frac{r}{p}(\cos f + e)\end{aligned}$$

Substituting the simplified κ expression back into Eq. (11.121), the reduced $\partial f/\partial \boldsymbol{v}$ expression is given by

$$\frac{\partial f}{\partial \boldsymbol{v}} = \frac{1}{he}\left(\frac{r}{p}(\cos f + e) + e\right)\boldsymbol{r}^T - \frac{r}{h^2 e}(p+r)\sin f\boldsymbol{v}^T \tag{11.122}$$

When computing the anomaly rates, we must take into account that the anomalies do have an unperturbed derivative. Thus

$$\begin{aligned}\frac{df}{dt} &= \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \boldsymbol{v}}\boldsymbol{a}_d \\ &= \frac{h}{r^2} + \frac{1}{he}\left(\frac{r}{p}(\cos f + e) + e\right)\boldsymbol{r}^T\boldsymbol{a}_d - \frac{r}{h^2 e}(p+r)\sin f\boldsymbol{v}^T\boldsymbol{a}_d\end{aligned} \tag{11.123}$$

To find the variation of the of the eccentric anomaly E, Eqs. (8.3) and (8.16) are combined to yield

$$\cos E = \frac{\cos f + e}{1 + e\cos f} \tag{11.124}$$

Again, we are able to express the partial derivative of E in terms of already derived partial derivatives. Similarly, Eqs. (8.4) and (8.17) are combined to yield

$$\sin E = \frac{b\sin f}{a(1+e\cos f} \tag{11.125}$$

Taking the partial derive of $\cos E$ yields

$$\sin E(1+e\cos f)\frac{\partial E}{\partial \boldsymbol{v}} = (e\cos E - 1)\sin f\frac{\partial f}{\partial \boldsymbol{v}} + (1 - \cos E\cos f)\frac{\partial e}{\partial \boldsymbol{v}} \tag{11.126}$$

After substituting the $\cos E$ and $\sin E$ expressions in Eqs. (11.124) and (11.125), the partial derivative of the eccentric anomaly E is written as

$$\frac{\partial E}{\partial \boldsymbol{v}} = \frac{r}{b}\frac{\partial f}{\partial \boldsymbol{v}} - \frac{ra}{pb}\sin f \frac{\partial e}{\partial \boldsymbol{v}} \tag{11.127}$$

After substituting the h and e partial derivative expressions and performing the typical algebraic reductions, the partial derivative of E with respect to $\boldsymbol{v}$ is

$$\frac{\partial E}{\partial \boldsymbol{v}} = \frac{r}{\mu b e}\left(\frac{h}{p}(\cos f + e)\boldsymbol{r}^T - (r+a)\sin f \boldsymbol{v}^T\right) \tag{11.128}$$

Using Eqs. (8.101) and (11.94), the eccentric anomaly variational equation is given by

$$\begin{aligned}\frac{dE}{dt} &= \frac{\partial E}{\partial t} + \frac{\partial E}{\partial \boldsymbol{v}}\boldsymbol{a}_d \\ &= \frac{na}{r} + \frac{r}{\mu b e}\left(\frac{h}{p}(\cos f + e)\boldsymbol{r}^T\boldsymbol{a}_d - (r+a)\sin f \boldsymbol{v}^T\boldsymbol{a}_d\right)\end{aligned} \tag{11.129}$$

To derive the variation of the mean anomaly, the mean anomaly definition in Eq. (8.103) is used. Taking the partial derivative of $M = E - e\sin E$ we find

$$\frac{\partial M}{\partial \boldsymbol{v}} = \frac{r}{a}\frac{\partial E}{\partial \boldsymbol{v}} - \sin E \frac{\partial e}{\partial \boldsymbol{v}} \tag{11.130}$$

After substituting the E and e partial derivatives and performing several algebraic reductions, $\partial M/\partial \boldsymbol{v}$ is given by

$$\frac{\partial M}{\partial \boldsymbol{v}} = \frac{rb}{ha^2 e}\left(\cos f \boldsymbol{r}^T - \frac{a}{h}(r+p)\sin f \boldsymbol{v}^T\right) \tag{11.131}$$

Finally, the mean anomaly variation is given by

$$\begin{aligned}\frac{dM}{dt} &= \frac{\partial M}{\partial t} + \frac{\partial M}{\partial \boldsymbol{v}}\boldsymbol{a}_d \\ &= n + \frac{rb}{ha^2 e}\left(\cos f \boldsymbol{r}^T\boldsymbol{a}_d - \frac{a}{h}(r+p)\sin f \boldsymbol{v}^T\boldsymbol{a}_d\right)\end{aligned} \tag{11.132}$$

Variation of the Argument of Perigee

The last variation to be derived is the variation of the argument of perigee. This is accomplished indirectly through the latitude argument $\theta = \omega + f$. The latitude angle θ is defined as the angle between the unit position vector $\hat{\boldsymbol{\imath}}_r$ and the ascending node direction $\hat{\boldsymbol{\imath}}_n$. Thus

$$\cos\theta = \hat{\boldsymbol{\imath}}_n^T \hat{\boldsymbol{\imath}}_r \tag{11.133}$$

Using Eq. (11.106), this is written as

$$\cos\theta = \cos\Omega(\hat{\boldsymbol{\imath}}_x^T \hat{\boldsymbol{\imath}}_r) + \sin\Omega(\hat{\boldsymbol{\imath}}_y^T \hat{\boldsymbol{\imath}}_r) \tag{11.134}$$

When taking the partial derivative of Eq. (11.134), keep in mind that the inertial unit vectors $\hat{\boldsymbol{\imath}}_x$ and $\hat{\boldsymbol{\imath}}_y$, along with the unit position vector $\hat{\boldsymbol{\imath}}_r$, are invariant with respect to the velocity vector $\boldsymbol{v}$. Thus

$$-\sin\theta\frac{\partial\theta}{\partial\boldsymbol{v}} = \left(-\sin\Omega(\hat{\boldsymbol{\imath}}_x^T\hat{\boldsymbol{\imath}}_r) + \cos\Omega(\hat{\boldsymbol{\imath}}_y^T\hat{\boldsymbol{\imath}}_r)\right)\frac{\partial\Omega}{\partial\boldsymbol{v}} \tag{11.135}$$

After substituting the unit position vector expression from Eq. (8.130), the partial derivative of the latitude angle with respect to $\boldsymbol{v}$ is

$$\frac{\partial\theta}{\partial\boldsymbol{v}} = -\cos i\frac{\partial\Omega}{\partial\boldsymbol{v}} \tag{11.136}$$

Using $\theta = \omega + f$, the partial derivative of the argument of perigee ω is

$$\frac{\partial\omega}{\partial\boldsymbol{v}} = -\frac{\partial f}{\partial\boldsymbol{v}} - \cos i\frac{\partial\Omega}{\partial\boldsymbol{v}} \tag{11.137}$$

After substituting the partial derivatives of f and Ω, the argument of perigee variation is given by

$$\begin{aligned}\frac{d\omega}{dt} = \frac{\partial\omega}{\partial\boldsymbol{v}}\boldsymbol{a}_d = &-\frac{1}{he}\left(\frac{r}{p}(\cos f + e) + e\right)\boldsymbol{r}^T\boldsymbol{a}_d \\ &-\frac{r}{h^2e}(p+r)\sin f\boldsymbol{v}^T\boldsymbol{a}_d - \frac{r\sin\theta}{h\tan i}\,\hat{\boldsymbol{\imath}}_h^T\boldsymbol{a}_d\end{aligned} \tag{11.138}$$

11.2.5 Gauss' Variational Equations

The variational equations developed in the previous section are written in a coordinate frame independent manner. A very convenient frame used in orbital mechanics is the rotating reference frame$\mathcal{R} = \{\hat{\boldsymbol{\imath}}_r, \hat{\boldsymbol{\imath}}_\theta, \hat{\boldsymbol{\imath}}_h\}$ where $\hat{\boldsymbol{\imath}}_r$ is along the orbit position vector, $\hat{\boldsymbol{\imath}}_h$ is along the orbit momentum vector and $\hat{\boldsymbol{\imath}}_\theta$ is perpendicular to the previous two satisfying the right-hand rule. This frame is often referred to as a Local-Vertical-Local-Horizon (LVLH) reference frame since it tracks the local horizon plane (spanned by $\hat{\boldsymbol{\imath}}_\theta$ and $\hat{\boldsymbol{\imath}}_h$ and the local vertical direction (i.e. $\hat{\boldsymbol{\imath}}_r$). Gauss' variational equations are a specific solution to the orbit element variation equation problem where the position vector $\boldsymbol{r}$, velocity vector $\boldsymbol{v}$ and the disturbance acceleration vector $\boldsymbol{a}_d$ are expressed in $\mathcal{R}$ frame components. The position vector is written as

$$\boldsymbol{r} = r\hat{\boldsymbol{\imath}}_r \tag{11.139}$$

Since the angular velocity of the $\mathcal{R}$ frame relative to the inertial frame $\mathcal{N}$ is $\boldsymbol{\omega}_{\mathcal{R}/\mathcal{N}} = \dot{\theta}\hat{\boldsymbol{\imath}}_h = \dot{f}\hat{\boldsymbol{\imath}}_h$, the velocity vector is expressed as

$$\dot{\boldsymbol{r}} = \dot{r}\hat{\boldsymbol{\imath}}_r + r\dot{f}\hat{\boldsymbol{\imath}}_\theta \tag{11.140}$$

Using the orbit equation in Eq. (8.6) and $h = r^2\dot{f}$, the orbit radius rate is written as

$$\dot{r} = \frac{r}{1+e\cos f}e\sin f\dot{f} = \frac{r^2}{p}e\sin f\dot{f} = \frac{\mu}{h}e\sin f \tag{11.141}$$

The tangential velocity component $r\dot{f}$ is rearranged into the form

$$r\dot{f} = \frac{h}{r} = \frac{\mu p}{hr} \tag{11.142}$$

The velocity vector $\boldsymbol{v}$ is now written as

$$\boldsymbol{v} = \dot{\boldsymbol{r}} = \frac{\mu}{h}\left(e\sin f\hat{\boldsymbol{i}}_r + \frac{p}{r}\hat{\imath}_\theta\right) \tag{11.143}$$

To express the semi-major axis variational equation in terms of LVLH frame components of $\boldsymbol{a}_d$, Eqs. (11.139) and (11.143) are substituted into Eq. (11.99) to yield

$$\frac{da}{dt} = \frac{2a^2}{h}\left(e\sin f a_r + \frac{p}{r}a_\theta\right) \tag{11.144}$$

After substituting Eqs. (11.139) and (11.143) into Eq. (11.105), the eccentricity variation is initially expressed as

$$\frac{de}{dt} = \frac{1}{\mu a e}\left(p\sin f a_r + \frac{(pa - r^2)(1 + e\cos f)}{ae}a_\theta\right) \tag{11.145}$$

After making use of the orbit equation in Eq. (8.6), the eccentricity variation expression is reduced to

$$\frac{de}{dt} = \frac{1}{h}\left(p\sin f a_r + ((p + r)\cos f + re)\, a_\theta\right) \tag{11.146}$$

The variational equations for the longitude of the ascending node Ω and the inclination angle i are obtained trivially from Eqs. (11.115) and (11.116) since $\hat{\boldsymbol{i}}_h^T\boldsymbol{a}_d = a_h$.

$$\frac{d\Omega}{dt} = \frac{r\sin\theta}{h\sin i}a_h \tag{11.147}$$

$$\frac{di}{dt} = \frac{r\cos\theta}{h}a_h \tag{11.148}$$

The variation of the argument of perigee ω is obtained by substituting Eqs. (11.139) and (11.143) into Eq. (11.138) and simplifying

$$\frac{d\omega}{dt} = -\frac{1}{he}\cos f p a_r + \frac{1}{he}(p + r)\sin f a_\theta - \frac{r\sin\theta\cos i}{h\sin i}a_h \tag{11.149}$$

The variational equations of the anomalies f, E and M are obtained by substituting Eqs. (11.139) and (11.143) into Eqs. (11.123), (11.129) and (11.132) respectively.

$$\frac{df}{dt} = \frac{h}{r^2} + \frac{1}{he}\left(p\cos f a_r - (p + r)\sin f a_\theta\right) \tag{11.150}$$

$$\frac{dE}{dt} = \frac{na}{r} + \frac{p}{bhe}\left(a(\cos f - e)a_r + (r + a)\sin f a_\theta\right) \tag{11.151}$$

$$\frac{dM}{dt} = n + \frac{b}{ahe}\left((p\cos f - 2re)a_r - (p + r)\sin f a_\theta\right) \tag{11.152}$$

Gauss' variation equations can be very convenient when the disturbance acceleration $\boldsymbol{a}_d$ is non-conservative. By mapping the acceleration vector into the LVLH frame, the orbit element variations can readily be integrated. Also, if the disturbance $\boldsymbol{a}_d$ is due to a control thrust, Gauss' variational equations show what effect such a control thrust would have on the orbit elements. For example, studying the variational equations of the ascending node Ω and i, it is clear the most efficient period during an orbit to make a nodal correction is during the polar crossing where $\sin\theta$ is maximized. Similarly, the most efficient period to adjust the orbit inclination is during the equator crossing where $\cos\theta$ is maximized. Since these equations are so convenient, they are summarized below one more time.

$$\frac{da}{dt} = \frac{2a^2}{h}\left(e\sin f a_r + \frac{p}{r}a_\theta\right) \tag{11.153a}$$

$$\frac{de}{dt} = \frac{1}{h}\left(p\sin f a_r + \left((p+r)\cos f + re\right)a_\theta\right) \tag{11.153b}$$

$$\frac{di}{dt} = \frac{r\cos\theta}{h}a_h \tag{11.153c}$$

$$\frac{d\Omega}{dt} = \frac{r\sin\theta}{h\sin i}a_h \tag{11.153d}$$

$$\frac{d\omega}{dt} = -\frac{1}{he}\cos f p a_r + \frac{1}{he}(p+r)\sin f a_\theta - \frac{r\sin\theta\cos i}{h\sin i}a_h \tag{11.153e}$$

$$\frac{dM}{dt} = n + \frac{b}{ahe}\left((p\cos f - 2re)a_r - (p+r)\sin f a_\theta\right) \tag{11.153f}$$

11.3 State Transition and Sensitivity Matrix

A state transition matrix $[\Phi(t,t_0)]$ provides a direct mapping from initial conditions $\boldsymbol{r}(t_0)$ to the final state vector $\boldsymbol{r}(t)$ at any particular time. This matrix can be viewed as the sensitivity matrix of the current state to the initial conditions. As such, it has many applications in perturbation theory since it can show, if setup properly, how initial trajectory errors will evolve over time. The state transition matrix, along with the associated sensitivity matrices, are also commonly used in control theory to drive initial trajectory errors to zero. This section presents basic the state transition matrix theory for both linear and nonlinear dynamical systems. Finally, an analytical solution is developed for the special case of Keplerian motion.

11.3.1 Linear Dynamic Systems

Homogeneous System

Let us begin by consideration of the homogeneous (constant coefficients) vector-matrix differential equation case

$$\frac{d\boldsymbol{x}}{dt} = \dot{\boldsymbol{x}} = [A]\boldsymbol{x} \qquad \boldsymbol{x}(t_0) = \boldsymbol{x}_0, \quad [A] = \text{constant} \tag{11.154}$$

where $\boldsymbol{x}(t)$ is a n-dimensional state vector. To establish the form of the general solution, we look at a Taylor's series solution.

$$\boldsymbol{x}(t) = \boldsymbol{x}(t_0) + \sum_{n=1}^{\infty} \left.\frac{d^2\boldsymbol{x}}{dt^n}\right|_{t_0} \frac{(t-t_0)^n}{n!} \tag{11.155}$$

By differentiating Eq. (11.154) and enforcing $[A]$ to be a constant matrix, note that the higher order derivatives of $\boldsymbol{x}(t)$ are expressed as

$$\frac{d^n\boldsymbol{x}(t)}{dt^n} = A^n\boldsymbol{x}(t_0) \tag{11.156}$$

Substituting Eq. (11.156) into the infinite series in Eq. (11.155) yields

$$\boldsymbol{x}(t) = \underbrace{\left([I] + \sum_{n=1}^{\infty} A^n \frac{(t-t_0)^n}{n!}\right)}_{\text{Matrix Exponential}} \boldsymbol{x}(t_0) \tag{11.157}$$

Note that the expression between the large brackets is precisely the definition of the matrix exponential function. Thus we are able to write the general solution for $\boldsymbol{x}(t)$ in the compact form

$$\boldsymbol{x}(t) = e^{[A](t-t_0)}\boldsymbol{x}(t_0) \tag{11.158}$$

Compare this solution to the solution of the scalar linear homogeneous equation

$$\dot{x} = ax \tag{11.159}$$

which has the well-known solution

$$x(t) = e^{a(t-t_0)}x_0 \tag{11.160}$$

Thus, except for the order of the matrix multiplication, the solution of Eq. (11.158) is analogous in many ways to the solution of the scalar case. However, caution must be exercised with this analogy since it isn't perfect.

Continuing with the constant $[A]$ case, let us now introduce a classical result which, if $[A]$ has distinct eigenvalues, transforms the computation of the matrix

exponential $e^{[A](t-t_0)}$ into a trivial exercise. Consider the transformation to a new n-dimensional state vector $\boldsymbol{\eta}$.

$$\boldsymbol{x}(t) = [T]\boldsymbol{\eta}(t) \tag{11.161}$$

where $[T]$ is a constant, non-singular $n \times n$ matrix. Substituting the definition of $\boldsymbol{x}(t)$ into the state differential equations in Eq. (11.154) yields

$$\dot{\boldsymbol{\eta}} = \left([T]^{-1}[A][T]\right)\boldsymbol{\eta} \tag{11.162}$$

The question now is how should the constant matrix $[T]$ be chosen such that the matrix matrix multiplication $[\Lambda] = [T]^{-1}[A][T]$ becomes diagonal? The answer is to chose the columns of the $[T]$ matrix to be the eigenvectors of $[A]$. To prove this, we write out the matrix multiplication as

$$[T]^{-1}[A][T] = [\Lambda] = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & & \lambda_n \end{bmatrix} \tag{11.163}$$

which leads to

$$[A][T] = [T][\Lambda] \tag{11.164}$$

Defining $\boldsymbol{t}_i$ to be the i-th column matrix of $[T]$, Eq. (11.164) is rewritten as a series of n equations.

$$[A]\boldsymbol{t}_i = \lambda_i \boldsymbol{t}_i \qquad i = 1, 2, \cdots, n \tag{11.165}$$

From Eq. (11.165) it is obvious that the matrix diagonal entries λ_i are the eigenvalues of $[A]$ and the columns $\boldsymbol{t}_i$ are the corresponding eigenvectors of $[A]$. Thus this state transformation transforms the originally coupled set of n differential equations into n uncoupled differential equations.

$$\dot{\boldsymbol{\eta}}(t) = \lambda_i \boldsymbol{\eta}(t) \tag{11.166}$$

Using the classical solution of a linear differential equation with constant coefficient in Eq. (11.160), the state vector $\boldsymbol{\eta}(t)$ is computed as

$$\boldsymbol{\eta}(t) = \begin{bmatrix} e^{\lambda_1(t-t_0)} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2(t-t_0)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & & e^{\lambda_2(t-t_0)} \end{bmatrix} \boldsymbol{\eta}(t_0) \tag{11.167}$$

with a matrix exponential expression which is trivial to compute. Substituting the state transformation definition in Eq. (11.161) back into the $\boldsymbol{\eta}(t)$ solution and equating it with the $\boldsymbol{x}(t)$ solution in Eq. (11.158), we are able to compute the complex matrix exponential $e^{[A](t-t_0)}$ through

$$e^{[A](t-t_0)} = [T][\text{diag}(e^{\lambda_i(t-t_0)})][T]^{-1} \tag{11.168}$$

Non-Homogeneous System

Now, having covered the solution to a homogeneous set of differential equations, we consider the more general case of non-homogeneous differential equations where the matrix $[A(t)]$ is time varying.

$$\dot{\boldsymbol{x}}(t) = [A(t)]\boldsymbol{x}(t) \tag{11.169}$$

Going back to the analogy with the scalar case $\dot{x} = a(t)x(t)$ which has the solution

$$x(t) = x(t_0)e^{\int_{t_0}^{t} a(\tau)d\tau} \tag{11.170}$$

we might expect that the solution of Eq. (11.169) be of the form

$$\boldsymbol{x}(t) = e^{\int_{t_0}^{t} [A(\tau)]d\tau}\boldsymbol{x}(t_0) \tag{11.171}$$

But Eq. (11.171) does not hold in general. It only holds for the special cases where either 1) $[A]$ is constant, 2) $[A]$ is diagonal and more generally 3) if $[A]\int[A]d\tau = \int[A]d\tau[A]$. Instead, in order to solve the set non-homogeneous differential equation $\dot{\boldsymbol{x}} = [A(t)]\boldsymbol{x}$, we seek a linear operator $[\Phi(t, t_0)]$ which maps the initial state vector $\boldsymbol{x}_0$ into $\boldsymbol{x}(t)$ as in

$$\boldsymbol{x}(t) = [\Phi(t, t_0)]\boldsymbol{x}(t_0) \qquad [\Phi(t_0, t_0)] = [I_{n\times n}] \tag{11.172}$$

This linear operator $[\Phi(t, t_0)]$ will be referred to as the *state transition matrix*. Substituting the proposed solution of $\boldsymbol{x}(t)$ in Eq. (11.172) back into the differential equation in Eq. (11.169) yields

$$\left([\dot{\Phi}(t, t_0)] - [A(t)][\Phi(t, t_0)]\right)\boldsymbol{x}(t_0) = \mathbf{0} \tag{11.173}$$

Since Eq. (11.173) must hold for any initial condition $\boldsymbol{x}_0$, it is apparent that the state transition matrix differential equation must satisfy

$$[\dot{\Phi}(t, t_0)] = [A(t)][\Phi(t, t_0)] \qquad [\Phi(t, t_0)] = [I_{n\times n}] \tag{11.174}$$

Thus, in the worst case we can solve Eq. (11.174) numerically to determine $[\Phi(t, t_0)]$. Also, notice from Eq. (11.172) that the state transition matrix can be defined as

$$[\Phi(t, t_0)] = \left[\frac{\partial \boldsymbol{x}(t)}{\partial \boldsymbol{x}(t_0)}\right] \tag{11.175}$$

Thus $[\Phi(t, t_0)]$ can be viewed as the sensitivity of the current state vector $\boldsymbol{x}(t)$ to the initial conditions $\boldsymbol{x}(t_0)$.

For the special case where $[A]$ is a constant matrix, we take the partial derivative of Eq. (11.158) with respect to $\boldsymbol{x}_0$ to find the state transition matrix for the homogeneous linear system.

$$[\Phi(t, t_0)] = \left[\frac{\partial \boldsymbol{x}(t)}{\partial \boldsymbol{x}(t_0)}\right] = e^{[A](t-t_0)} \tag{11.176}$$

To verify that this is indeed correct, the state transition matrix $[\Phi(t,t_0)]$ is expanded as a Taylor's series.

$$[\Phi(t,t_0)] = \underbrace{[\Phi(t_0,t_0)]}_{[I_{n\times n}]} + \sum_{n=1}^{\infty} \frac{d^n[\Phi(t,t_0)]}{dt^n}\frac{(t-t_0)^n}{n!} \tag{11.177}$$

Since $[A]$ is a constant matrix for this special case, the higher derivatives of the state transition matrix are given by

$$\frac{d^n[\Phi(t,t_0)]}{dt^n} = [A]^n[\Phi(t,t_0)] \tag{11.178}$$

Substituting these higher derivatives back into Eq. (11.177) yields the expected definition of a state transition matrix for a linear homogeneous system.

$$[\Phi(t,t_0)] = \underbrace{[I_{n\times n}] + \sum_{n=1}^{\infty}[A]^n\frac{(t-t_0)^n}{n!}}_{\text{Matrix Exponential}} = e^{[A](t-t_0)} \tag{11.179}$$

In general, note that the state transition matrix $[\Phi(t_j,t_i)]$ maps the state vector at time t_i to a state vector at time t_j.

$$\boldsymbol{x}(t_j) = [\Phi(t_j,t_i)]\boldsymbol{x}(t_i) \tag{11.180}$$

Thus, given the three times t_1, t_2 and t_3, we are able to write

$$\boldsymbol{x}(t_2) = [\Phi(t_2,t_1)]\boldsymbol{x}(t_1) \tag{11.181}$$
$$\boldsymbol{x}(t_3) = [\Phi(t_3,t_2)]\boldsymbol{x}(t_2) \tag{11.182}$$
$$\boldsymbol{x}(t_3) = [\Phi(t_3,t_1)]\boldsymbol{x}(t_1) = [\Phi(t_3,t_2)][\Phi(t_2,t_1)]\boldsymbol{x}(t_1) \tag{11.183}$$

We conclude that $[\Phi(t_j,t_i)]$ abides by the group property

$$[\Phi(t_k,t_i)] = [\Phi(t_k,t_j)][\Phi(t_j,t_i)] \tag{11.184}$$

Also, note that the inverse of the state transition matrix is simply defined as

$$[\Phi(t_j,t_i)]^{-1} = [\Phi(t_i,t_j)] \tag{11.185}$$

Next, let us consider a linear differential equation with a forcing term $\boldsymbol{u}(t)$.

$$\dot{\boldsymbol{x}} = [A(t)]\boldsymbol{x}(t) + [B(t)]\boldsymbol{u}(t) \tag{11.186}$$

For the special case where $\boldsymbol{u}(t) = \boldsymbol{0}$, we have already established that

$$\boldsymbol{x}(t) = [\Phi(t,t_0)]\boldsymbol{x}(t_0) \tag{11.187a}$$
$$[\dot{\Phi}(t,t_0)] = [A(t)][\Phi(t,t_0)] \qquad [\Phi(t_0,t_0)] = [I_{n\times n}] \tag{11.187b}$$

For $\boldsymbol{u}(t) \neq \boldsymbol{0}$, we seek to replace $\boldsymbol{x}(t_0)$ by a function of time $\boldsymbol{C}(t)$ to make Eq. (11.186) satisfy the forced differential equation in Eq. (11.186). Employing Lagrange's method of variation of parameters, we seek to find out what $\boldsymbol{C}(t)$ will make the solution of Eq. (11.186) have the the form

$$\boldsymbol{x}(t) = [\Phi(t, t_0)]\boldsymbol{C}(t) \qquad \boldsymbol{C}(t_0) = \boldsymbol{x}_0 \tag{11.188}$$

Differentiating Eq. (11.188) and making use of Eq. (11.187b) yields

$$\dot{\boldsymbol{x}}(t) = [A(t)][\Phi(t, t_0)]\boldsymbol{C}(t) + [\Phi(t, t_0)]\dot{\boldsymbol{C}}(t) \tag{11.189}$$

Substituting Eq. (11.187a) into Eq. (11.186) and comparing the resulting $\dot{\boldsymbol{x}}$ expression to Eq. (11.189), the differential equation for $\boldsymbol{C}(t)$ must satisfy

$$\dot{\boldsymbol{C}}(t) = [\Phi(t, t_0)]^{-1}[B(t)]\boldsymbol{u}(t) \tag{11.190}$$

Integrating $\dot{\boldsymbol{C}}$ while making use of $\boldsymbol{C}(t_0) = \boldsymbol{x}_0$ and the state transition matrix inverse property in Eq. (11.185), the function $\boldsymbol{C}(t)$ is now expressed as

$$\boldsymbol{C}(t) = \boldsymbol{x}(t_0) + \int_{t_0}^{t} [\Phi(t_0, \tau)][B(\tau)]\boldsymbol{u}(\tau)d\tau \tag{11.191}$$

Substituting this $\boldsymbol{C}(t)$ function definition back into Eq. (11.188), the state vector $\boldsymbol{x}(t)$ is expressed as a function of the initial conditions and the forcing function $\boldsymbol{u}(t)$ as

$$\boldsymbol{x}(t) = [\Phi(t, t_0)]\boldsymbol{x}(t_0) + \int_{t_0}^{t} [\Phi(t, \tau)][B(\tau)]\boldsymbol{u}(\tau)d\tau \tag{11.192}$$

11.3.2 Nonlinear Dynamic Systems

While the linear systems theory allows us to predict the response of a large class of systems, often the dynamical systems of interest are nonlinear in their nature. For example, the prime differential equation of interest for Keplerian two-body motion is $\ddot{\boldsymbol{r}} = -\mu/r^3\boldsymbol{r}$ which is nonlinear. In the following development we will illustrate how it is still possible to describe time evolution of a state vector in terms of a state transition matrix.

Let us consider the case of a forced nonlinear dynamical system of the form

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(t, \boldsymbol{x}(t), \boldsymbol{p}, \boldsymbol{u}(t)) \qquad \boldsymbol{x}(t_0) = \boldsymbol{x}_0 \tag{11.193}$$

where the n-dimensional vector $\boldsymbol{p}$ is a force model parameter vector. It contains time invariant parameters which affect the effectiveness of the control vector $\boldsymbol{u}$ such as the moment arm for example. Integrating this differential equation yields

$$\boldsymbol{x}(t) = \boldsymbol{x}(t_0) + \int_{t_0}^{t} \boldsymbol{f}(\tau, \boldsymbol{x}(\tau), \boldsymbol{p}, \boldsymbol{u}(\tau))d\tau \tag{11.194}$$

We are interested in the sensitivities of the current state vector $\boldsymbol{x}(t)$ with respect to either the initial state vector $\boldsymbol{x}(t_0)$ or the force model parameter vector $\boldsymbol{p}$.

$$[\Phi(t,t_0)] \equiv \left[\frac{\partial \boldsymbol{x}(t)}{\partial \boldsymbol{x}(t_0)}\right] \tag{11.195}$$

$$[\Psi(t,t_0)] \equiv \left[\frac{\partial \boldsymbol{x}(t)}{\partial \boldsymbol{p}}\right] \tag{11.196}$$

The sensitivity matrix $[\Psi(t,t_0)]$ could be useful in determining how to modify the $\boldsymbol{p}$ vector to enhance the $\boldsymbol{x}(t)$ trajectory. Differentiating Eqs. (11.195) and (11.196) gives

$$[\Phi(t,t_0)] = [I_{n\times n}] + \int_{t_0}^{t} \left[\frac{\partial \boldsymbol{f}()}{\partial \boldsymbol{x}(\tau)}\right]\left[\frac{\partial \boldsymbol{x}(\tau)}{\partial \boldsymbol{x}(t_0)}\right] d\tau \tag{11.197}$$

$$[\Psi(t,t_0)] = [0_{n\times n}] + \int_{t_0}^{t} \left(\left[\frac{\partial \boldsymbol{f}()}{\partial \boldsymbol{x}(\tau)}\right]\left[\frac{\partial \boldsymbol{x}(\tau)}{\partial \boldsymbol{x}(t_0)}\right] + \left[\frac{\partial \boldsymbol{f}()}{\partial \boldsymbol{p}}\right]\right) d\tau \tag{11.198}$$

where $\boldsymbol{f}() = \boldsymbol{f}(\tau, \boldsymbol{x}(\tau), \boldsymbol{p}, \boldsymbol{u}(\tau))$ is implied. Let us introduce the following definitions for this nonlinear system.

$$[A(t)] \equiv \left[\frac{\partial \boldsymbol{f}()}{\partial \boldsymbol{x}(t)}\right] \tag{11.199}$$

$$[C(t)] \equiv \left[\frac{\partial \boldsymbol{f}()}{\partial \boldsymbol{p}}\right] \tag{11.200}$$

Differentiating Eqs. (11.197) and (11.198) with respect to time leads to the following sensitivity matrix time rates:

$$[\dot{\Phi}(t,t_0)] = [A(t)][\Phi(t,t_0)] \quad , \; [\Phi(t_0,t_0)] = [I_{n\times n}] \tag{11.201}$$

$$[\dot{\Psi}(t,t_0)] = [A(t)][\Psi(t,t_0)] + [C(t)] \quad , \; [\Psi(t_0,t_0)] = [0_{n\times n}] \tag{11.202}$$

If it is necessary to compute the solution of $\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(t, \boldsymbol{x}(t), \boldsymbol{p}, \boldsymbol{u}(t))$ numerically, then it is generally also impossible to solve analytically for either $[\Phi(t,t_0)]$ or $[\Psi(t,t_0)]$. One standard algorithm is to simply integrate the three differential equations provided in Eqs. (11.193), (11.201) and (11.202) simultaneously. Now, if the differential equation in Eq. (11.193) can be solved analytically for $\boldsymbol{x}(t)$, then it is usually not necessary to directly solve the differential equations of $[\Phi(t,t_0)]$ and $[\Psi(t,t_0)]$. Rather, we can simply determine these sensitivities by partial differentiation of the analytical solution $\boldsymbol{x}(t)$.

Eqs. (11.201) and (11.202) define the state transition matrix and force parameter $\boldsymbol{p}$ sensitivity matrix for the actual trajectory $\boldsymbol{x}(t)$ of a nonlinear system. However, occasionally it is convenient to study the evolution of deviations $\delta\boldsymbol{x}(t)$ from a reference trajectory $\boldsymbol{x}_{ref}(t)$. This trajectory is generated using a reference initial condition $\boldsymbol{x}_{ref}(t_0)$, along with a reference control vector $\boldsymbol{u}_{ref}(t)$ and reference force model parameter $\boldsymbol{p}_{ref}$. We formally indicate this trajectory as

$$\boldsymbol{x}_{ref}(t) = \boldsymbol{x}_{ref}(t_0) + \int_{t_0}^{t} \boldsymbol{f}(\tau, \boldsymbol{x}_{ref}(\tau), \boldsymbol{p}_{ref}, \boldsymbol{u}_{ref}(\tau)) d\tau \tag{11.203}$$

Let the actual trajectory be provided by the nonlinear equation in Eq. (11.194). Deviations of this trajectory to the reference trajectory are then defined as

$$\delta \boldsymbol{x}(t) = \boldsymbol{x}(t) - \boldsymbol{x}_{ref}(t) \tag{11.204}$$

Using a first order Taylor series expansion of Eq. (11.194) yields

$$\boldsymbol{x}(t) \approx \boldsymbol{x}(t_0) + \int_{t_0}^{t} \Big(\boldsymbol{f}(\tau, \boldsymbol{x}_{ref}(\tau), \boldsymbol{p}_{ref}, \boldsymbol{u}_{ref}(t)) + \left[\frac{\partial \boldsymbol{f}()}{\partial \boldsymbol{x}}\right]_{ref} \delta \boldsymbol{x} + \left[\frac{\partial \boldsymbol{f}()}{\partial \boldsymbol{u}}\right]_{ref} \delta \boldsymbol{u} + \left[\frac{\partial \boldsymbol{f}()}{\partial \boldsymbol{p}}\right]_{ref} \delta \boldsymbol{p} \Big) \, d\tau \tag{11.205}$$

Defining the following partial derivative short hand notations

$$[A(t)] \equiv \left[\frac{\partial \boldsymbol{f}()}{\partial \boldsymbol{x}}\right]_{ref} \tag{11.206}$$

$$[B(t)] \equiv \left[\frac{\partial \boldsymbol{f}()}{\partial \boldsymbol{u}}\right]_{ref} \tag{11.207}$$

$$[C(t)] \equiv \left[\frac{\partial \boldsymbol{f}()}{\partial \boldsymbol{p}}\right]_{ref} \tag{11.208}$$

the trajectory deviations $\delta \boldsymbol{x}(t)$ are approximated as

$$\delta \boldsymbol{x}(t) = \delta \boldsymbol{x}(t_0) + \int_{t_0}^{t} \Big([A(\tau)]\delta \boldsymbol{x} + [B(\tau)]\delta \boldsymbol{u} + [C(\tau)]\delta \boldsymbol{p} \Big) d\tau \tag{11.209}$$

Note that the $\delta \boldsymbol{x}(t)$ derivative is now given by the linear, time dependent differential equation

$$\delta \dot{\boldsymbol{x}}(t) = [A(t)]\delta \boldsymbol{x} + [B(t)]\delta \boldsymbol{u} + [C(t)]\delta \boldsymbol{p} \tag{11.210}$$

Since Eq. (11.210) is of the form assumed in Eq. (11.186), we can write the solution to this differential equation in terms of the state transition matrix as

$$\delta \boldsymbol{x}(t) = [\Phi(t, t_0)]\delta \boldsymbol{x}(t_0) + \int_{t_0}^{t} [\Phi(t, \tau)] \Big([B(\tau)]\delta \boldsymbol{u} + [C(\tau)]\delta \boldsymbol{p} \Big) d\tau \tag{11.211}$$

Differentiating Eq. (11.209) with respect to $\delta \boldsymbol{x}(t_0)$, the state transition matrix is defined for the departure motion as

$$[\Phi(t, t_0)] = \left[\frac{\partial(\delta \boldsymbol{x}(t))}{\partial(\delta \boldsymbol{x}(t_0))}\right] = [I_{n\times n}] + \int_{t_0}^{t} [A(\tau)] \left[\frac{\partial(\delta \boldsymbol{x}(\tau))}{\partial(\delta \boldsymbol{x}(t_0))}\right] d\tau \tag{11.212}$$

Thus the state transition matrix assumes the now familiar form

$$[\dot{\Phi}(t, t_0)] = [A(t)][\Phi(t, t_0)], \qquad [\Phi(t_0, t_0)] = [I_{n\times n}] \tag{11.213}$$

The sensitivity matrix $[\Psi(t,t_0)]$ of the trajectory deviations with respect to the force model parameter vector differences $\delta\boldsymbol{p}$ is given by

$$[\Psi(t,t_0)] = \left[\frac{\partial(\delta\boldsymbol{x}(t))}{\partial(\delta\boldsymbol{p})}\right] = [0_{n\times n}] + \int_{t_0}^{t}\Big([A(\tau)][\Phi(\tau,t_0)] + [C(\tau)]\Big)d\tau \quad (11.214)$$

The differential equation of $[\Psi(t,t_0)]$ again assumes the form

$$[\dot{\Psi}(t,t_0)] = [A(t)][\Phi(t,t_0)] + [C(t)], \qquad [\Psi(t,t_0)] = [0_{n\times n}] \quad (11.215)$$

11.3.3 Symplectic State Transition Matrix

Discussing the direction cosine matrices in Chapter 3 we found that they had the wonderful property of being orthogonal. Thus, their inverse is simply the transpose of the matrix. If the state transition matrix satisfies some specific properties, we can show that it is symplectic. This means that it too will have a simple analytic matrix inverse formula. Let $[\Phi]$ and $[J]$ be a $2n \times 2n$ matrices with $[J]$ being defined as the skew-symmetric matrix

$$[J] = \begin{bmatrix} 0_{n\times n} & I_{n\times n} \\ -I_{n\times n} & 0_{n\times n} \end{bmatrix} \quad (11.216)$$

Note that $[J]^2 = [J][J] = -[I_{2n\times 2n}]$. The matrix $[\Phi]$ is called symplectic if it satisfies the following condition:

$$[\Phi]^T[J][\Phi] = [J] \quad (11.217)$$

The significance of this condition is that if we pre-multiply it by $[J]$ and post-multiply it by $[\Phi]^{-1}$, then the matrix inverse of $[\Phi]$ is given by

$$[\Phi]^{-1} = -[J][\Phi]^T[J] \quad (11.218)$$

Partitioning the matrix $[\Phi]$ into $n \times n$ submatrices Φ_{ij} through

$$[\Phi] = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \quad (11.219)$$

the matrix inverse of $[\Phi]$ is expressed analytically through the simple expression

$$[\Phi]^{-1} = \begin{bmatrix} \Phi_{22}^T & -\Phi_{12}^T \\ -\Phi_{21}^T & \Phi_{11}^T \end{bmatrix} \quad (11.220)$$

The following development will show necessary conditions which will guarantee that the state transition matrix $[\Phi(t,t_0)]$ is symplectic. Let the dynamical system be given through the second order differential equation

$$\ddot{\boldsymbol{r}} = \dot{\boldsymbol{v}} = \boldsymbol{f}(t,\boldsymbol{r}) \quad (11.221)$$

where $\boldsymbol{r}$ is a n-dimensional state vector and $\boldsymbol{v} = \dot{\boldsymbol{r}}$. Let the $2n$ dimensional state vector $\boldsymbol{x}$ be defined as

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{r} \\ \boldsymbol{v} \end{pmatrix} \tag{11.222}$$

Then the second order differential equation in Eq. (11.221) can be written in first order form analogous to Eq. (11.193) as

$$\dot{\boldsymbol{x}} = \boldsymbol{F}(t, \boldsymbol{x}) = \begin{pmatrix} \boldsymbol{v} \\ \boldsymbol{f}(t, \boldsymbol{r}) \end{pmatrix} \tag{11.223}$$

Using Eq. (11.199), the linear state plant matrix $[A(t)]$ is given by

$$[A(t)] = \left[\frac{\partial \boldsymbol{F}(t, \boldsymbol{x}(t))}{\partial \boldsymbol{x}(t)}\right] = \begin{bmatrix} 0_{n\times n} & I_{n\times n} \\ G & 0_{n\times n} \end{bmatrix} \tag{11.224}$$

The $n \times n$ matrix $[G]$ is defined here as

$$[G] = \left[\frac{\partial \boldsymbol{f}(t, \boldsymbol{r})}{\partial \boldsymbol{r}}\right] \tag{11.225}$$

Using the results from Eq. (11.201), the state transition matrix for the dynamical system defined in Eq. (11.223) must satisfy

$$[\dot{\Phi}(t, t_0)] = [A(t)][\Phi(t, t_0)], \qquad [\Phi(t_0, t_0)] = [I_{2n\times 2n}] \tag{11.226}$$

The question now is, under what conditions will this state transition matrix satisfy the symplectic condition shown in Eq. (11.217). By inspection, it is clear that this condition is satisfied at t_0 by the $[\Phi(t_0, t_0)]$ given in Eq. (11.226). To complete the proof that $[\Phi(t, t_0)]$ is symplectic, we must next show that

$$\frac{d}{dt}\left([\Phi(t, t_0)]^T [J][\Phi(t, t_0)]\right) \equiv [0_{2n\times 2n}] \tag{11.227}$$

Substituting the $[\dot{\Phi}(t, t_0)]$ definition in Eq. (11.226), we find that

$$\begin{aligned} &[\dot{\Phi}(t, t_0)]^T [J][\Phi(t, t_0)] + [\Phi(t, t_0)]^T [J][\dot{\Phi}(t, t_0)] \\ &= [\Phi(t, t_0)]^T [A(t)]^T [J][\Phi(t, t_0)] + [\Phi(t, t_0)]^T [J][A(t)][\Phi(t, t_0)] \\ &= [\Phi(t, t_0)]^T \left([A(t)]^T [J] + [J][A(t)]\right) [\Phi(t, t_0)] \\ &= [\Phi(t, t_0)]^T \begin{bmatrix} G - G^T & 0_{n\times n} \\ 0_{n\times n} & 0_{n\times n} \end{bmatrix} [\Phi(t, t_0)] \equiv [0_{2n\times 2n}] \end{aligned} \tag{11.228}$$

Thus, the only condition necessary for the state transition matrix $[\Phi(t, t_0)]$ of the dynamical system given in Eq.(11.223) to be symplectic is that $[G] = [G]^T$ is symmetric.

Example 11.4: Again, let us return to our favorite differential equation describing Keplerian two-body motion.

$$\ddot{\boldsymbol{r}} = -\frac{\mu}{r^3}\boldsymbol{r} = -\nabla V(\boldsymbol{r})$$

This second order differential equation of of the form assumed in Eq. (11.221). Let $\boldsymbol{r} = \boldsymbol{r}(x, y, z)$, then the matrix $[G]$ is expressed as

$$[G] = -\left[\frac{\partial(\nabla V(\boldsymbol{r}))}{\partial \boldsymbol{r}}\right] = -\begin{bmatrix} \frac{\partial^2 V}{\partial x^2} & \frac{\partial^2 V}{\partial x \partial y} & \frac{\partial^2 V}{\partial x \partial z} \\ \frac{\partial^2 V}{\partial x \partial y} & \frac{\partial^2 V}{\partial y^2} & \frac{\partial^2 V}{\partial y \partial z} \\ \frac{\partial^2 V}{\partial x \partial z} & \frac{\partial^2 V}{\partial y \partial z} & \frac{\partial^2 V}{\partial y^2} \end{bmatrix}$$

Since $[G] = [G]^T$, the state transition matrix of the Keplerian 2-body problem is guaranteed to be symplectic and posses the elegant matrix inverse shown in Eq. (11.220). More generally, the symplectic $[\Phi(t, t_0)]$ property is a property of conservative systems and "natural" coordinates.

11.3.4 State Transition Matrix of Keplerian Motion

Since there exists an analytical solution to the two-body problem, it is also possible to develop an analytical solution to the state transition matrix of the two-body problem. Using the F and G solution developed in section 8.4.3, we are able to express the position vector $\boldsymbol{r}(t)$ and velocity vector $\boldsymbol{v}(t)$ as a nonlinear function of time and the initial conditions $\boldsymbol{r}_0 = \boldsymbol{r}(t_0)$ and $\boldsymbol{v}_0 = \boldsymbol{v}(t_0)$.

$$\boldsymbol{x}(t) = \begin{pmatrix} \boldsymbol{r}(t) \\ \boldsymbol{v}(t) \end{pmatrix} = \begin{bmatrix} F \cdot I_{3\times 3} & G \cdot I_{3\times 3} \\ \dot{F} \cdot I_{3\times 3} & \dot{G} \cdot I_{3\times 3} \end{bmatrix} \boldsymbol{x}_0 \tag{11.229}$$

The scalar F and G coefficients, along with their derivatives, are given by

$$F = 1 - \frac{a}{r_0}(1 - \cos\hat{E}) \qquad G = \Delta t + \sqrt{\frac{a^3}{\mu}}(\sin\hat{E} - \hat{E}) \tag{11.230}$$

$$\dot{F} = -\frac{\sqrt{\mu a}}{r r_0}\sin\hat{E} \qquad \dot{G} = 1 + \frac{a}{r}(\cos\hat{E} - 1) \tag{11.231}$$

with $\Delta t = t - t_0$ and $\hat{E} = E - E_0$. The state transition matrix $[\Phi(t, t_0)]$ for this nonlinear system is defined as

$$[\Phi(t, t_0)] = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} = \left[\frac{\partial \boldsymbol{x}(t)}{\partial \boldsymbol{x}_0}\right] \tag{11.232}$$

At first glance, it might appear that $[\Phi(t, t_0)]$ would simply be given by

$$[\Phi(t, t_0)] = \begin{bmatrix} F \cdot I_{3\times 3} & G \cdot I_{3\times 3} \\ \dot{F} \cdot I_{3\times 3} & \dot{G} \cdot I_{3\times 3} \end{bmatrix}$$

for the two-body problem. However, this conclusion neglects the fact that the F and G functions themselves also depend on the initial state vector. Only for linear dynamical systems does the state transition matrix map directly between the initial and current state vector. For a nonlinear system, the state transition matrix provides the sensitivity matrix of the current state vector with respect to the initial condition vector. Thus, to find the state transition matrix for Keplerian motion, we must find the various partial derivatives of the F and G functions with respect to the initial state vectors.

Subdividing the 6×6 state transition matrix into four 3×3 matrices Φ_{ij}, and using the F and G solution to compute the required partial derivatives leads to the following result:

$$\Phi_{11} = \frac{\partial \boldsymbol{r}(t)}{\partial \boldsymbol{r}_0} = F \cdot I_{3\times3} + \boldsymbol{r}_0 \frac{\partial F}{\partial \boldsymbol{r}_0} + \boldsymbol{v}_0 \frac{\partial G}{\partial \boldsymbol{r}_0} \tag{11.233a}$$

$$\Phi_{12} = \frac{\partial \boldsymbol{r}(t)}{\partial \boldsymbol{v}_0} = G \cdot I_{3\times3} + \boldsymbol{r}_0 \frac{\partial F}{\partial \boldsymbol{v}_0} + \boldsymbol{v}_0 \frac{\partial G}{\partial \boldsymbol{v}_0} \tag{11.233b}$$

$$\Phi_{21} = \frac{\partial \boldsymbol{v}(t)}{\partial \boldsymbol{r}_0} = \dot{F} \cdot I_{3\times3} + \boldsymbol{r}_0 \frac{\partial \dot{F}}{\partial \boldsymbol{r}_0} + \boldsymbol{v}_0 \frac{\partial \dot{G}}{\partial \boldsymbol{r}_0} \tag{11.233c}$$

$$\Phi_{22} = \frac{\partial \boldsymbol{v}(t)}{\partial \boldsymbol{v}_0} = \dot{G} \cdot I_{3\times3} + \boldsymbol{r}_0 \frac{\partial \dot{F}}{\partial \boldsymbol{v}_0} + \boldsymbol{v}_0 \frac{\partial \dot{G}}{\partial \boldsymbol{v}_0} \tag{11.233d}$$

Before we tackle the complex partial derivatives of F and G, along with their derivative functions, we develop the partial derivatives of various scalar parameters. The partial derivatives of the initial orbit radius r_0 and initial orbit velocity magnitude v_0 are given by

$$\frac{\partial r_0}{\partial \boldsymbol{r}_0} = \frac{1}{r_0}\boldsymbol{r}_0^T \qquad \frac{\partial r_0}{\partial \boldsymbol{v}_0} = \boldsymbol{0}^T \tag{11.234}$$

$$\frac{\partial v_0}{\partial \boldsymbol{r}_0} = \boldsymbol{0}^T \qquad \frac{\partial v_0}{\partial \boldsymbol{v}_0} = \frac{1}{v_0}\boldsymbol{v}_0^T \tag{11.235}$$

Using the definition of $\sigma_0 \equiv (1/\sqrt{\mu})\boldsymbol{r}_0^T\boldsymbol{v}_0$, the partial derivative of σ_0 with respect to either $\boldsymbol{r}_0$ or $\boldsymbol{v}_0$ is given by

$$\frac{\partial \sigma_0}{\partial \boldsymbol{r}_0} = \frac{1}{\sqrt{\mu}}\boldsymbol{v}_0^T \qquad \frac{\partial \sigma_0}{\partial \boldsymbol{v}_0} = \frac{1}{\sqrt{\mu}}\boldsymbol{r}_0^T \tag{11.236}$$

To find the sensitivities of the semi-major axis a with respect to the initial state vectors, we write the vis-viva equation (Eq. (8.77)) at the initial time t_0 as

$$\frac{1}{a} = \frac{2}{r_0} - \frac{v_0^2}{\mu} \tag{11.237}$$

To simplify the development of the various partial derivatives, we introduce the place-holder vector $\boldsymbol{\alpha}$. This vector can be either $\boldsymbol{r}_0$ or $\boldsymbol{v}_0$. Taking the partial derivative of Eq. (11.237) with respect to the generic $\boldsymbol{\alpha}$ vector yields

$$-\frac{1}{a^2}\frac{\partial a}{\partial \boldsymbol{\alpha}} = -\frac{2}{r_0^2}\frac{\partial r_0}{\partial \boldsymbol{\alpha}} - \frac{2v_0}{\mu}\frac{\partial v_0}{\partial \boldsymbol{\alpha}} \tag{11.238}$$

Substituting Eqs. (11.234) and (11.235), the partial derivatives of a with respect to either $\boldsymbol{r}_0$ or $\boldsymbol{v}_0$ are

$$\frac{\partial a}{\partial \boldsymbol{r}_0} = \frac{2a^2}{r_0^3}\boldsymbol{r}_0^T \qquad\qquad \frac{\partial a}{\partial \boldsymbol{v}_0} = \frac{2a^2}{\mu}\boldsymbol{v}_0^T \tag{11.239}$$

To find the sensitivity of $\hat{E}$ with respect to the initial state vectors, we make use of the modified Kepler's equations derived for the F and G solution in Eq. (8.176).

$$\sqrt{\frac{\mu}{a^3}}\Delta t = \hat{E} + \left(\frac{r_0}{a} - 1\right)\sin\hat{E} + \frac{\sigma_0}{\sqrt{a}}\left(1 - \cos\hat{E}\right) \tag{11.240}$$

Taking the partial derivative of Eq. (11.240) with respect to the generic vector $\boldsymbol{\alpha}$, we find

$$\begin{aligned} -\frac{3}{2}\sqrt{\frac{\mu}{a^5}}\Delta t\frac{\partial a}{\partial \boldsymbol{\alpha}} = &\left(1 + \left(\frac{r_0}{a} - 1\right)\cos\hat{E} + \frac{\sigma_0}{\sqrt{a}}\sin\hat{E}\right)\frac{\partial \hat{E}}{\partial \boldsymbol{\alpha}} \\ &+ \left(\frac{1}{a}\sin\hat{E}\right)\frac{\partial r_0}{\partial \boldsymbol{\alpha}} + \left(-\frac{r_0}{a^2}\sin\hat{E} - \frac{1}{2}\frac{\sigma_0}{a^{3/2}}(1 - \cos\hat{E})\right)\frac{\partial a}{\partial \boldsymbol{\alpha}} \\ &+ \frac{1}{\sqrt{a}}(1 - \cos\hat{E})\frac{\partial \sigma_0}{\partial \boldsymbol{\alpha}} \end{aligned} \tag{11.241}$$

Substituting the orbit radius expression in Eq. (8.172) and solving for the partial derivative of $\hat{E}$, we find

$$\begin{aligned} \frac{\partial \hat{E}}{\partial \boldsymbol{\alpha}} = &\left(-\frac{3}{2}\sqrt{\frac{\mu}{a^3}}\frac{\Delta t}{r} + \frac{r_0}{ra}\sin\hat{E} + \frac{\sigma_0}{2\sqrt{a}r}(1 - \cos\hat{E})\right)\frac{\partial a}{\partial \boldsymbol{\alpha}} \\ &- \frac{\sin\hat{E}}{r}\frac{\partial r_0}{\partial \boldsymbol{\alpha}} - \frac{\sqrt{a}}{r}(1 - \cos\hat{E})\frac{\partial \sigma_0}{\partial \boldsymbol{\alpha}} \end{aligned} \tag{11.242}$$

To find the sensitivity of the orbit radius r at time t, we make use of the orbit radius expression in terms of the eccentric anomaly difference $\hat{E}$ in Eq. (8.172).

$$r = a\left(1 + \left(\frac{r_0}{a} - 1\right)\cos\hat{E} + \frac{\sigma_0}{\sqrt{a}}\sin\hat{E}\right) \tag{11.243}$$

Taking the partial derivative of r with respect to the generic vector $\boldsymbol{\alpha}$ yields

$$\begin{aligned} \frac{\partial r}{\partial \boldsymbol{\alpha}} = &\left(1 - \cos\hat{E} + \frac{1}{2}\frac{\sigma_0}{\sqrt{a}}\sin\hat{E}\right)\frac{\partial a}{\partial \boldsymbol{\alpha}} + \cos\hat{E}\frac{\partial r_0}{\partial \boldsymbol{\alpha}} \\ &\sqrt{a}\sin\hat{E}\frac{\partial \sigma_0}{\partial \boldsymbol{\alpha}} + a\left(\left(1 - \frac{r_0}{a}\right)\sin\hat{E} + \frac{\sigma_0}{\sqrt{a}}\cos\hat{E}\right)\frac{\partial \hat{E}}{\partial \boldsymbol{\alpha}} \end{aligned} \tag{11.244}$$

At this point we are able to compute the partial derivatives of the scalar functions F and G, along with the derivatives of the associate derivative functions, using the intermediate results we just derived. Using Eqs. (11.230) and

(11.231), the sensitivities of the F, G, $\dot{F}$ and $\dot{G}$ scalar parameters to the initial state vectors are expressed as:

$$\frac{\partial F}{\partial \boldsymbol{\alpha}} = \frac{1}{r_0}(1-\cos\hat{E})\left(-\frac{\partial a}{\partial \boldsymbol{\alpha}} + \frac{a}{r_0}\frac{\partial r_0}{\partial \boldsymbol{\alpha}}\right) - \frac{a}{r_0}\sin\hat{E}\frac{\partial \hat{E}}{\partial \boldsymbol{\alpha}} \tag{11.245a}$$

$$\frac{\partial G}{\partial \boldsymbol{\alpha}} = \frac{3}{2}\sqrt{\frac{a}{\mu}}(\sin\hat{E}-\hat{E})\frac{\partial a}{\partial \boldsymbol{\alpha}} - \sqrt{\frac{a^3}{\mu}}(1-\cos\hat{E})\frac{\partial \hat{E}}{\partial \boldsymbol{\alpha}} \tag{11.245b}$$

$$\frac{\partial \dot{F}}{\partial \boldsymbol{\alpha}} = \sqrt{\mu a}\frac{\sin\hat{E}}{r r_0}\left(-\frac{1}{2a}\frac{\partial a}{\partial \boldsymbol{\alpha}} + \frac{1}{r_0}\frac{\partial r_0}{\partial \boldsymbol{\alpha}} + \frac{1}{r}\frac{\partial r}{\partial \boldsymbol{\alpha}} - \cot\hat{E}\frac{\partial \hat{E}}{\partial \boldsymbol{\alpha}}\right) \tag{11.245c}$$

$$\frac{\partial \dot{G}}{\partial \boldsymbol{\alpha}} = -\frac{1}{r}(1-\cos\hat{E})\left(\frac{\partial a}{\partial \boldsymbol{\alpha}} - \frac{a}{r}\frac{\partial r}{\partial \boldsymbol{\alpha}}\right) - \frac{a}{r}\sin\hat{E}\frac{\partial \hat{E}}{\partial \boldsymbol{\alpha}} \tag{11.245d}$$

As was the convention earlier, the vector $\boldsymbol{\alpha}$ in Eq. (11.245) is replaced with either $\boldsymbol{r}_0$ or $\boldsymbol{v}_0$. Substituting the partial derivatives in Eq. (11.245) into Eq. (11.233), an analytical solution to the two-body problem state transition matrix is found.

While the presented analytical method to compute the state transition matrix is relatively straight forward to program, it is not a very compact or elegant solution. Richard Battin develops in Reference 1 an elegant and compact analytical expression of the two-body state transition matrix by performing extensive algebraic simplifications. Defining the scalar parameter C as

$$C = a\sqrt{\frac{a^3}{\mu}}\left(3\sin\hat{E} - (2+\cos\hat{E})\hat{E}\right) - \Delta t a(1-\cos\hat{E}) \tag{11.246}$$

and the position and velocity vector differences as

$$\delta\boldsymbol{r} = \boldsymbol{r} - \boldsymbol{r}_0 \qquad \delta\boldsymbol{v} = \boldsymbol{v} - \boldsymbol{v}_0 \tag{11.247}$$

he is able to express the state transition submatrices Φ_{ij} through

$$\Phi_{11} = \frac{r}{\mu}\delta\boldsymbol{v}\delta\boldsymbol{v}^T + \frac{1}{r_0^3}\left(r_0(1-F)\boldsymbol{r}\boldsymbol{r}_0^T + C\boldsymbol{v}\boldsymbol{r}_0^T\right) + F\cdot I_{3\times 3} \tag{11.248a}$$

$$\Phi_{12} = \frac{r_0}{\mu}(1-F)\left(\delta\boldsymbol{r}\boldsymbol{v}_0^T - \delta\boldsymbol{v}\boldsymbol{r}_0^T\right) + \frac{C}{\mu}\boldsymbol{v}\boldsymbol{v}_0^T + G\cdot I_{3\times 3} \tag{11.248b}$$

$$\begin{aligned}\Phi_{21} = &-\frac{1}{r_0^2}\delta\boldsymbol{v}\boldsymbol{r}_0^T - \frac{1}{r^2}\boldsymbol{r}\delta\boldsymbol{v}^T - \frac{\mu C}{r^3 r_0^3}\boldsymbol{r}\boldsymbol{r}_0^T \\ &+ \dot{F}\left(I_{3\times 3} - \frac{1}{r^2}\boldsymbol{r}\boldsymbol{r}^T + \frac{1}{\mu r}(\boldsymbol{r}\boldsymbol{v}^T - \boldsymbol{v}\boldsymbol{r}^T)\boldsymbol{r}\delta\boldsymbol{v}^T\right)\end{aligned} \tag{11.248c}$$

$$\Phi_{22} = \frac{r_0}{\mu}\delta\boldsymbol{v}\delta\boldsymbol{v}^T + \frac{1}{r^3}\left(r_0(1-F)\boldsymbol{r}\boldsymbol{r}_0^T - C\boldsymbol{r}\boldsymbol{v}_0^T\right) + \dot{G}\cdot I_{3\times 3} \tag{11.248d}$$

with the state transition matrix being defined through

$$[\Phi] = \begin{bmatrix}\Phi_{11} & \Phi_{12}\\ \Phi_{21} & \Phi_{22}\end{bmatrix} = \begin{bmatrix}\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{r}_0} & \frac{\partial \boldsymbol{r}}{\partial \dot{\boldsymbol{r}}_0}\\ \frac{\partial \dot{\boldsymbol{r}}}{\partial \boldsymbol{r}_0} & \frac{\partial \dot{\boldsymbol{r}}}{\partial \dot{\boldsymbol{r}}_0}\end{bmatrix} \tag{11.249}$$

Problems

11.1 Observe that the F and G solution in Eqs. (8.148) and (8.149), using rectangular coordinates, can be written in the form:

$$\begin{pmatrix} x \\ y \\ z \\ \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} = \underbrace{\begin{bmatrix} F & 0 & 0 & G & 0 & 0 \\ 0 & F & 0 & 0 & G & 0 \\ 0 & 0 & F & 0 & 0 & G \\ \dot{F} & 0 & 0 & \dot{G} & 0 & 0 \\ 0 & \dot{F} & 0 & 0 & \dot{G} & 0 \\ 0 & 0 & \dot{F} & 0 & 0 & \dot{G} \end{bmatrix}}_{[\Theta(t,t_0)]} \begin{pmatrix} x_0 \\ y_0 \\ z_0 \\ \dot{x}_0 \\ \dot{y}_0 \\ \dot{z}_0 \end{pmatrix} \tag{11.250}$$

Comparing Eq. (11.250) with the usual state transition matrix form, we conclude that $[\Theta(t,t_0)]$ "looks like" a state transition matrix $[\Phi(t,t_0)]$ in that it maps initial conditions into the instantaneous state. But $[\Phi(t,t_0)]$, as we developed in this chapter, approximately maps the linearized departure motion in the sense

$$\begin{pmatrix} \delta x \\ \delta y \\ \delta z \\ \delta \dot{x} \\ \delta \dot{y} \\ \delta \dot{z} \end{pmatrix} = [\Phi(t,t_0)] \begin{pmatrix} \delta x_0 \\ \delta y_0 \\ \delta z_0 \\ \delta \dot{x}_0 \\ \delta \dot{y}_0 \\ \delta \dot{z}_0 \end{pmatrix}$$

Note that $[\Phi(t,t_0)]$ is fully populated, whereas $[\Theta(t,t_0)]$ has an elegant sparse structure. Establish the relationship between these two matrices.

11.2 Program the state transition matrix computation of the two-body problem using Battin's method in Eqs. (11.246)–(11.249).

Bibliography

[1] Battin, R. H., *An Introduction to the Mathematics and Methods of Astrodynamics*, AIAA Education Series, New York, 1987.

[2] McCuskey, S. W., *Introduction to Celestial Mechanics*, Assison-Wesley Publishing Company, Inc., Reading, MA, 1963.

[3] Brouwer, D., "Solution of the Problem of Artificial Satellite Theory Without Drag," *The Astronautical Journal*, Vol. 64, No. 1274, 1959, pp. 378–397.

Chapter Twelve

Transfer Orbits

PLANET is latin for "wanderer." As seen by an Earth based observer, the planets appear to wander across the sky on smooth, but seemingly arbitrary paths. Aristarchus of Samos (310–230 B.C.) was a Greek astronomer and mathematician who developed a heliocentric universe in which the Earth rotated about the Sun. He even developed an ingenious geometric method to determine the distance between the Sun, Earth and the moon. However, this knowledge got mostly forgotten during the afterward until Nicholas Copernicus (1473–1543 A.D.) rediscovered the fact that the Earth rotated about the Sun. Galileo Galilei (1564–1642 A.D.) later confirmed through the use of his telescope that the Earth revolved about the Sun. Understanding this was crucial to understand the orbits of the planets. Since the planets each have a different mean heliocentric orbit radius, they travel each at different rates. As seen from the rotating Earth frame, observing the planet's trajectories yields interesting and complex geometric paths.

At first glance it would appear very daunting to attempt to travel between planets with their relative trajectories being so complex. Of course, choosing an inertial Sun centered reference frame is the first step toward simplification of the interplanetary motion. Similar considerations govern transfers between Earth centered orbits. This chapter discusses some basic methods used to design interplanetary transfer orbits. Various methods of minimum energy transfer orbits are discussed, as well as the two-point two-body boundary value problem. The method of pathed-conic's is a convenient method to perform preliminary mission analysis. Here the total transfer orbit from one body to another is broken up into various stages where there is only one dominating gravitational influence. The section on patched-conic orbits will discuss issues in developing the interplanetary transfer orbit, as well as issues in designing orbits to escape and enter a planet's sphere of influence. While most theory is typically applied to travel among different solar system planets, it can also be applied to travel between other celestial bodies such as moons or comets.

12.1 Minimum Energy Orbit

Often it is of interest to find a suitable transfer orbit that will connect two points in space. The following development will derive the concept of a minimum energy transfer orbit. As discussed by Battin in Reference 1, this minimum energy orbit becomes a good initial guess at an orbit to start the numerical iteration to find the solution to the two-point boundary value problem. It is also convenient to find other specialized orbit transfers such as the Hohmann transfer.

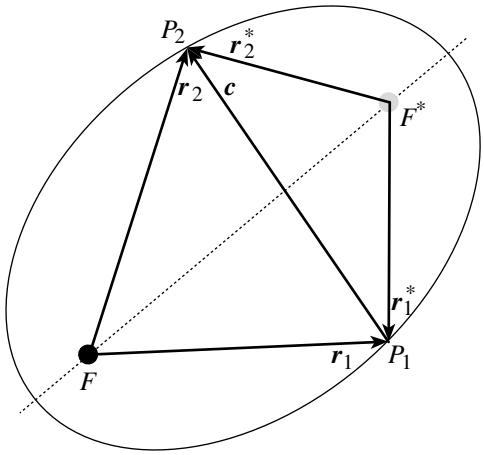


Figure 12.1: Illustration of Lambert's Problem

Consider the general elliptical orbit that connect points P_1 and P_2 shown in Figure 12.1. Let $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ be the corresponding position vectors relative to the occupied focus F, and let $\boldsymbol{r}_1^*$ and $\boldsymbol{r}_2^*$ be the position vectors relative to the un-occupied focus F^*. We begin the development of the minimum energy transfer orbit by recalling a key geometrical property of an ellipse. The sum of the two radial distances from any point on the ellipse to each focal point is constant and equal to $2a$. Thus, we are able to write

$$\underbrace{FP_1}_{r_1} + \underbrace{F^*P_1}_{r_1^*} = 2a \tag{12.1a}$$

$$\underbrace{FP_2}_{r_2} + \underbrace{F^*P_2}_{r_2^*} = 2a \tag{12.1b}$$

where the notation $r_i = |\boldsymbol{r}_i|$ is used. Summing Eqs. (12.1a) and (12.1b) we obtain the following geometric result which must hold for any ellipse:

$$\underbrace{r_1 + r_2}_{\text{fixed}} + r_1^* + r_2^* = 4a \tag{12.2}$$

Since the radial distances r_1 and r_2 are specified by the two-point boundary value problem statement, we are only free to chose the parameters r_1^*, r_2^* and a. Note that the radial distances r_i^* to the un-occupied focus are related to the chord length c through the inequality constraint

$$c \leq r_1^* + r_2^* \tag{12.3}$$

Recall the elliptic orbit energy equation, also referred to the vis-viva equation, given in Eq. (8.77) through

$$\frac{v^2}{2} - \frac{\mu}{r} = E \equiv -\frac{1}{2a} \tag{12.4}$$

where E is the total orbit energy and $v = |\dot{\boldsymbol{r}}|$. Studying Eq. (12.4) it is clear that in order to minimize E, we seek a transfer orbit with the smallest semi-major axis parameter a. Revisiting the elliptic orbit condition in Eq. (12.2), minimizing a means that the sum $r_1^* + r_2^*$ must be minimized. However, according to the constraint in Eq. (12.3), the smallest possible value that the sum $r_1^* + r_2^*$ can achieve is the chord length c. Thus we conclude that the un-occupied focus of the minimum energy transfer orbit must lie on the chord vector $\boldsymbol{c}$. The semi-major axis a_m of this minimum energy orbit is then given by[1]

$$a_m = \frac{1}{4}\left(r_1 + r_2 + c\right) \tag{12.5}$$

Notice that $(r_1 + r_2 + c)$ is the perimeter of the triangle FP_1P_2. Given the initial and final position vectors $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, we can compute c using

$$c = |\boldsymbol{r}_2 - \boldsymbol{r}_1| \tag{12.6}$$

If we know the true anomaly difference Δf between the points P_1 and P_2, then we can use the law of cosines to compute c through the scalar values r_1 and r_2.

$$c = \sqrt{r_2^2 + r_1^2 - 2r_1 r_2 \cos\theta} \tag{12.7}$$

Example 12.1: Let us consider the special minimal energy orbit transfer case where $r_1 = r_2 = r$ as illustrated in Figure 12.2. Note that this problem is essentially the minimal energy ballistic missile problem which was discussed in Example 8.3. Therefore we can assume that we are attempting to fire a projectile at time t_1 and target a point on Earth's surface that is Δf degrees away. For this special case where the initial and final orbit radius are equal (the Earth is assumed to be spherical here), the cord length determined using Eq. (12.7) to be

$$c = r\sqrt{2(1 - \cos\Delta f)}$$

Using the energy equation, the normalized initial velocity magnitude v_0 is given by

$$v_{0_{min}}^2 = 2 - \frac{r}{a_m}$$

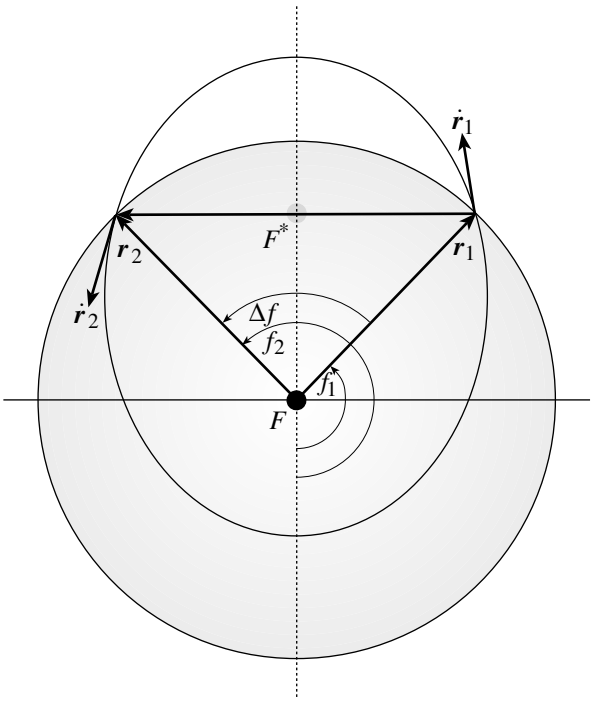


Figure 12.2: Illustration of Lambert's Problem with the $r_1 = r_2$ Constraint

where the velocity magnitude is normalized by $\sqrt{\mu/r}$. Since $1/a_m$ is given by

$$\frac{1}{a_m} = \frac{4}{2r + c}$$

the minimum energy transfer orbit initial velocity v_0 is expressed in terms of the range angle Δf as

$$v_{0_{min}}^2 = 2 - \frac{4}{2 + \sqrt{2(1 - \cos \Delta f)}}$$

Note the two simple special cases where Δf is either 0 or 180 degrees. For the case where $\Delta f = 0$ degrees, we find that v_0 is equal to zero. This makes intuitive sense since it should take no energy for an object to remain at the same location. If $\Delta f = 180$ degrees, then the normalized v_0 becomes 1. This corresponds to the projective being in a circular orbit and just skimming across the surface.

The minimum velocity function developed in the Example 8.3 is expressed in terms of the semi-range angle $\phi = \Delta f/2$.

$$v_{0_{min}}^2 = 2 \tan^2 \phi \left(\frac{1}{|\sin \phi|} - 1 \right)$$

After performing some extensive trigonometric algebra, it can be shown that the two velocity expressions are identical.

12.2 The Hohmann Transfer Orbit

Let us consider the special transfer orbit case where the Δf is equal to 180 degrees. The minimum energy transfer case is illustrated in Figure 12.3. Note that since the un-occupied focus F^* must lie on the cord $\boldsymbol{c}$, the points P_1 and P_2 become either the apoapses or periapses of the transfer orbit, depending on which r_i is larger. This type of minimum energy transfer orbit is commonly referred to as a *Hohmann transfer* orbit.[2, 3] Walter Hohmann (1880–1945) showed in 1925 that an elliptic transfer orbit requires the least Δv to transfer between two circular orbits. His book entitled *Die Erreichbarkeit the Himmelskörper* was a pioneering work that showed how to perform interplanetary travel.

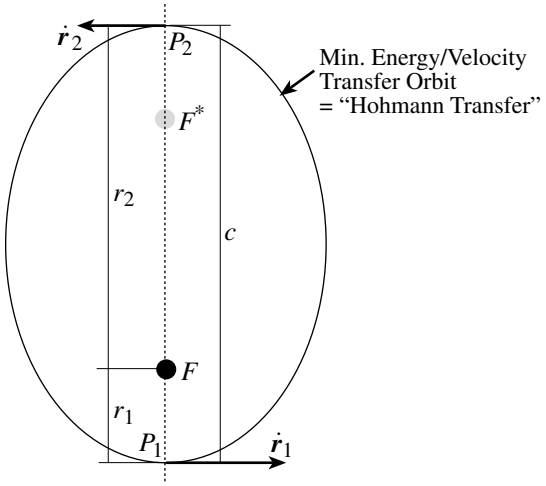


Figure 12.3: Illustration of a Special Case of Lambert's Problem with $\Delta f = 180^{\circ}$

Hohmann transfer orbits are commonly used when increasing or decreasing the radius of a circular orbit. Since $\Delta f = 180^{\circ}$, note that

$$c = r_1 + r_2 \tag{12.8}$$

The minimum energy orbit semi-major axis a_m is then given by

$$a_m = \frac{1}{4}(r_1 + r_2 + c) = \frac{c}{2} \tag{12.9}$$

Using the orbit energy equation and the identity in Eq. (12.8), we are able to write the velocity magnitude v_{m_1} at point P_1 as

$$\begin{aligned} v_{m_1}^2 &= \mu\left(\frac{2}{r_1} - \frac{1}{a_m}\right) \\ &= 2\mu\left(\frac{c - r_1}{r_1 c}\right) \\ &= \frac{2\mu}{c}\left(\frac{r_2}{r_1}\right) \end{aligned} \tag{12.10}$$

Similarly, we are able to write the velocity magnitude v_{m_2} at point P_2 as

$$v_{m_2}^2 = \frac{2\mu}{c}\left(\frac{r_1}{r_2}\right) \tag{12.11}$$

Note that the v_{m_i} expressions in Eqs. (12.10) and (12.11) only depend on the initial and final orbit radii r_1 and r_2 since $c = r_1 + r_2$.

Now consider not just the minimum velocity orbit, but the family of all orbits which pass through the points P_1 and P_2. Further results for the 180° special case are obtained by studying the equation of a conic $p = r(1 + e\cos f)$. Since $f_2 = f_1 + \pi$, we find that

$$p = r_1(1 + e\cos f_1) \tag{12.12a}$$

$$p = r_2(1 - e\cos f_1) \tag{12.12b}$$

Adding Eqs. (12.12a) and (12.12b) we find

$$\frac{p}{r_1} + \frac{p}{r_2} = 2 \tag{12.13}$$

Solving for the semi-latus rectum p yields

$$p = \frac{2r_1 r_2}{r_1 + r_2} = \frac{2r_1 r_2}{c} \tag{12.14}$$

which states that p is the harmonic mean of the initial and final orbit radius. Note that the above equation only holds for the special case being considered where $\Delta f = 180°$.

Next we investigate the radial and tangential velocity components. Let the position vector be given by $\boldsymbol{r} = r\hat{\boldsymbol{\imath}}_r$, the velocity vector $\boldsymbol{v}$ is given by

$$\boldsymbol{v} = \dot{\boldsymbol{r}} = \dot{r}\hat{\boldsymbol{\imath}}_r + r\dot{\theta}\hat{\boldsymbol{\imath}}_\theta = v_r\hat{\boldsymbol{\imath}}_r + v_\theta\hat{\boldsymbol{\imath}}_\theta \tag{12.15}$$

where $\dot{\theta}$ is the true latitude rate. Using the angular momentum definition $h = \sqrt{\mu p} = r^2\dot{\theta}$, we write the transverse velocity magnitude v_θ for any orbit transfer angle Δf as

$$v_{\theta_1}^2 = r_1^2\dot{\theta}_1^2 = \frac{r_1^4\dot{\theta}_1^2}{r_1^2} = \frac{h^2}{r_1^2} = \frac{\mu p}{r_1^2} \tag{12.16}$$

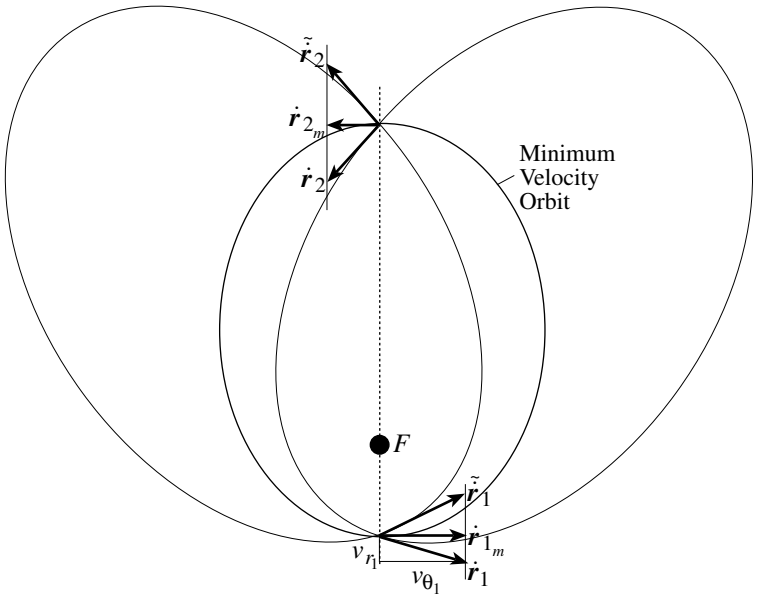


Figure 12.4: Possible Orbit Solutions to Lambert Problem with $\Delta f = 180^{\circ}$

Specializing the transverse velocity expression for the special case of having $\Delta f = 180^{\circ}$, we substitute the semi-latus rectum definition in Eq. (12.14) into Eq. (12.16) to find

$$v_{\theta_1}^2 = \frac{2\mu}{c}\left(\frac{r_2}{r_1}\right) \qquad \text{for} \quad \Delta f = 180^{\circ} \tag{12.17}$$

Note that Eq. (12.17) must hold true for *any* orbit that connects the given points P_1 and P_2 which are $\Delta f = 180^{\circ}$ apart. The transverse velocity magnitude only depends on the orbit radii r_1 and r_2. Comparing v_{θ_1} in Eq. (12.17) to the minimum energy orbit velocity magnitude v_{m_1} in Eq. (12.10), we find that

$$v_{\theta_1}^2 \equiv v_{m_1}^2 \tag{12.18}$$

Similarly we find

$$v_{\theta_2}^2 = \frac{2\mu}{c}\left(\frac{r_1}{r_2}\right) \equiv v_{m_2}^2 \tag{12.19}$$

Thus we can conclude that for the special case of having $\Delta f = 180^{\circ}$, that all transfer orbits passing through P_1 and P_2 have the same transverse velocity v_θ.

Next we investigate the radial take-off and arrival velocity v_{r_1} and v_{r_2}. Using Eq. (11.141), we are able to write the general orbit radial speed as

$$v_{r_i} = \dot{r}_i = \frac{he \sin f_1}{p} \tag{12.20}$$

Adding the radial velocities at P_1 and P_2 we find

$$v_{r_1} + v_{r_2} = \frac{he}{p}\left(\sin f_1 + \sin(f_1 + \pi)\right) \tag{12.21}$$

Since $\sin(f_1 + \pi) = -\sin f_1$, the above equation reduces to the simple relationship

$$v_{r_1} = -v_{r_2} \qquad \text{for } \Delta f = 180^{\circ} \tag{12.22}$$

From this equation we can conclude that the radial take-off and arrival speeds will have equal magnitude and opposite sign for *all* transfer orbits.

This very elegant take-off and arrival behavior for the $\Delta f = 180^{\circ}$ special case is illustrated in Figure 12.4. Note that

$$|\dot{\boldsymbol{r}}_i| = v_{\theta_i}$$

is only true if the transfer orbit is a minimum energy transfer orbit. The locus of the take-off and arrival velocity vectors is a straight line. From the energy equation we find that

$$v_i^2 = v_{r_i}^2 + v_{\theta_i}^2 = \mu\left(\frac{2}{r_i} - \frac{1}{a}\right) \tag{12.23}$$

Substituting $v_{\theta_i} = v_{m_i}$, the radial take-off and arrival speeds are expressed for general transfer orbits as

$$v_{r_i}^2 = \mu\left(\frac{2}{c} - \frac{1}{a}\right) \tag{12.24}$$

Note the elegant similarity of the above equation to the energy equation. If the minimum energy transfer orbit semi-major axis $a_m = c/2$ is chosen (which is specialized for the $\Delta f = 180^{\circ}$ case), then the v_{r_i} expressions become zero again. This repeats the above conclusion that a minimum energy take-off and arrival velocity will only have transverse components.

Example 12.2: Let us consider the case where a circular orbit of radius r_1 is to be boosted to a higher orbit or radius r_2. The minimum energy transfer orbit is a Hohmann orbit where $\Delta f = 180^{\circ}$. The initial and final circular orbits, as well as the Hohmann transfer orbit, is illustrated in Figure 12.5. The orbital velocities v_1 and v_2 of the respective initial and final circular orbits are given through the energy equation in Eq. (8.82) as

$$v_i = \sqrt{\frac{\mu}{r_i}}$$

Note that for this case the chord is given by $c = r_1 + r_2$. Using the minimum energy transfer orbit to reach the higher circular orbit, the transverse velocity v_{θ_1} that a spacecraft must have at point P_1 is given by Eq. (12.17).

$$v_{\theta_1} = \sqrt{\frac{2\mu}{c}\left(\frac{r_2}{r_1}\right)} = v_1\sqrt{\frac{2r_2}{c}}$$

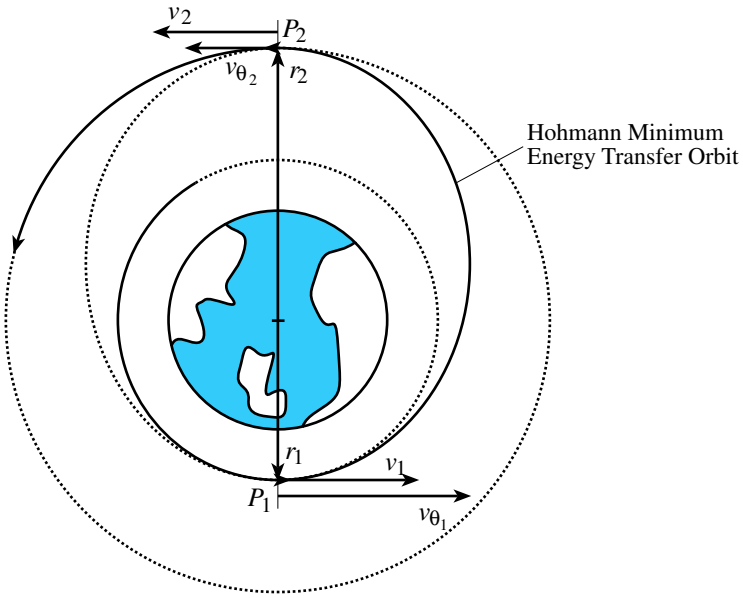


Figure 12.5: Illustration of a Hohmann Transfer Orbit

Being a minimum energy transfer orbit, no radial velocity will be present at this point. Thus, for the spacecraft leave its circular orbit of radius r_1 and enter the elliptic transfer orbit of semi-major axis $a_m = c/2$, a change in velocity Δv_1 is required.

$$\Delta v_1 = v_{\theta_1} - v_1 = v_1 \left(\sqrt{\frac{2r_2}{c}} - 1 \right) \tag{12.25}$$

Note that since $r_2 > r_1$ that Δv_1 will be positive. Once the spacecraft reaches point P_2, a second burn will be required to correct the orbit velocity. The spacecraft will have the velocity v_{θ_2} of the transfer orbit given by

$$v_{\theta_2} = \sqrt{\frac{2\mu}{c} \left(\frac{r_1}{r_2} \right)} = v_2 \sqrt{\frac{2r_1}{c}}$$

The change in velocity Δv_2 of the second burn to enter the higher circular orbit is computed through

$$\Delta v_2 = v_2 - v_{\theta_2} = v_2 \left(1 - \sqrt{\frac{2r_1}{c}} \right) \tag{12.26}$$

Again, note that this second change in velocity is magnitude is also positive since $r_2 > r_1$. The cost of an orbit transfer is typically given in terms of the sum of all required changes in velocities (i.e. burns). For this maneuver, the total Δv is computed using

$$\Delta v = |\Delta v_1| + |\Delta v_2|$$

Assume the initial orbit had a radius of $r_1 = 7000$ km, and the final circular orbit has a radius of $r_2 = 7200$ km. The first burn to enter the elliptic Hohmann transfer orbit would require $\Delta v_1 = 52.3541$ m/s, while the second burn would require a Δv_2 of 51.9867 m/s. The total cost for raising the orbit radius of the circular orbit is $\Delta v = 104.351$ m/s.

12.3 Lambert's Problem

The two-point boundary value problem of the two-body problem is a classic celestial mechanics challenge that was first stated and solved by Johann Heinrich Lambert (1728–1779). The goal is to find an orbit which connects to points in space with a given flight time. Today this problem is commonly solved when controlling and targeting spacecraft or directing missiles in an inverse-square gravity field.

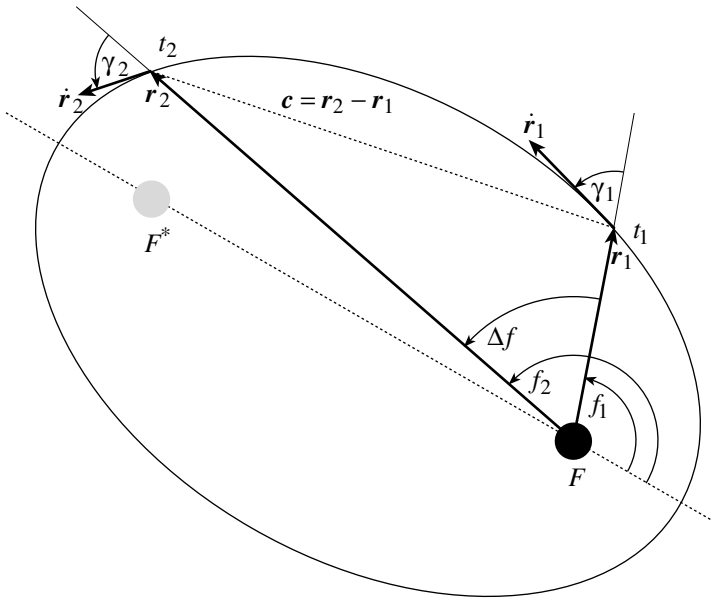


Figure 12.6: Illustration of the Two-Point Boundary Value Problem of the Two-Body Problem

The geometry of Lambert's Problem is illustrated in Figure 12.6. The initial position vector is given by $\boldsymbol{r}_1$ and the final position vector is $\boldsymbol{r}_2$. All position vectors are measured relative to the source of the dominant inverse gravity field. As such, this source will be the focus F of any elliptic orbit that will connect position $\boldsymbol{r}_1$ to position $\boldsymbol{r}_2$. For example, if the objective is to travel between two planets, then the gravity source during the interplanetary flight would be the Sun. The unoccupied focus is denoted by F^*. The true anomalies f_i are

measured relative to periapses, with Δf being the angular change between the initial and final position vectors. The chord vector $\boldsymbol{c}$ is relative vector between the final and initial positions. Given $\boldsymbol{r}_1$, $\boldsymbol{r}_2$ and a time of flight $\Delta t = t_2 - t_1$, Lambert's problem seeks to find the orbit which will connect the two positions at the given times.

To solve this problem, this section presents a general numerical iteration technique to solve this two-point boundary value problem. This method applies to un-perturbed gravity field case, as well as the perturbed gravity field case. The reader is referred to Battin's chapter on Lambert's problem in Reference 1, where he develops a very elegant analytical solution to Lambert's problem. However, this solution is only applicable if there are no perturbations to the inverse-square gravity field. The general solution presented in this chapter makes use of the state transition matrix and can be applied even if gravitational perturbations are present. To start the iteration, the elegant and convenient concept of the minimum energy orbits is employed. As suggested by Battin, the convergence and stability of the numerical iteration is improved if this orbit is used as the initial solution guess of the two-point boundary value problem. The solution to Lambert's problem is full of very elegant and beautiful properties. This section will discuss two some elegant properties of the departure and arrival velocity vectors of the two-point two-body boundary value problem.

12.3.1 General Problem Solution

A common method to solve a general two-point boundary value problem is to employ a numerical iteration technique called the "shooting-method". Given the initial and final states $x(t_1)$ and $x(t_2)$, as well as a desired transfer time Δt, the shooting method technique starts out with a guess of the initial velocity $\dot{x}(t_1)$. After integrating the trajectory to obtain the state $\hat{x}(t_2)$, the final targeting error

$$\delta x(t_2) = \hat{x}(t_2) - x(t_2) \tag{12.27}$$

is computed. Using the sensitivity of the final position to the initial velocity, the initial velocity estimate is updated using the local-linearization based Newton method:

$$\underbrace{\dot{x}(t_1)}_{new} = \underbrace{\dot{x}(t_1)}_{old} - \left(\frac{\partial x(t_2)}{\partial \dot{x}(t_1)}\right)^{-1} \delta x(t_2) \tag{12.28}$$

The initial velocity is successively updated until the target error $\delta x(t_2)$ has become sufficiently small. The success of this iterative technique depends on the nonlinearity of the governing differential equations, as well as the quality of the initial guess. If the guess is very poor and the problem is highly nonlinear, then the shooting method may not converge to the true answer. The more linear the problem is, the more accurate the initial velocity updates in Eq. (12.28) will be and the more successful the application of the shooting method will be to solve the two-point boundary value problem.

To solve Lambert's problem, a slightly modified version of the shooting method is proposed since the governing differential equations are relatively sensitive to the initial velocity vector guess. A good estimate of this vector is required for the standard shooting method to converge. The problem statement provides us with the initial position vector $\boldsymbol{r}_1$ and the desired position vector $\boldsymbol{r}_2$ with a flight time of Δt. To start the numerical iteration, we don't simply choose an arbitrary initial velocity vector $\dot{\boldsymbol{r}}_1$. Instead, we choose a velocity vector which results in a motion that will reach point $\boldsymbol{r}_2$ precisely, though not necessarily at the desired time t_2. We will show how to construct such a solution using a minimum energy orbit transfer. This will provide us with the required initial velocity vector. Note that in principle any transfer orbit could be used. It is not required to use the minimum energy orbit as the initial guess. However, as pointed out by Battin in Reference 1, the convergence and stability of solving Lambert's problem is increased if this minimum energy transfer orbit is used as the initial guess.

A flow diagram of the modified shooting method to solve Lambert's problem is shown in Figure 12.7. This continuation method is often used in numerical iterations. Assume the initial velocity estimate $\dot{\tilde{\boldsymbol{r}}}_1$ results in a transfer time of $\Delta\tilde{t}$ and zero final tracking error. Whereas the standard shooting method attempts to solve the two-point boundary value problem in one go, the modified shooting method employed here solves a series of neighboring two-point boundary value problems which gradually lead to the desired answer. The benefit here is greatly increased convergence and stability of the numerical iteration method. The transfer time $\Delta\hat{t}$ of each successive two-point boundary value problem is swept linearly from the initial transfer time of $\Delta\tilde{t}$ to the desired transfer time of Δt. The initial guess will result in no final tracking error. As the transfer time $\Delta\hat{t}$ slowly approaches the desired transfer time Δt, the initial velocity vector $\dot{\boldsymbol{r}}_1$ is iteratively adjusted such that the final tracking errors remain within a certain tolerance ϵ. By starting out with a set of states $(\dot{\tilde{\boldsymbol{r}}}_1, \Delta\tilde{t})$ that do reach the desired $\boldsymbol{r}_2$, and then gradually adjusting the flight time to the desired flight time, the numerical iteration technique will never encounter very large final tracking errors. This makes the state transition matrix computed initial velocity corrections more accurate and the overall convergence time is decreased.

To compute the state transition matrix $[\partial\hat{\boldsymbol{r}}_2/\partial\dot{\boldsymbol{r}}_1]$, the analytical solution given in Eq. (11.248c) could be used. This equation is valid if forces are perturbing the two-body Keplerian motion. If large perturbations are present, then this state transition matrix should be computed using standard numerical techniques.

To find a initial velocity vector which results in an orbit that connects the position vectors $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ precisely, we recall the Lagrange F and G solution to the orbit motion in terms of a true anomaly difference Δf. Using Eqs. (8.148) and (8.177) we are able to relate the velocity vectors $\dot{\boldsymbol{r}}_1$ and $\dot{\boldsymbol{r}}_2$ to the *specified* position vectors $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ through

$$\boldsymbol{r}_2 = F\boldsymbol{r}_1 + G\dot{\boldsymbol{r}}_1 \tag{12.29}$$

$$\boldsymbol{r}_1 = \dot{G}\boldsymbol{r}_2 - G\dot{\boldsymbol{r}}_2 \tag{12.30}$$

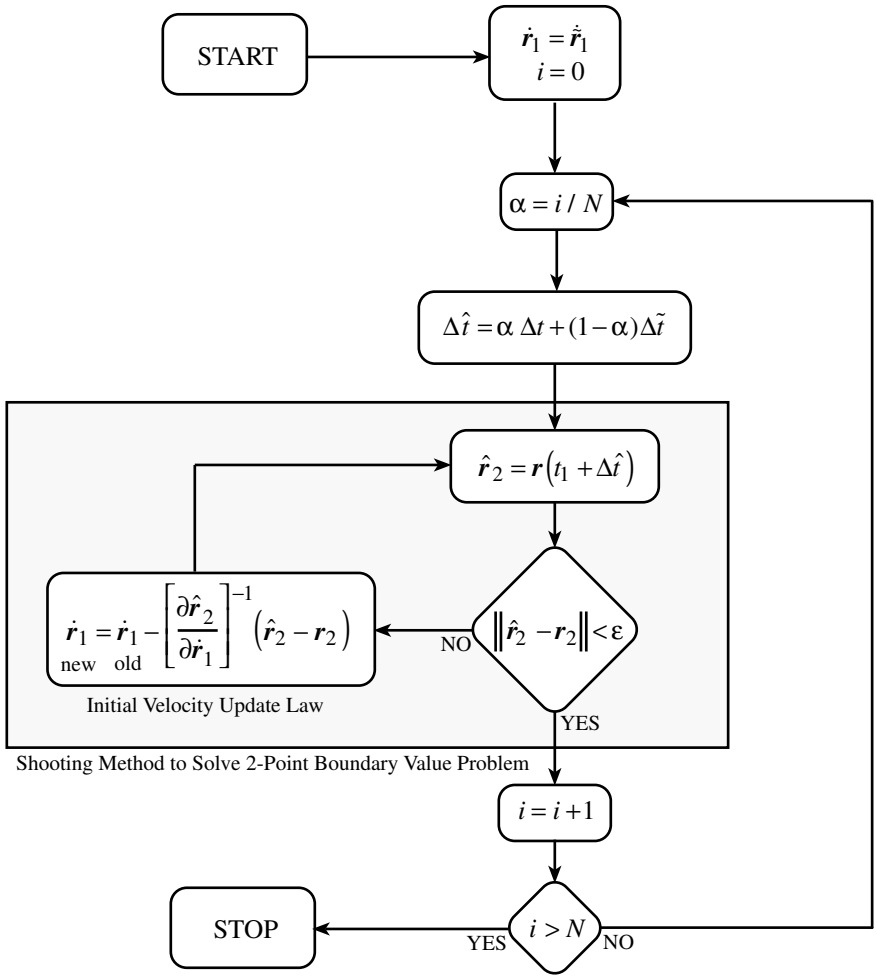


Figure 12.7: Flow Diagram of the Modified Shooting-Method to Solve Lambert's Problem

Solving these two equations for the velocity vectors of interests yields

$$\dot{\boldsymbol{r}}_1 = \frac{1}{G}\left(\boldsymbol{r}_2 - F\boldsymbol{r}_1\right) \tag{12.31}$$

$$\dot{\boldsymbol{r}}_2 = \frac{1}{G}\left(-\boldsymbol{r}_1 + \dot{G}\boldsymbol{r}_2\right) \tag{12.32}$$

with the functions F, G and $\dot{G}$ being defined in Eqs. (8.183), (8.184) and (8.187) as

$$F = 1 - \frac{r_2}{p}\left(1 - \cos\Delta f\right) \tag{12.33}$$

$$G = \frac{r_1 r_2}{\sqrt{\mu p}}\sin\Delta f \tag{12.34}$$

$$\dot{G} = 1 - \frac{r_1}{p}\left(1 - \cos\Delta f\right) \tag{12.35}$$

Substituting Eqs. (12.33) - (12.35) into the velocity vector expressions in Eqs. (12.31) and (12.32) yields the desired initial and final velocity expressions

$$\dot{\boldsymbol{r}}_1 = \frac{\sqrt{\mu p}}{r_1 r_2 \sin\Delta f}\left[\boldsymbol{c} + \frac{r_2}{p}\left(1 - \cos\Delta f\right)\boldsymbol{r}_1\right] \tag{12.36}$$

$$\dot{\boldsymbol{r}}_2 = \frac{\sqrt{\mu p}}{r_1 r_2 \sin\Delta f}\left[\boldsymbol{c} + \frac{r_1}{p}\left(1 - \cos\Delta f\right)\boldsymbol{r}_2\right] \tag{12.37}$$

where $\boldsymbol{c} = \boldsymbol{r}_2 - \boldsymbol{r}_1$ is the chord vector. Eq. (12.36) provides us with a convenient method to compute an initial velocity vector guess which will generate an orbit that connects the two position vectors $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. However, this orbit will generally not have the required transfer time. Note that the only unknown parameter in Eq. (12.36) is the parameter p. All other parameters are specified through the problem statement. For the non-perturbed Keplerian orbit case, solving Lambert's problem can be reduced to a one-dimensional search for an orbit with the proper energy such that the desired transfer time is achieved. The transfer orbit plane itself is defined by the vectors $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. Battin presents in Reference 1 Lagrange's classical solution to this two-point boundary value problem which involves a one-dimensional search for the proper transfer orbit semi-major axis. The development of this elegant solution to the unperturbed Keplerian motion case is not covered in this chapter.

While Eq. (12.36) will yield a suitable initial velocity vector guess $\dot{\tilde{\boldsymbol{r}}}_1$, we still require the associated transfer time $\Delta\tilde{t}$. Given the orbit equation in Eq. (8.6), we are able to establish the following two identities at times t_1 and t_2:

$$e\cos f_1 = \frac{p}{r_1} - 1 = \frac{a(1-e^2)}{r_1} - 1 \tag{12.38}$$

$$e\underbrace{\left(\cos f_1\cos\Delta f - \sin f_1\sin\Delta f\right)}_{\cos f_2} = \frac{p}{r_2} - 1 = \frac{a(1-e^2)}{r_2} - 1 \tag{12.39}$$

These two equations contain the three unknown parameters e, f_1 and p. If $p = a(1-e^2)$ is used, then the semi-major axis a replaces the semi-latus rectum p as a free parameter. By choosing either an initial $\tilde{p}$ or $\tilde{a}$, the two equations in Eqs. (12.38) and (12.39) must be solved simultaneously for the corresponding orbit eccentricity $\tilde{e}$ and initial true anomaly $\tilde{f}_1$. Given $\tilde{f}_1$ and $\tilde{f}_2 = f_1 + \Delta f$, as well as the eccentricity $\tilde{e}$, we are able to compute the corresponding eccentric and mean anomalies. The transfer time of the initial transfer orbit guess is then given through Kepler's equation:

$$\Delta\tilde{t} = \sqrt{\frac{\tilde{a}^3}{\mu}}\left(\tilde{M}_2 - \tilde{M}_1\right) \tag{12.40}$$

To use the minimum energy orbit as an initial guess for solving Lambert's problem, we first compute a_m and then solve Eqs. (12.38) and (12.39) for the transfer orbit initial true anomaly f_1 and eccentricity e. With these parameters, we are then able to compute the required initial velocity vector $\dot{\boldsymbol{r}}_1$ using Eq. (12.36) and the minimum energy orbit transfer time $\Delta\tilde{t}$ using Eq. (12.40).

This completes the required steps to solve Lambert's two-point boundary value problem of the two-body problem using the modified shooting method. While this method works with a standard two-body problem, it will also work for more general boundary value problems where the gravity field contains some perturbations.

12.3.2 Elegant Velocity Properties

In Eqs. (12.36) and (12.37) we developed the necessary orbit velocities such that the corresponding transfer orbit will contain the points P_1 and P_2. Some very elegant geometric interpolations can be arrived at by investigating these velocity vectors in the non-orthogonal coordinate system $(\hat{\boldsymbol{\imath}}_c, \hat{\boldsymbol{\imath}}_{r_i})$, where $\hat{\boldsymbol{\imath}}_c$ is the unit vector of the chord vector and $\hat{\boldsymbol{\imath}}_{r_i}$ is the unit vector of either the initial or final orbit radius vector. This coordinate system is illustrated in Figure 12.8.

Using these unit coordinates, the initial and final velocity vector of any transfer orbit is written in the compact form

$$\dot{\boldsymbol{r}}_1 = v_\rho \hat{\boldsymbol{\imath}}_{r_1} + v_c \hat{\boldsymbol{\imath}}_c \tag{12.41}$$

$$\dot{\boldsymbol{r}}_2 = -v_\rho \hat{\boldsymbol{\imath}}_{r_2} + v_c \hat{\boldsymbol{\imath}}_c \tag{12.42}$$

with the velocity vector component v_ρ along the orbit radius being given by

$$v_\rho = \sqrt{\frac{\mu}{p}}\left(\frac{1-\cos\Delta f}{\sin\Delta f}\right) \tag{12.43}$$

and the vector component v_c along the chord c being given by

$$v_c = \frac{c\sqrt{\mu p}}{r_1 r_2 \sin\Delta f} \tag{12.44}$$

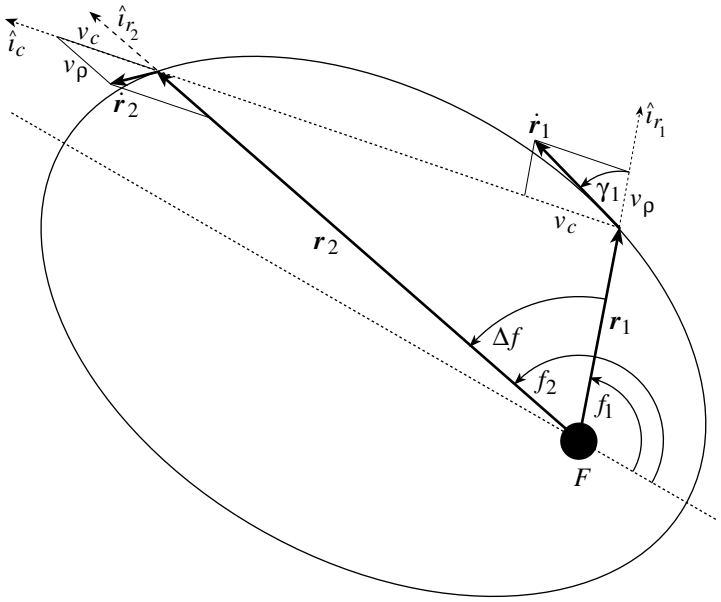


Figure 12.8: Geometric Interpretation of the Solution to the Two-Point Boundary Value Problem of the Two-Body Problem

Note that v_c must be equal for any orbits that connects the points P_1 and P_2. The initial and final velocity vector components along the initial and final orbit position vectors have the same magnitude, but opposite sign.

To investigate on what curve all possible take-off and arrival velocity vectors will lie, we look at the product of v_ρ and v_c:[1]

$$v_c v_\rho = \frac{\mu c}{r_1 r_2} \left(\frac{1 - \cos \Delta f}{\sin \Delta f} \right) \tag{12.45}$$

Using several trigonometric identities, this product can be written in the compact form

$$v_c v_\rho = \frac{\mu c}{2 r_1 r_2} \sec^2 \left(\frac{\Delta f}{2} \right) \tag{12.46}$$

Note that the right hand side of this product depends solely upon the triangle formed by the position vectors $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. Thus, all infinity of orbits passing through P_1 and P_2 have the property that $v_c v_\rho = (\text{const})^2$. At first glance this property may not be recognized. However, Eq. (12.46) is the equation of a hyperbola in asymptotic coordinates.

Let us digress briefly and study the hyperbola expressed in terms of asymptotic coordinates to establish this truth. Figure 12.9 shows a hyperbola with coordinates (x, y) relative to an orthogonal base vector set $\{\hat{\boldsymbol{e}}_x, \hat{\boldsymbol{e}}_y\}$ and with asymptotic coordinates (X, Y) relative to the non-orthogonal base vector set

$\{\hat{\boldsymbol{e}}_X, \hat{\boldsymbol{e}}_Y\}$. Note that the hyperbola will asymptotically approach the direction vectors $\hat{\boldsymbol{e}}_X$ and $\hat{\boldsymbol{e}}_Y$, with ψ being the slope angle of the hyperbola asymptote relative to the $\hat{\boldsymbol{e}}_x$ axis.

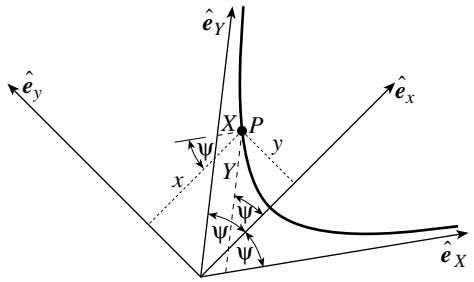


Figure 12.9: Equation of a Hyperbola in Asymptotic Coordinates

Studying Figure 12.9 carefully, we find the relationship between the orthogonal and non-orthogonal coordinates to be

$$x = (X + Y)\cos\psi \tag{12.47a}$$

$$y = (Y - X)\sin\psi \tag{12.47b}$$

The equation of a hyperbola in terms of orthogonal coordinates (x, y) is

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \tag{12.48}$$

where the hyperbolic semi-major axis a is defined to be a negative quantity and $b = a\sqrt{1 - e^2}$. Substituting Eqs. (12.47a) and (12.47b) into Eq. (12.48) we find

$$\frac{(X^2 + 2XY + Y^2)\cos^2\psi}{a^2} - \frac{(X^2 - 2XY + Y^2)\sin^2\psi}{b^2} = 1 \tag{12.49}$$

In Eq. (8.26) we found the useful relationship

$$b^2 = a^2\tan^2\psi \tag{12.50}$$

Using this identity, the equation of a hyperbola is written in terms of the asymptotic coordinates (X, Y) as

$$XY = \frac{a^2}{4}\sec^2\psi = \text{constant} \tag{12.51}$$

Comparing Eqs. (12.46) and (12.51), it is apparent that the product $v_c v_\rho$ indeed describes a hyperbola in terms of asymptotic coordinates with the hyperbolic semi-major axis being

$$a = \sqrt{\frac{2\mu c}{r_1 r_2}} \tag{12.52}$$

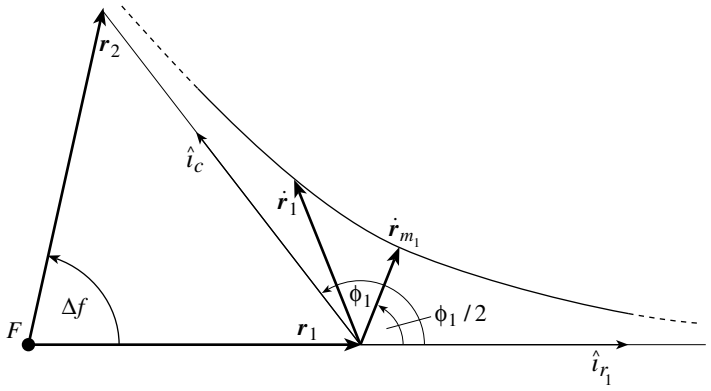


Figure 12.10: Illustration of the Velocity Vector Locus

The geometric interpretation of this hyperbola is shown in Figure 12.10. Note that the loci of all velocity vectors leaving point P_1 going through point P_2 lie along a hyperbola. Additionally, this hyperbola is completely established by the triangle FP_1P_2. Studying Figure 12.10 the minimum velocity orbit is easily found to be such that

$$v_\rho = v_c \tag{12.53}$$

Also, the direction of the minimum velocity vector $\dot{\boldsymbol{r}}_{m_1}$ is such that its unit vector bisects the angle ϕ_1.

Compare the general minimum velocity transfer orbit condition in Eq. (12.53) to the minimum velocity orbit conditions for the special case where Δf is precisely 180 degrees. As $\Delta f \to 180^\circ$, the hyperbola of take-off velocity vectors loci becomes a rectilinear curve collinear aligned with the $\boldsymbol{r}_1$ position vector. Simultaneously the asymptotic velocity coordinates v_c and v_ρ will go to zero for the minium velocity solution. This agrees with the $\Delta f = 180^\circ$ special case which concluded that a corresponding minimum velocity orbit would have zero radial take-off velocity. Similarly, as $\Delta f \to 180^\circ$, the minimum velocity take-off vector $\dot{\boldsymbol{r}}_{m_1}$ will become perpendicular to the $\boldsymbol{r}_1$ position vector. Again this agrees with the $\Delta f = 180^\circ$ special case condition which states that the minimum velocity transfer orbit will only have a transverse velocity component.

12.4 Rotating the Orbit Plane

The previous orbit maneuvers all had in common that they strive to move a spacecraft from one point in space to another point. In this section we will investigate a different type of orbit maneuver. Here the goal is to change the orbit plan orientation, without necessarily affecting the orbit geometry itself.

Assuming the spacecraft is undergoing Keplerian (non-perturbed) motion, then its orbit plane is fixed in inertial space. The orientation of any plane is

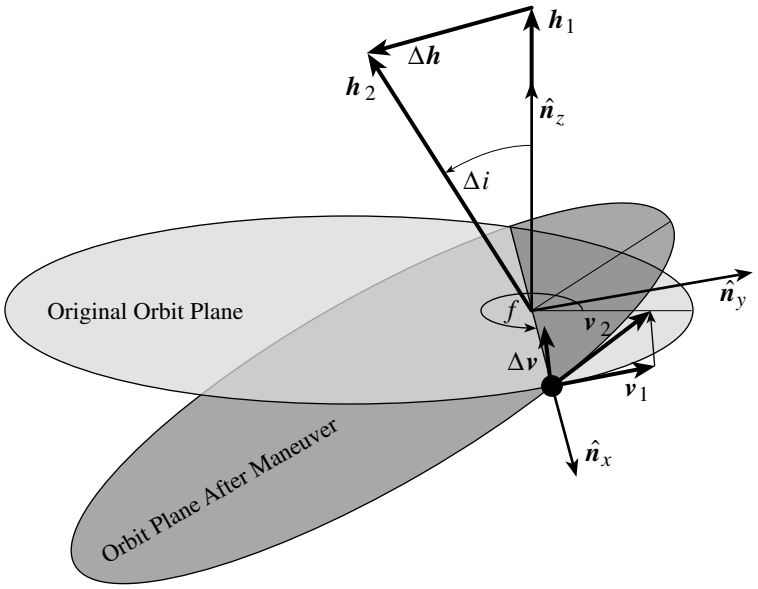


Figure 12.11: Illustration of an Orbit Plane Change Maneuver

completely prescribed through a unit normal vector to this plane. For the case of having a Keplerian motion, the angular momentum vector $\boldsymbol{h}_1 = \boldsymbol{r} \times \boldsymbol{v}_1$ provides such an orbit-plane normal vector. Let the desired orbit plane orientation be given through the angular momentum vector $\boldsymbol{h}_2 = \boldsymbol{r} \times \boldsymbol{v}_2$. These two orbit planes are illustrated in Figure 12.11. The difference in angular momentum vectors is given by

$$\Delta\boldsymbol{h} = \boldsymbol{h}_2 - \boldsymbol{h}_1 \tag{12.54}$$

Let the axis which which intersects the two orbit planes of interest be given be described through the unit direction vector $\hat{\boldsymbol{n}}_x$. Note that since both orbit planes are inertially fixed, so is $\hat{\boldsymbol{n}}_x$ inertially fixed. Let $\hat{\boldsymbol{n}}_z$ be the unit direction vector of the original angular momentum vector $\boldsymbol{h}_1$, then we can define the inertial frame $\mathcal{N} : \{\hat{\boldsymbol{n}}_x, \hat{\boldsymbol{n}}_y, \hat{\boldsymbol{n}}_z\}$, where $\hat{\boldsymbol{n}}_y$ completes the right-hand coordinate system. The orbit plane orientation change is described through the scalar angular parameter Δi. As shown in Eq. (2.38), the rate of change of the angular momentum vector $\boldsymbol{h}$ is equal to the torque $\boldsymbol{L}$ applied to the spacecraft.

$$\dot{\boldsymbol{h}} = \boldsymbol{L} = \boldsymbol{r} \times \dot{\boldsymbol{v}} \tag{12.55}$$

Assuming the orbit plane change is applied impulsively, we are able to rewrite Eq. (12.55) as a difference equation between angular momentum and velocity vectors.

$$\Delta\boldsymbol{h} = \boldsymbol{r} \times \Delta\boldsymbol{v} = \boldsymbol{h}_2 - \boldsymbol{h}_1 \tag{12.56}$$

To determine what velocity change $\Delta\boldsymbol{v}$ must be applied at which point $\boldsymbol{r}$ of the original orbit, we write these two vectors with inertial $\mathcal{N}$ components as

$$\boldsymbol{r} = r_x\hat{\boldsymbol{n}}_x + r_y\hat{\boldsymbol{n}}_y \tag{12.57}$$

$$\Delta\boldsymbol{v} = \Delta v_x\hat{\boldsymbol{n}}_x + \Delta v_y\hat{\boldsymbol{n}}_y + \Delta v_z\hat{\boldsymbol{n}}_z \tag{12.58}$$

Studying Figure 12.11, it is clear that the angular momentum vector change $\Delta\boldsymbol{h}$ will not have any vector component along the $\hat{\boldsymbol{n}}_x$ axis. Thus, $\Delta\boldsymbol{h}$ can be written in terms of inertial $\mathcal{N}$ frame components as

$$\Delta\boldsymbol{h} = -\Delta h_y\hat{\boldsymbol{n}}_y - \Delta h_z\hat{\boldsymbol{n}}_z \tag{12.59}$$

with the angular momentum change vector components Δh_y and Δh_z being given by

$$\Delta h_y = h_1 \sin\Delta i \tag{12.60}$$

$$\Delta h_z = h_1\left(1 - \cos\Delta i\right) \tag{12.61}$$

Substituting Eqs. (12.57) and (12.58) into Eq. (12.56), the angular momentum difference $\Delta\boldsymbol{h}$ is expressed as

$$\Delta\boldsymbol{h} = r_y\Delta v_z\hat{\boldsymbol{n}}_x - r_x\Delta v_z\hat{\boldsymbol{n}}_y + \left(r_x\Delta v_y - r_y\Delta v_x\right)\hat{\boldsymbol{n}}_z \tag{12.62}$$

By comparing the $\mathcal{N}$ frame components of Eqs. (12.59) and (12.62), we are able to make the following conclusions. Since the angular momentum change along the $\hat{\boldsymbol{n}}_y$ axis is non-zero, neither the position vector component r_y nor the velocity vector difference component Δv_z can be zero. Thus, since there is no angular momentum change along the $\hat{\boldsymbol{n}}_x$ axis, the position vector component r_y must be zero. This provides us with the intuitive solution that the orbit plane change will occur at the point $\boldsymbol{r} = r_x\hat{\boldsymbol{n}}_x$ where where the two orbit planes intersect. Note that r_x could be either positive or negative, depending on which of the possible two intersection points is used to perform the orbit plane change.

Let us write the orbit velocity vector $\boldsymbol{v}_1$ in terms of $\mathcal{N}$ frame components as

$$\boldsymbol{v}_1 = \dot{r}_x\hat{\boldsymbol{n}}_x + r_x\dot{f}\hat{\boldsymbol{n}}_y = v_x\hat{\boldsymbol{n}}_x + v_y\hat{\boldsymbol{n}}_y \tag{12.63}$$

Using the angular momentum property $h_1 = r_x^2\dot{f}$, and by comparing the vector components along the remaining two axis of the $\mathcal{N}$ frame, the velocity vector change vector components are in terms of the desired orbit plane difference angle Δi as

$$\Delta v_x = \text{arbitrary} \tag{12.64}$$

$$\Delta v_y = -v_y\left(1 - \cos\Delta i\right) \tag{12.65}$$

$$\Delta v_z = v_y \sin\Delta i \tag{12.66}$$

Since Δv_x is arbitrary, while Δv_y and Δv_z must have specific values, the minimum energy orbit plane change is achieved by setting Δv_x equal to zero. Let us

further examine what happens to the orbit energy v_i^2. Using the orbit velocity $\mathcal{N}$ frame vector components shown in Eq. (12.63), the orbit energy v_2^2 is found to be

$$\begin{aligned} v_2^2 &= (v_x + \Delta v_x)^2 + v_y^2 \left(1 - (1 - \cos\Delta i)\right)^2 + v_y^2 \sin^2 \Delta i \\ &= (v_x + \Delta v_x)^2 + v_y^2 = v_1^2 + 2 v_x \Delta v_x + \Delta v_x^2 \end{aligned} \tag{12.67}$$

Thus, for the minimum energy orbit plane change where $\Delta v_x = 0$, we find that $v_2^2 = v_1^2$. Therefore the orbit energy is not changed during such a plane change. Practically, this means that a minimum energy orbit plane change will only change the orbit plane itself, but will not affect the orbit geometry (semi-major axis and eccentricity).

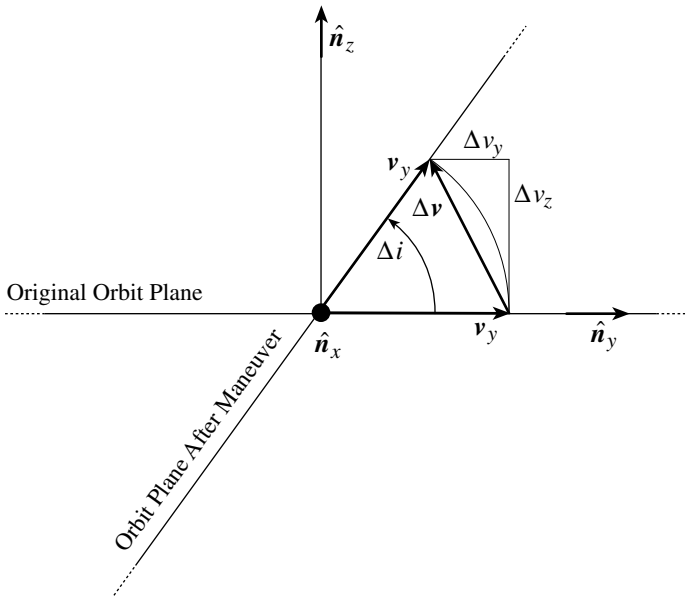


Figure 12.12: Side View of Velocity Vectors Involved in a Minimal Energy Orbit Plane Change Maneuver

Figure 12.12 illustrates what happens to the orbit velocity vectors during the orbit plane change maneuver. The tangential velocity component v_y is simply rotated by the $\Delta \boldsymbol{v}$ velocity correction to lie in the desired orbit plane. The radial velocity component v_x is unaffected by the orbit plane change if Δv_x is equal to zero. However, since the magnitudes of the $\Delta \boldsymbol{v}$ components are directly a function of the tangential orbit velocity v_y, performing a orbit plane change is a very costly maneuver that requires a relatively large $\Delta \boldsymbol{v}$.

Example 12.3: To illustrate how much Δv is required to perform a orbit plane change, we revisit initial circular orbit used in Example 12.2. In that

example, an initial circular orbit radius of 7000 km was increased to a circular orbit of 7200 km through a Hohmann orbit transfer. The total velocity change required for that maneuver was $\Delta v_H = 104.351$ m/s. Using the same fuel cost, let us see just how far we would be able to rotate the initial orbit plane. Summing the terms in Eqs. (12.65) and (12.66), the total fuel budget for the minimal energy orbit plane change is given by

$$\Delta v = v_y \sqrt{2(1 - \cos \Delta i)} \tag{12.68}$$

Solving this equation for the desired Δi we find

$$\Delta i = \cos^{-1} \left(1 - \frac{1}{2} \left(\frac{\Delta v}{v_y} \right)^2 \right)$$

Given an orbit radius of $r_x = 7000$ km, the tangential orbit velocity v_y is 7460 m/s. For a given allowable Δv_H, the achievable orbit plan rotation angle Δi is only 0.80°. This corresponds to a maximum out-of-plane separation from the original orbit of 97.92 m.

Example 12.4: Using Gauss' variation equations in Eqs. (11.153), we saw how a continuous external disturbance acceleration $\boldsymbol{a} = (a_r, a_\theta, a_h)$ will affect the orbit element rates. We could also use Eq. (11.153c) to find how a continuous thrust is used to change the orbit inclination angle i. This equation is repeated here for convenience.

$$\frac{di}{dt} = \frac{r \cos \theta}{h} a_h$$

where θ is the true latitude angle. The orbit inclination angle is easiest to adjust while crossing the equatorial plane with $\theta = 0$. Using $h = r v_y$, we find the Δv requirement for a desired Δi inclination change to be

$$\Delta v_h = v_y \Delta i$$

Note that this equation is obtained by integrating the continuous disturbance equation over a small time interval Δt. At first glance this equation appears to differ from Eq. (12.68). However, if we linearize Eq. (12.68), i.e. assume that the orbit inclination change Δi is small, it agrees with the Δv requirement developed in this example.

Note another important detail. Gauss' equations form a set of continuous differential equations for the orbit elements in terms of the disturbance acceleration vector components a_r, a_θ and a_h. Note that these vector components are taken relative to the rotating spacecraft fixed reference frame $\{\hat{\boldsymbol{\imath}}_r, \hat{\boldsymbol{\imath}}_\theta, \hat{\boldsymbol{\imath}}_h\}$. Since the impulsive thrust is approximated to be a continuous thrust over a small period of time Δt, the orientation vector $\hat{\boldsymbol{\imath}}_h$ is time varying through-out this maneuver. Thus, the Δv_h requirement shown above is to to be taken in a straight-line manner. Rather, it forms the arc length as the current velocity vector component v_y is rotated to the desired orbit inclination angle.

12.5 Patched-Conic Orbit Solution

In Section 10.5 we discussed the concept of gravitational spheres of influence. The idea is that the gravity field of a multi-body system can be locally approximated about a single body through the inverse square gravitational field. This concept is illustrated in Figure 12.13. Assume that $m_1 >> m_2$, then the local gravitational sphere of influence about m_2 is approximated in Eq. (10.80) as

$$r = \left(\frac{m_2}{m_1}\right)^{2/5} r_{12} \tag{12.69}$$

Note that this formula was derived seeking the surface where the gravitational accelerations due to either body are equal. Different formula exists to compute the approximated spherical region of influence. However, to be used in the method of patched-conic orbits, the resulting energy approximations are rather insensitive to the choice of sphere of influence radius formula selection.

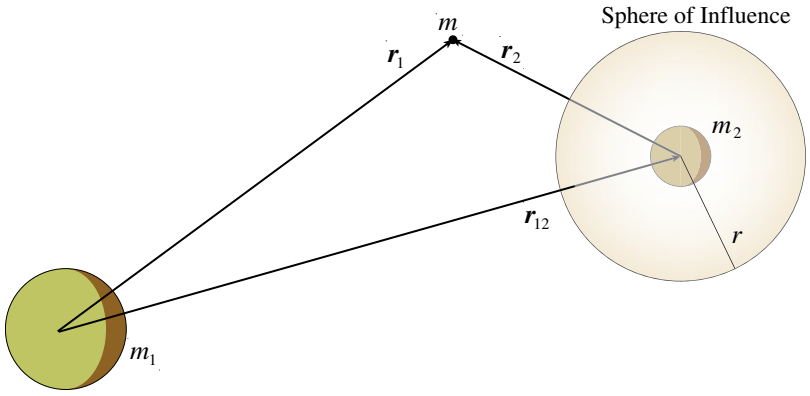


Figure 12.13: Illustration of the Concept of Gravitational Spheres of Influence

When a spacecraft is traveling among several celestial bodies, then the transfer orbit can be dissected into a finite number of section using the concept of gravitational spheres of influence. At any instance of time, the gravitational attraction acting on the spacecraft is assumed to be originating solely from the local dominant gravitational influence. For example, using the illustration in Figure 12.13, assume that m_1 is Earth and m_2 is the Moon. As the spacecraft enters a transfer orbit from the Earth to the Moon, its initial trajectory is essentially a solution of the classical Keplerian two-body problem. Only as the spacecraft becomes sufficiently close to the Moon is its gravitational attraction dominated by the Moon. To find an initial transfer orbit guess, or to approximately evaluate what Δv's would be involved in reaching the Moon, the transfer orbit can be split into two regions where the spacecraft is considered to be under the sole influence of either the Earth or the Moon. This method of dissecting a transfer orbit into various sections of two-body solutions is referred to a the

method of patched conics. To determine a precise transfer orbit, a numerical solution technique must be employed which incorporates the gravitational influence of all celestial bodies, as well as any other existing perturbations, at any instance of time. The most common use of the patched-conic orbit solution is to determine approximately what Δv's would be required for a proposed mission or to find an initial orbit guess that will start a numerical orbit search algorithm. We mention that the patched conic method is known to be much more valid in estimating the magnitude of Δv, rather than establishing the direction and timing of the velocity changes.

While in transit between two planets, an interplanetary spacecraft will spend most of its time under the dominant gravitational influence of the Sun . Investigating the feasibility of interplanetary missions, the concept of the patch-conic orbits is very useful. A sample orbit between Earth and Mars is illustrated in Figure 12.14. While the spacecraft is in the Earth's sphere on influence, it is shown to be on a hyperbolic orbit relative to Earth. As the spacecraft leaves the Earth's sphere of influence, it continues on under the gravitational influence of the Sun. Even though we had a hyperbolic orbit relative to Earth, at this point the spacecraft velocity (energy) relative to the Sun is only sufficient to yield an elliptic orbit with the Sun as its focus. Although both the Earth and the Sun attract the spacecraft, the Sun is ignored inside the Earth's sphere of influence, and the Earth is ignored outside it's sphere of influence. After the long heliocentric orbit transit phase, the spacecraft finally enters Mars' sphere of influence. What type of orbit the spacecraft will have relative to Mars depends on the relative velocity magnitude of the craft. However, since it's velocity at infinity is nonzero, we can anticipate a hyperbolic planet centric orbit from the onset. The illustration in Figure 12.14 depicts the expected hyperbolic orbit of the spacecraft relative to the target planet.

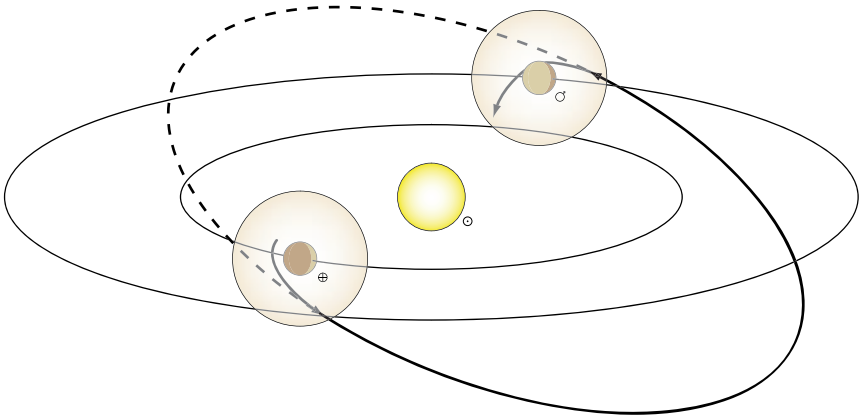


Figure 12.14: Approximating a Trajectory Among Multiple Celestial Bodies Through Gravitational Spheres of Influences

12.5.1 Establishing the Heliocentric Departure Velocity

Note that the transit orbit between the two planets can be chosen to fit the mission requirements. To find the minimum energy transfer orbit to boost a spacecraft from Earth's orbit to another planets orbit, a Hohmann transfer ellipse would be chosen. A sample Hohmann transfer orbit from Earth to Mars is illustrated in Figure 12.15. A major benefit of Hohmann transfer orbits is that the spacecraft approach trajectory will asymptotically approach the target planet trajectory at the rendez vous point. The approach speed will also be relatively slow. Without the target planet's gravitational field present, the craft would not have the proper heliocentric velocity to remain in this orbit. However, if guided properly, it is possible for the spacecraft to enter the local planet's gravity well and, with a modest energy charge, remain in a closed orbit about this planet. Depending on the target orbit, only small Δv orbit corrections would be necessary to achieve this final planetary orbit.

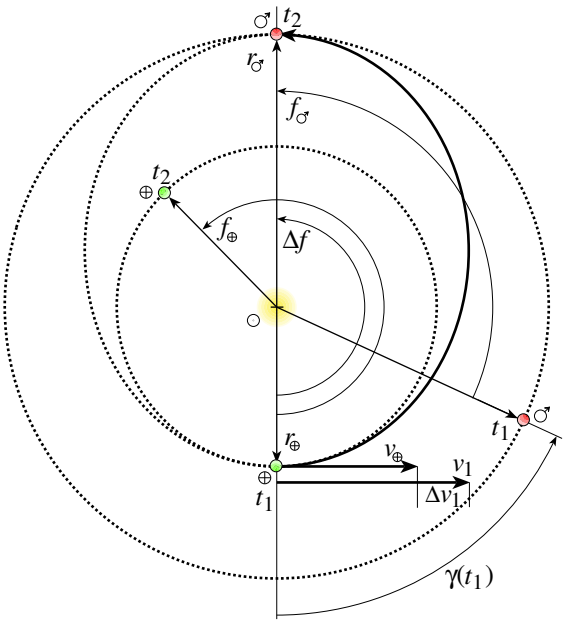


Figure 12.15: Hohmann Transfer Orbit Illustration between Earth and Mars

Example 12.2 illustrates how to compute the total Δv required to perform a Hohmann transfer between two circular orbits. For the interplanetary Hohmann transfer, we only make use of the first Δv_1 calculation in Eq. (12.25) that yields the required heliocentric departure velocity v_1. At the end of the Hohmann transfer, no second burn is performed to circularize the spacecraft orbit about the Sun. Instead, the spacecraft is guided and controlled in such a way that the craft is captured by the target planets gravity well. The total Δv requirement

to perform this capture depends on the target orbit geometry and orientation.

Using the average planetary orbit radii shown in Table 10.1, we can compute the minimum heliocentric (measured relative to the Sun) Δv requirement to depart Earth's orbit and arrive at any other planet in our solar system. The required Earth departure Δv's are computed using Eq. (12.25). The results are shown in Table 12.1. Note that a negative Δv means that the spacecraft must slow down relative to Earth in order to reach this planet. As a comparison, the Earth's heliocentric speed is 29.77 km/s. The negative Δv entries for Mercury and Venus indicate that the craft must exit the Earth's sphere of influence in the opposite direction to the Earth's counter-clockwise motion. Through this the apofocus counter-clockwise velocity will be appropriately reduced so that the craft "falls" interior to the Earth's orbit and arrives at either planet at perifocus.

Table 12.1: Minimum Δv Requirements to Reach Other Planets While Departing From Earth

Planet	Departure Δv [km/s]	Transfer Time [years]
Mercury ☿	-7.53	0.29
Venus ♀	-2.50	0.40
Mars ♂	2.94	0.71
Jupiter ♃	8.79	2.73
Saturn ♄	10.29	6.07
Uranus ♅	11.28	16.06
Neptune ♆	11.65	30.71
Pluto ♇	11.81	45.66

The results in Table 12.1 provide only rough estimates of the minimum energy Δv requirements and flight times. This simple calculation ignores any orbit plane changes that might be required, as well the Δv requirement to escape the departure planets local gravity field or the Δv requirement to park in an orbit about the arrival planet. Even so, it is clear that the fuel requirements to reach all the outer solar system planets with a minimum energy transfer orbit are about the same. However, the flight time to reach Uranus, Neptune and Pluto become very large and vary substantially from each other. The reason for the small change in fuel requirements to the outer planets is that the Sun's gravitational attraction is much smaller in the outer reaches of our solar system then in the proximity of Earth. In all cases to the outer planets the Earth relative departure velocity is approaching the escape speed of approximately 12.33 km/s. The interplanetary spacecraft will require most of the fuel to travel through the inner solar system. Traveling from Saturn on to the outer planets (or escaping the solar system) will only take a relatively minor addition in Δv requirements.

To have the spacecraft actually intercept the target planet when it reaches the desired heliocentric orbit radius, the departure and target planets true

anomaly angles must have a specific phase difference $\gamma(t_1)$ at the time of the spacecraft launch corresponding to the trip time. Let Δf be the angular distance that the spacecraft will travel relative to the Sun while in transit between planets. The associated travel time is given by ΔT. Let n_1 be the mean angular rate of the departure planet about the Sun, while n_2 is the mean angular rate of the target planet. Since the target planet will have traveled an angular distance of $n_2\Delta T$ while the spacecraft is in transit, the initial phase angle between the departure and target planets must be

$$\gamma(t_1) = \Delta f - n_2 \Delta T \tag{12.70}$$

If a launch cannot be performed while the planets are at this particular phase angle, then the mission planners must wait until the planets have rotated sufficiently for this launch condition to repeat itself (or choose to use a sub-optimal transfer orbit). How long this wait will be depends on the synodic period between the departure and target planet. A synodic period T_s is defined as the time required for a particular phase angle between the planets to repeat itself.[4] To compute the synodic period T_s, let us assume that the planets are at a desired phase angle γ at the initial time. For this angle to repeat itself, after a period T_s the angular difference between the planets must have changed by $\pm 2\pi$.

$$\gamma(T_s) = \gamma(t_0) + n_2 T_s - n_1 T_s = \gamma(t_0) \pm 2\pi \tag{12.71}$$

From this condition, the synodic period T_s is expressed in terms of the planets heliocentric rotations rates n_i as[4]

$$T_s = \frac{2\pi}{|n_2 - n_1|} \tag{12.72}$$

Assuming that the interplanetary spacecraft is departing from Earth, Table 12.2 shows the synodic periods for the various launch windows to repeat themselves. Note that if there is a large difference in the planetary rotation rates n_i, then the synodic periods will be relatively short. For example, the angular rate $n_{\mathbb{Q}}$ of the planet Mercury is about 4 times larger than that of Earth. By the time that Mercury finishes one revolution about the Sun, Earth will only have rotated a relatively small distance. To catch up to the required phase angle γ will only take a short time. This is why the synodic period between Mercury and Earth is only slightly larger than the revolution period of Mercury. On the other hand, the synodic periods between Earth and the outer planets is essentially one Earth year. In this case Earth is considered to be the fast planet, while the outer planets are almost standing still in comparison. It takes Neptune and Pluto over 100 years to finish one revolution about the Sun. However, as the difference in heliocentric rotation rates becomes small, then it can take a long time for the required planetary phase condition to repeat itself. This is why the synodic periods between Earth and Mars or Venus are the largest.

Besides requiring potentially long transit times and perhaps extended waits for favorable planetary positions, using a pure Hohmann transfer for interplanetary travel has other drawbacks. If the spacecraft is to travel by another planet

Table 12.2: Synodic Periods between Earth and Other Planets

Planet	Heliocentric Ang. Rate [deg/year]	Revolution Period about the Sun [years]	Synodic Period [years]
Mercury ☿	1493.04	0.24	0.318
Venus ♀	584.60	0.62	1.600
Mars ♂	191.20	1.88	2.138
Jupiter ♃	30.30	11.88	1.093
Saturn ♄	12.18	29.57	1.036
Uranus ♅	4.27	84.17	1.013
Neptune ♆	2.17	165.40	1.007
Pluto ♇	1.45	248.81	1.005

and then return to Earth, the craft will return to the point in space where Earth was when the spacecraft was launched. However, in the mean time Earth will have moved on to a new position about the Sun. For a free-return fly-by type mission, the Hohmann transfer orbit would need to be abandoned to guarantee that the spacecraft will reach both the other planet and Earth during its return flight.

Example 12.5: Let us investigate a simplified heliocentric Hohmann transfer orbit between Earth and Mars as shown in Figure 12.15. This example illustrates how the values in Table 12.1 were obtained. Both planetary orbits are assumed to lie in the same plane and have zero eccentricity. The simplified geometry is illustrated in Figure 12.15.

The gravitational constant of the Sun is $\mu_\odot = 1.326 \cdot 10^{11}$ km^3/s^2. Earth's average heliocentric radius is $r_\oplus = 149.60 \cdot 10^6$ km, while Mars' average radius is $r_\mars = 227.94 \cdot 10^6$ km. The planet's mean rotation rate about the Sun is then computed through

$$n_\oplus = \sqrt{\frac{\mu_\odot}{r_\oplus^3}} = 0.985 \text{ deg/day}$$

$$n_\mars = \sqrt{\frac{\mu_\odot}{r_\mars^3}} = 0.985 \text{ deg/day}$$

The Earth's heliocentric velocity magnitude $v_\oplus$ is

$$v_\oplus = \sqrt{\frac{\mu_\odot}{r_\oplus}} = 29.77 \text{ km/s}$$

The semi-major axis of the minimum energy transfer orbit is computed using Eq. (12.9).

$$a = \frac{r_\oplus + r_\mars}{2}$$

Since the Hohmann transfer ellipse is only traveled from the periapses to the apoapses (half an orbit), the transfer time ΔT is

$$\Delta T = \frac{1}{2}\left(2\pi\sqrt{\frac{a^3}{\mu_\odot}}\right) = 258.98 \text{ days}$$

The angular distances that the planets travel while the spacecraft is completing its transfer orbit are

$$f_{\oplus} = n_{\oplus} \cdot \Delta T = 255.13 \text{ deg}$$
$$f_{\sigma} = n_{\sigma} \cdot \Delta T = 135.66 \text{ deg}$$

Studying these angles, it is evident how critical timing is when performing any interplanetary missions. When the spacecraft is departing the Earth's sphere of influence, Mars must be 135.66 degrees away from the rendez vous point. This is generally not the case. Mission planners must therefore look at the Earth and Mars motion and plan their launch windows accordingly. If the different orbit inclinations and eccentricities are also taken into account, then the launch window calculation become more complex.

To be able to intercept a planet at a specific time, Lambert's two-point boundary value problem would need to be solved. This is illustrated in Figure 12.16 as a non-Hohmann transfer orbit between Earth and Mars. This method allows us to specify the spacecraft to be at a desired location at a desired time. However, care must be taken when choosing the intercept time and place to avoid large Δv requirements. To find an optimum solution taking both fuel consumption and time of flight considerations into account, a numerical optimization is typically performed taking all gravitational attractions and other perturbations into account.

Another concern of such faster transfer orbits is that the spacecraft's approach trajectory tangent will not asymptotically approach the arrival planets trajectory target. Instead, the craft may approach the target planets trajectory at an oblique angle. The relative approach speeds are often very high and have different directions. So, while the transfer orbit time has been reduced, it may take longer to insert the craft in a desired orbit about the target planet. Due to deceleration and/or propulsion constraints, this is may be performed in multiple steps.

12.5.2 Escaping the Departure Planet's Sphere of Influence

Previously we discussed methods to find the required heliocentric velocity v_1 that a spacecraft will need to travel from one planet and travel to another. In the following discussion we will investigate how a spacecraft will escape the gravitational influence of the departure planet. As a general notation heliocentric velocity magnitudes are expressed as v_i, while velocity magnitudes relative to either the departure or target planets are expressed as ν_i.

Let us examine how a spacecraft would travel through the gravitational sphere of influence of the departure planet. As the spacecraft exits the planets sphere of influence, its velocity must have a know magnitude v_1 and direction. Without loss of generality, let us assume that the interplanetary transfer is a minimum energy Hohmann transfer ellipse and that the departure planet is Earth. According to Eq. (12.17), the heliocentric departure velocity vector $\boldsymbol{v}_1$

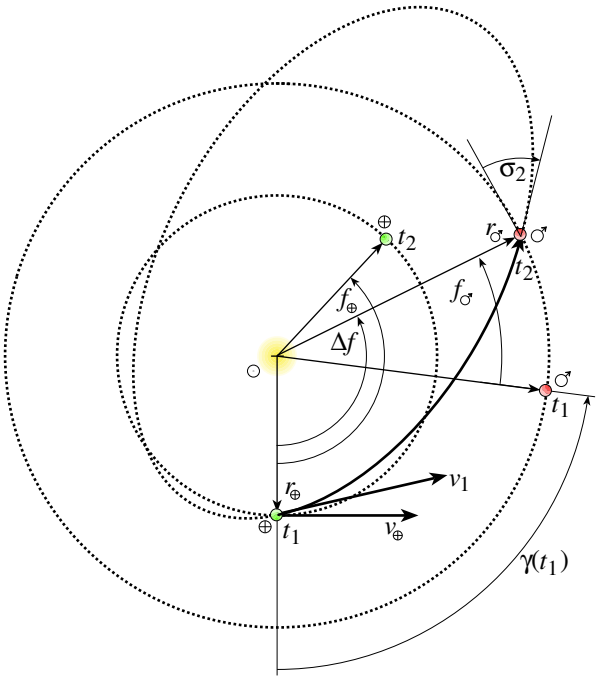


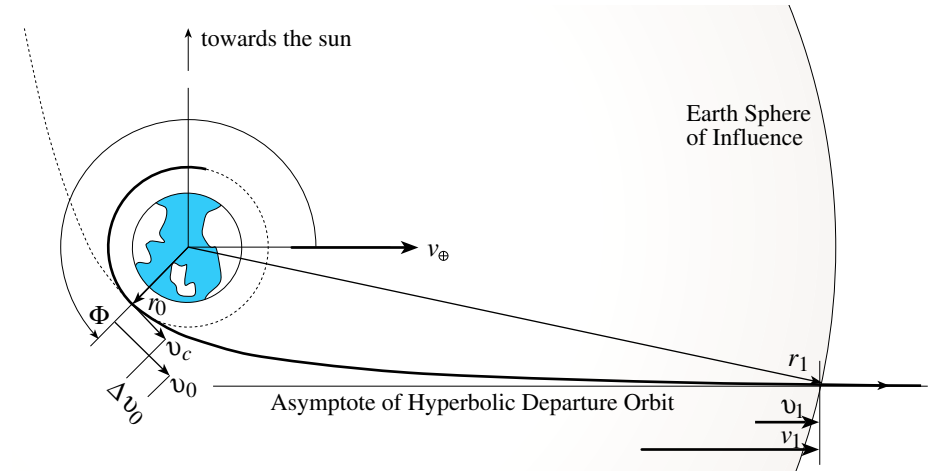
Figure 12.16: Non-Minimum Energy Transfer Orbit Illustration between Earth and Mars

must be aligned with the Earth's heliocentric velocity vector. Further, let us assume that the spacecraft is initially in a circular orbit about the Earth at a radius r_0.

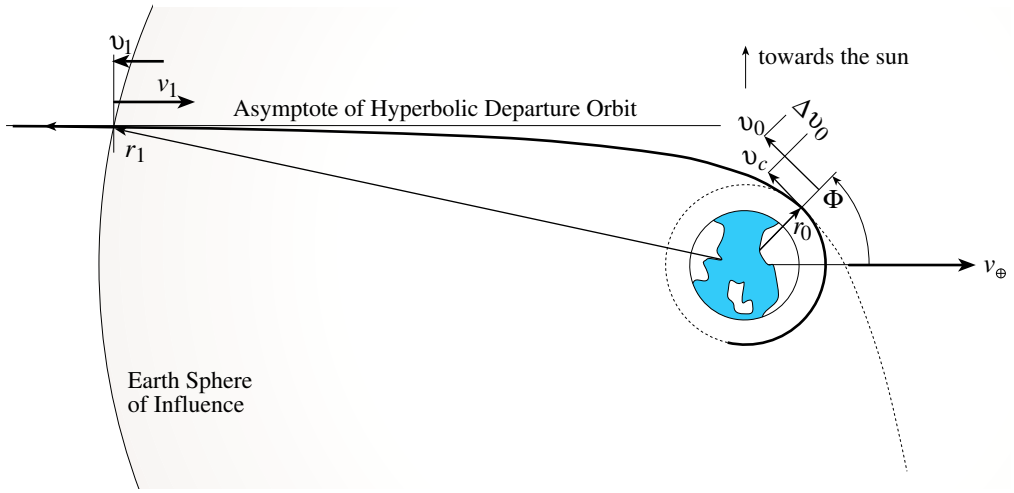
To depart Earth's sphere of influence, either a parabolic or hyperbolic orbit is required relative to Earth. As the Earth relative position vector grows large, we required the velocity vector to be aligned with the heliocentric Earth velocity $\boldsymbol{\nu}_\oplus$. Figure 12.17 illustrates the departure hyperbola from Earth both an outer or inner solar system planet. If the spacecraft is to travel to an outer planet, Table 12.1 shows that a positive Δv is required. Thus the spacecraft needs to accelerate relative to the heliocentric Earth velocity. If the spacecraft is to travel to an inner planet, then a negative Δv is to be applied and the craft has to slow down relative to Earth.

As the departing spacecraft approaches Earth's sphere of influence, its velocity vector must have converged sufficiently to the required magnitude v_1 and direction. If the parking orbit radius r_0 is large compared to the sphere of influence radius, then this may not occur.

Let the time t_0 be the time where the spacecraft performs a burn to leave its circular orbit about Earth and enter a hyperbolic departure orbit. The time t_1 is defined as the an instance where the departure orbit intersects the planets sphere of influence. The required Earth relative velocity nu_1 that the spacecraft



(i) Departure Orbit to Reach Outer Planet



(ii) Departure Orbit to Reach Inner Planet

Figure 12.17: Earth Relative Hyperbolic Departure Orbit Illustration

must possess as it approaches the sphere of influence is computed using

$$\nu_1 = v_1 - v_\oplus \tag{12.73}$$

The vis-viva equation of a hyperbolic orbit about Earth is given by

$$\nu_i^2 = \frac{2\mu_\oplus}{r_i} - \frac{\mu_\oplus}{a} \tag{12.74}$$

where the semi-major axis of a hyperbola is defined to be a negative quantity. Since the spacecraft trajectory will have asymptotically approached its hyperbolic asymptote at t_1, we can approximate $r_1 \approx \infty$. Using Eq. (12.74), the Earth relative spacecraft velocity ν_1 is then given by

$$\nu_1 = \sqrt{\frac{2\mu_\oplus}{r_1} - \frac{\mu_\oplus}{a}} \approx \sqrt{-\frac{\mu_\oplus}{a}} \tag{12.75}$$

The semi-major axis a of the departure hyperbola is then expressed in terms of the departure velocity ν_1 or v_1 through

$$a = -\frac{\mu_\oplus}{\nu_1^2} = -\frac{\mu_\oplus}{(v_1 - v_\oplus)^2} \tag{12.76}$$

The Earth relative speed ν_0 that the spacecraft must have after the burn to enter the hyperbolic orbit at t_0 is

$$\nu_0 = \sqrt{\frac{2\mu_\oplus}{r_0} - \frac{\mu_\oplus}{a}} \tag{12.77}$$

After substituting Eq. (12.76), the speed ν_0 is expressed as

$$\nu_0^2 = \nu_1^2 + \frac{2\mu_\oplus}{r_0} \tag{12.78}$$

While in the circular Earth parking orbit, the spacecraft has an orbit speed of

$$\nu_c = \sqrt{\frac{\mu_\oplus}{r_0}} \tag{12.79}$$

The burn $\Delta\nu_0$ at t_0 to enter the departure hyperbolic orbit is computed using

$$\Delta\nu_0 = \nu_0 - \nu_c = \sqrt{2\nu_c^2 + \nu_1^2} - \nu_c \tag{12.80}$$

As the parking orbit radius r_0 becomes smaller, then the corresponding circular orbit velocity ν_c becomes larger. If r_0 is sufficiently small such that $\nu_c >> \nu_1$, then the departure burn $\Delta\nu_0$ can be approximated as

$$\Delta\nu_0 \approx \nu_c \tag{12.81}$$

The point at which the $\Delta\nu_0$ burn must be applied is defined through the angle Φ. Since the spacecraft velocity must asymptotically align itself with the

Earth heliocentric velocity vector , this burn angle Φ is the hyperbolic asymptote slope angle ϕ computed in Eq. (8.24). Note that the magnitude of this angle is computed differently for burns sending spacecraft to outer planets versus burns sending craft to inner planets. For transfers to inner planets, the burn angle Φ is defined as

$$\Phi = \cos^{-1}\left(\frac{1}{e}\right) \tag{12.82}$$

For transfers to outer planets, the burn angle Φ must have a phase angle π added to it.

$$\Phi = \cos^{-1}\left(\frac{1}{e}\right) + \pi \tag{12.83}$$

Eq. (12.82) corresponds to launching "in the evening" so that we exit out the "back door" of the Earth's sphere of influence. Whereas Eq. (12.83) corresponds to launching "in the morning" so that we exit out of the "front door" of the Earth's sphere of influence (see Figure 12.17).

Note that the hyperbolic eccentricity e has a value greater than 1. Further, note that the spacecraft hyperbolic injection point does not have to be in the Earth orbit plane about the Sun as illustrated in Figure 12.17. Let the angle Φ describe a cone about the Earth heliocentric velocity vector as illustrated in Figure 12.18. The craft can be in a general Earth orbit initially, as long as its trajectory intersects this cone. With each hyperbolic departure orbit, the spacecraft is initially in a circular parking orbit before receiving a tangential burn as shown in Figure 12.17. As prescribed, the hyperbolic departure orbit achieves the required escape velocity ν_1 at the Earth's sphere of influence with its direction being along either the positive or negative Earth velocity direction. The distance between the spacecraft and the Earth heliocentric velocity direction is negligible here. With the patched-conic interplanetary orbit solution, this distance is minor compared to the large distance to the other planet. However, to find a precise interplanetary transfer orbit using a numerical solution technique, this distance must be taken into consideration.

Given the heliocentric departure velocity v_1 required of the spacecraft to perform the interplanetary mission, Eq. (12.76) defines the semi-major axis of the hyperbolic departure orbit, while Eq. (12.78) defines the initial Earth relative velocity ν_0 at the hyperbolic perigee. Lastly, to define the hyperbolic trajectory geometry, as well as compute the burn angle Φ, we need to determine the departure orbit eccentricity. Using the definitions of angular momentum, as well as Eqs. (8.28) and (8.29), we are able to express h through

$$h^2 = \mu_\oplus p = \mu_\oplus a(1-e^2) = \mu_\oplus r_p(1+e) \tag{12.84}$$

Since the injection burn point at t_0 is the periapses of the departure orbit, note that $r_0 = r_p$. The angular momentum can then be expressed as

$$h^2 = r_0^2 \nu_0^2 \tag{12.85}$$

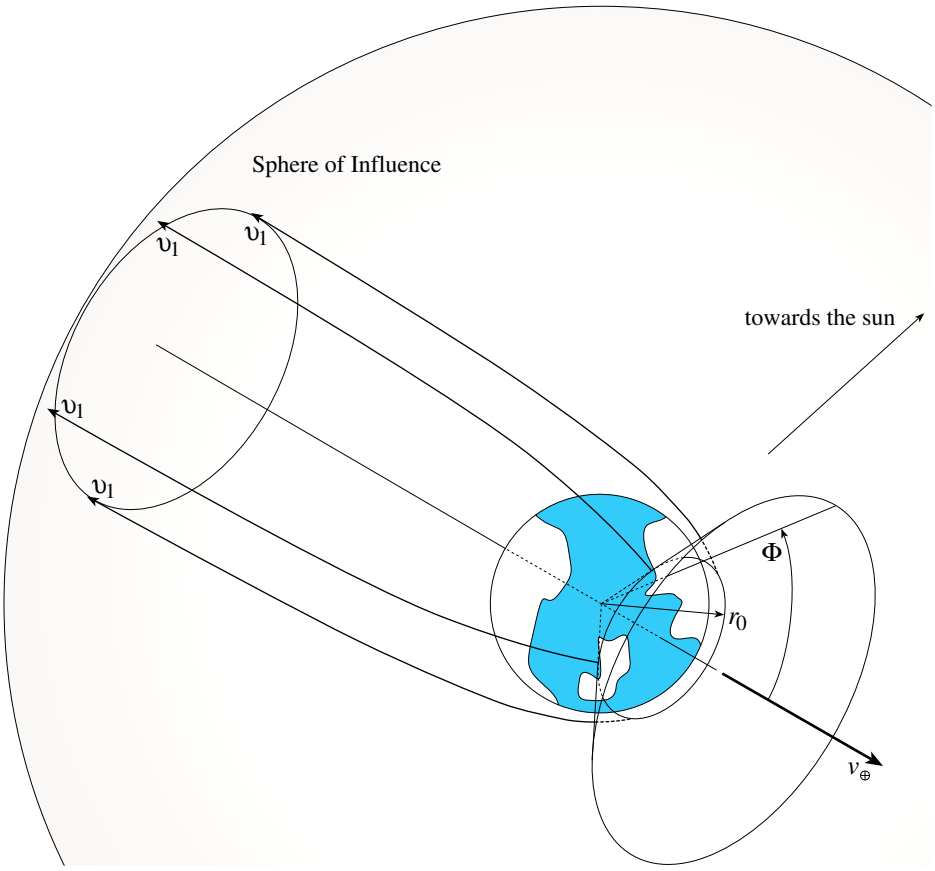


Figure 12.18: Three-Dimensional Illustration of the Departure Hyperbolic Orbits

Substituting Eq. (12.85) into Eq. (12.84), we are able to solve for the eccentricity e in terms of the initial hyperbolic orbit speed ν_0.

$$e = \frac{r_0 \nu_0^2}{\mu_\oplus} - 1 \tag{12.86}$$

Substituting Eq. (12.76) into Eq. (12.84), we can solve for e in terms of the escape velocity ν_1 that the spacecraft must have as it travels through the Earth's sphere of influence.

$$e = \frac{r_0 \nu_1^2}{\mu_\oplus} + 1 \tag{12.87}$$

Either formula can be used to compute the hyperbolic departure orbit eccentricity.

12.5.3 Enter the Target Planet's Sphere of Influence

After a long interplanetary transit phase of the mission, assume the spacecraft is entering the target planet's sphere of influence. Figure 12.19 illustrates the arrival of the spacecraft as seen relative to the target planet. Assume the interplanetary transfer orbit is a near minimum energy type orbit. If the spacecraft is traveling from Earth to an inner planet such as Venus, then the craft will arrive at the target planet at the perifocus of its heliocentric transfer orbit. Thus the heliocentric velocity of the spacecraft will be larger than the heliocentric orbit velocity of the planet. To let the craft be captured by the planet, we typically approach the planet through the "back door". We we were ahead of the planet, our large velocity would make us out run the planet and we would never enter its sphere of influence (see Figure 12.19(i)). If the spacecraft is traveling to an outer planet such as Mars, then the planet is reached at the apofocus of the heliocentric transfer orbit. Thus our velocity will be less than the planet's orbit velocity and we must position ourselves ahead of it. This way the planet will overtake us and allow us to enter its sphere of influence (see Figure 12.19(ii)).

Without loss of generality, let us assume that the spacecraft travels from Earth to Venus. Further, the planets orbits are once again assumed to be circular. To find the heliocentric arrival velocity v_2, need to know the departure planet heliocentric orbit radius $r_\oplus$ and the heliocentric departure velocity v_1. Using the vis-viva equation in Eq. (8.82) and given the target planet's heliocentric orbit radius r_φ, the arrival velocity v_2 is expressed as:

$$v_2 = \sqrt{\mu_\odot \left(\frac{1}{r_\varphi} - \frac{1}{r_\oplus} \right) + v_1^2} \tag{12.88}$$

To compute the spacecraft heading angle σ_2 relative to the Sun normal direction, we recall the definition of the angular momentum vector $\boldsymbol{h}$.

$$\boldsymbol{h} = \boldsymbol{r} \times \boldsymbol{v} \tag{12.89}$$

Assuming that the interplanetary mission began with a minimum energy burn along the Earth velocity vector, then $h = r_\oplus v_1$. Here planets sphere of influence radius is considered to be much smaller than the planetary heliocentric orbit radius. The angular momentum of the spacecraft as it reaches Venus' sphere of influence is then expressed as

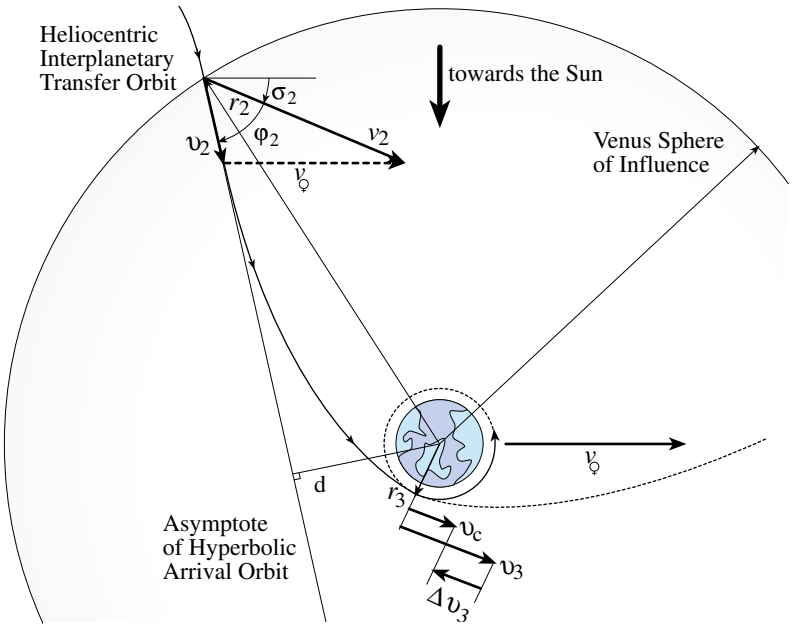
$$h = |\boldsymbol{r}_\varphi \times \boldsymbol{v}_2| = r_\varphi \cdot v_2 \cdot \sin(90^\text{o} - \sigma_2) = r_\varphi v_2 \cos\sigma_2 \tag{12.90}$$

Using Eq. (12.90), the heading angle σ_2 is written as

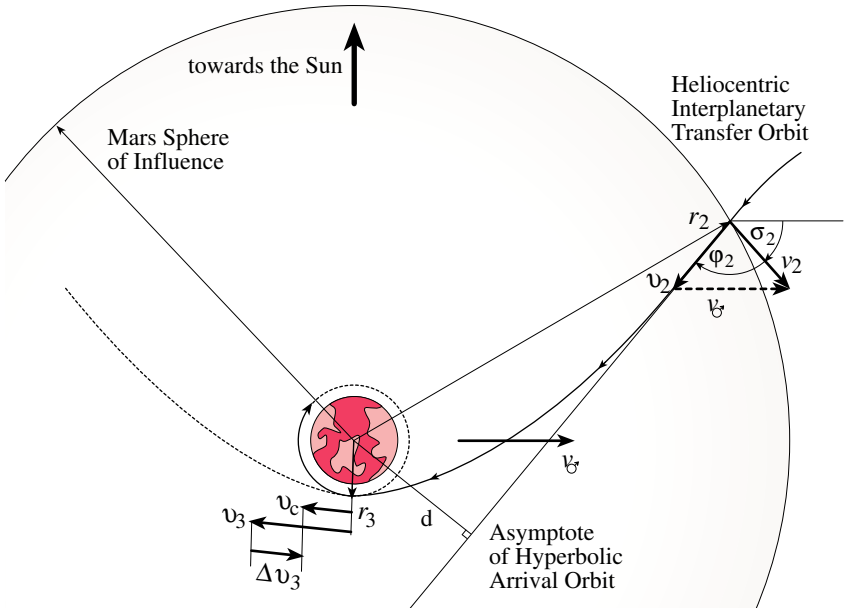
$$\sigma_2 = \cos^{-1}\left(\frac{h}{r_\varphi v_2} \right) \tag{12.91}$$

To compute the Venus relative velocity vector $\boldsymbol{\nu}_2$ of the spacecraft as it enters the planet's sphere of influence, the Venus heliocentric velocity $\boldsymbol{v}_\varphi$ must be subtracted from the heliocentric velocity $\boldsymbol{v}_2$ of the craft.

$$\boldsymbol{\nu}_2 = \boldsymbol{v}_2 - \boldsymbol{v}_\varphi \tag{12.92}$$



(i) Arrival Orbit to the Inner Planet Venus



(ii) Arrival Orbit to the Outer Planet Mars

Figure 12.19: Hyperbolic Arrival Orbit Illustration as seen by the Target Planet

Using the law of cosines, we can compute the magnitude of $\boldsymbol{n}u_2$ through

$$\nu_2 = \sqrt{v_2^2 + v_{\varphi}^2 - v_2 v_{\varphi} \cos\sigma_2} \tag{12.93}$$

The heading angle φ_2 between the velocity vectors $\boldsymbol{v}_2$ and $\boldsymbol{\nu}_2$ is found using the law of sines.

$$\varphi_2 = \sin^{-1}\left(\frac{v_{\varphi}}{\nu_2}\sin\sigma_2\right) \tag{12.94}$$

The velocity ν_2 is the velocity of the spacecraft relative to the target planet as it enters the sphere of influence. Using the energy equation, we can express the semi-major axis of the approach trajectory through

$$\frac{1}{a} = \frac{2}{r_2} - \frac{\nu_2^2}{\mu_{\varphi}} \tag{12.95}$$

Assuming that the approach trajectory is a hyperbolic orbit, which is typically the case, we can set $r_2 \to \infty$ and approximate a as

$$a = -\frac{\mu_{\varphi}}{\nu_2^2} \tag{12.96}$$

In order to achieve a final orbit about the target planet, it is obviously important that the spacecraft is not aimed directly at the target planet. Instead, it's heliocentric trajectory is designed such that it will miss the target planet by a certain miss-distance d_m. This distance is measured along the planets heliocentric orbit path as shown in Figure 12.20. The illustrations used in this discussion all assume the spacecraft is going to approach the target planet from behind (as seen by the planets heliocentric velocity direction). It is also possible for the spacecraft to be aimed to intercept the target planets trajectory ahead of the planet. However, the resulting orbit about the target planet will have the opposite direction. To compute the shortest distance d_a between the hyperbolic approach asymptote and the target planet we use

$$d_a = d_m \sin(\varphi_2 + \sigma_2) \tag{12.97}$$

The hyperbolic asymptote angle Φ is determined through

$$\Phi = \cos^{-1}\left(\frac{1}{e}\right) \tag{12.98}$$

To determine the hyperbolic eccentricity, we investigate again the constant spacecraft's angular momentum relative to the target planet. As the spacecraft enters the sphere of influence, the momentum h is given by

$$h = |\boldsymbol{r}_2 \times \boldsymbol{\nu}_2| = d_a \nu_2 \tag{12.99}$$

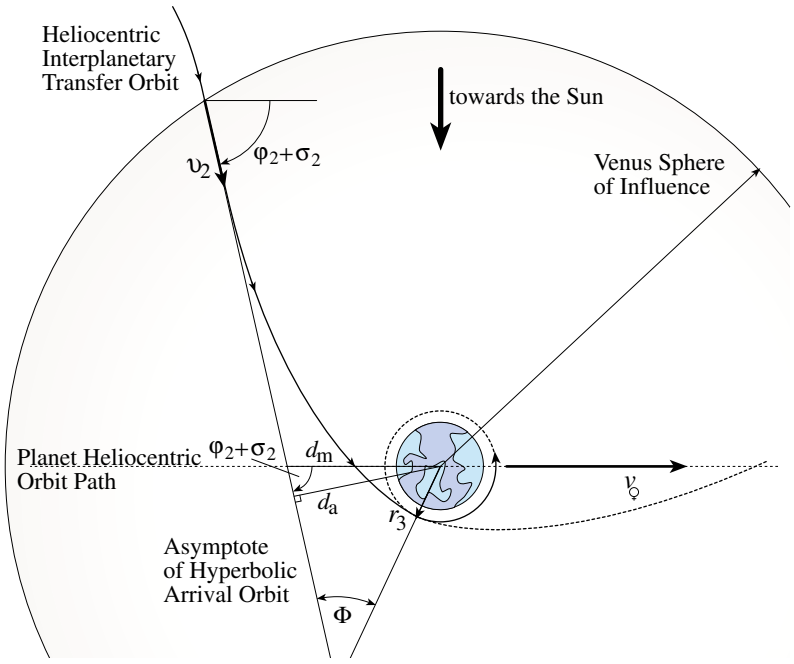


Figure 12.20: Illustration of the Asymptotic Approach Distance and Planet Miss-Distance

Substituting Eqs. (12.96) and (12.99) into Eq. (12.84), the hyperbolic eccentricity e is expressed as

$$e = \sqrt{1 + \frac{d_a^2 \nu_2^2}{\mu_{\varphi}^2}} \tag{12.100}$$

The periapses radius r_p is determined by substituting the semi-major axis a definition in Eq. (12.96) into the angular momentum expression in Eq. (12.84).

$$r_p = \frac{\mu_{\varphi}}{\nu_2^2}(e-1) \tag{12.101}$$

The orbit mission is typically designed such that the periapses radius r_p is also the desired circular orbit radius about the target planet. This state is controlled both by the approach speed ν_2 and the eccentricity e of the hyperbolic approach orbit. Since e depends on the miss-distance d_m, the periapses radius can be set by aiming the spacecraft an appropriate distance ahead or behind the planet.

Assume we wish to have a final circular parking orbit of radius r_3 about the target planet, where r_3 is also the periapses radius r_p of the hyperbolic approach orbit. Let time t_3 be the point where the approach trajectory touches this desired parking orbit. Using the energy equation, we can express the spacecraft

velocity v_3 at that instance.

$$\nu_3 = \sqrt{2\frac{\mu_\varphi}{r_3} + \nu_2^2} \tag{12.102}$$

The impulsive orbit correction required at perigee to circularize the orbit about the target planet is

$$\Delta\nu_3 = \nu_c - \nu_3 \tag{12.103}$$

where the circular orbit speed ν_c is given by

$$\nu_c = \sqrt{\frac{\mu_\varphi}{r_3}} \tag{12.104}$$

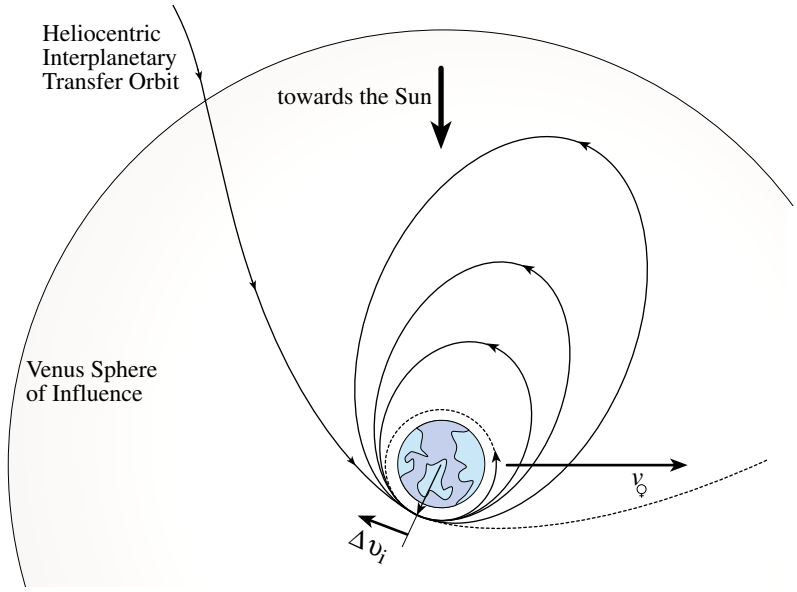


Figure 12.21: Illustration of Using a Staged Injection Burn with Elliptic Intermediate Orbits

Depending on the incoming hyperbolic velocity magnitude, the $\Delta\nu_3$ computed in Eq. (12.103) may cause excessive deceleration forces on the crew and spacecraft structure. Also, it is possible that the spacecraft engines cannot produce a large enough Δv over the short burning time. What is commonly done is to achieve the desired orbit about the target planet through several stages of $\Delta\nu$ burns. During the first periapses passage, a large enough $\Delta\nu$ is applied to change the target planet relative orbit from being hyperbolic to being elliptic. As the spacecraft revisits the periapses, additional burns are performed to gradually reduce the semi-major axis of the elliptic orbit to the desired value.

This concept is illustrated in Figure 12.21. Through this sequential adjustment approach, the deceleration forces can be kept smaller to be achievable by the trusters and, possibly, to be bearable to crew and spacecraft structure. Note that in order to use the patched-conic orbits to compute $\Delta\nu$ requirements, it is important that the staging ellipses remain well within the sphere of influence of the target planet. Otherwise we can no longer assume Keplerian two-body motion. Instead, the complete influence of both the target planet and the Sun must be taking into account in a numerical simulation.

Instead of performing thrusting maneuvers to decelerate the spacecraft, it is also possible to use the target atmosphere (if present) to decelerate the spacecraft. Such maneuvers are referred to as aerobraking maneuvers. Their effectiveness depends on the approach speed and mass of the spacecraft. If the either is too high, than the atmosphere may not be able to slow down the spacecraft sufficiently to achieve an elliptic orbit and the craft will exit the planets sphere of influence. Structural heating and stress restrictions dictate how low a spacecraft can dip into an atmosphere to slow down.

12.5.4 Planetary Fly-By's

All interplanetary orbits discussed so far had the spacecraft travel directly from the departure planet to the target planet. Due to the sphere's of influence concept, the gravitational attraction of the other solar system planets was ignored. This greatly simplified the initial mission analysis and allowed to predict some basic Δv estimates.

However, omitting the gravitational influence of other planets excludes a very attractive type of interplanetary transfer orbit. It is possible to make use of other planets' gravitational attraction to accelerate or decelerate the spacecraft relative to the Sun. Since several of these planets are quite massive (e.g. Jupiter), the favorable orbit perturbations can be large indeed. As the spacecraft approaches a planet, it will be slung around the planet and leave with a different heliocentric velocity direction and magnitude. This type of orbit maneuver is called a planetary fly-by and is illustrated in Figure 12.22. Instead of having a clean point-to-point interplanetary mission, the spacecraft is sent on a game of cosmic pinball. The major benefit of a planetary fly-by is that the spacecraft can be accelerated to reach a target planet faster without requiring more Δv's. The velocity direction changes are especially important, but the magnitude of the velocity can also be favorably affected. For example, to reach Mars it is possible to use Venus to sling-shot around and reach Mars in a much shorter period of time than what is typically possible for a given Δv and a direct interplanetary approach. Also, for a deep space probe to reach the outer solar system planets, it is common to have them fly by Jupiter first to receive an extra boost to their heliocentric velocity.

Assuming that the spacecraft is approaching a planet with a hyperbolic speed, it should be clear that having the craft sling around the planet should naturally change the velocity vector direction without having to fire any boosters. What may not be clear at first glance is why would this also change the

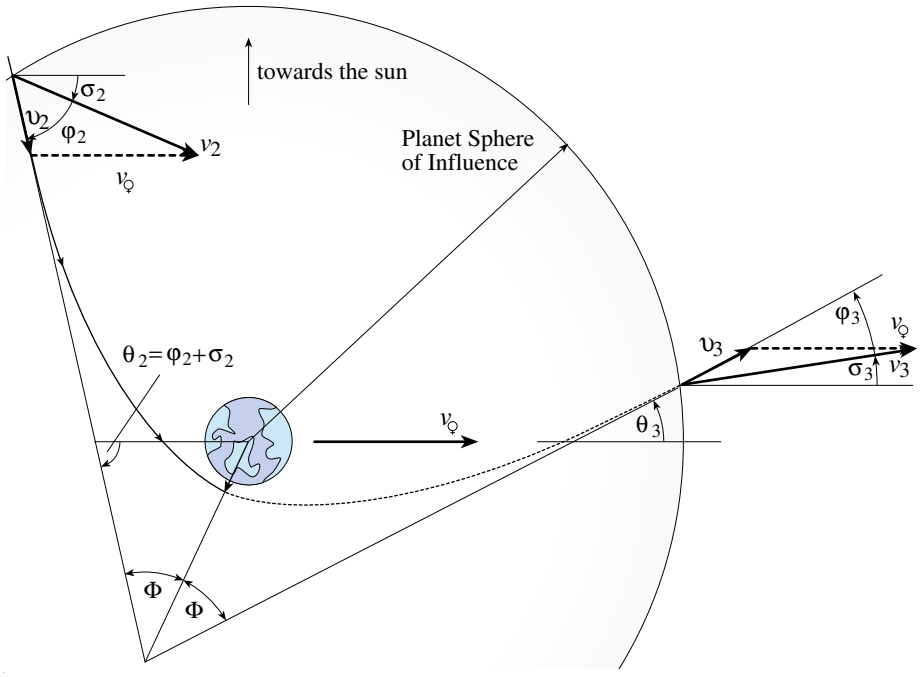


Figure 12.22: Illustration of a Sample Planetary Fly-By Maneuver about Venus

magnitude of the velocity vector. Figure 12.22 illustrates how a planetary fly-by would be seen by the moving planet. Let us use the same notation as was used developing the hyperbolic approach trajectories to a target planet. As the spacecraft as it enters the planets sphere of influence, its heliocentric velocity is given by $\boldsymbol{v}_2$. The planet relative velocity vector is $\boldsymbol{\nu}_2 = \boldsymbol{v}_2 - \boldsymbol{v}_{\varphi}$. Studying Keplerian two-body motion, we know that the velocity ν_2 with which the spacecraft approach the planet will also be the asymptotic velocity that the craft will approach as it leaves the planet on a hyperbolic orbit. The question is, how is the spacecraft being accelerate or decelerate here? The key observation to make here is that for interplanetary missions, we are concerned with the heliocentric velocity of a spacecraft $\boldsymbol{v}_i$, not with the velocity ν_i relative to some planet. As the craft departs the planets sphere of influence, it's *planet relative velocity* $\boldsymbol{\nu}_3$ will have the same magnitude as the approach velocity vector $\boldsymbol{\nu}_2$. The two velocity vectors $\boldsymbol{\nu}_i$ will only differ in their direction. Thus we have

$$\boldsymbol{\nu}_3 \neq \boldsymbol{\nu}_2 \qquad \nu_3 = \nu_2 \tag{12.105}$$

The spacecraft's heliocentric velocity as it departs the planet's sphere of influence is given by

$$\boldsymbol{v}_3 = \boldsymbol{\nu}_3 + \boldsymbol{v}_{\varphi} \tag{12.106}$$

Depending on the angle between the planetary heliocentric velocity vector and the flight angle of the spacecraft, the velocity $\boldsymbol{v}_3$ can either have a larger or smaller magnitude than $\boldsymbol{v}_2$. The illustration in Figure 12.22 shows a case where the planet is approached from behind (as seen relative to the planet's heliocentric velocity vector). This type of approach will result in the spacecraft picking up some heliocentric speed. If the craft approaches the planet from the front and departs rearward, then the craft would loose some heliocentric speed.

How to compute the incoming spacecraft heading angles σ_2 and φ_2, as well as the hyperbolic asymptote angle Φ, have already been shown in section 12.5.3. For notational convenience, let us define θ_i to be

$$\theta_i = \varphi_i + \sigma_i \tag{12.107}$$

Studying Figure 12.22, the angle θ_3 describes the planet relative departure velocity vector $\boldsymbol{\nu}_3$ direction relative to the planet's heliocentric velocity vector and is found to be

$$\theta_3 = 180^{\text{o}} - 2\Phi - \theta_2 \tag{12.108}$$

Using the law of cosines, the heliocentric velocity magnitude v_3 is expressed as

$$v_3 = \sqrt{\nu_3^2 + v_{\text{♀}}^2 + 2\nu_2 v_{\text{♀}} \cos\theta_3} \tag{12.109}$$

Using the law of sines, the angle φ_3 between the planet relative and the Sun relative velocity vectors is given by

$$\varphi_3 = \sin^{-1}\left(\sin\theta_3 \frac{v_{\text{♀}}}{v_3}\right) \tag{12.110}$$

Thus, the direction angle σ_3 of the spacecraft's heliocentric departure velocity vector is computed using

$$\sigma_3 = \theta_3 - \varphi_3 \tag{12.111}$$

The heliocentric angular momentum of the spacecraft as it enters the planets sphere of influence is

$$h(t_2) = r_{\text{♀}} v_2 \cos\sigma_2 \tag{12.112}$$

As the spacecraft exits the sphere of influence, it's angular momentum about the Sun is given by

$$h(t_3) = r_{\text{♀}} v_3 \cos\sigma_3 \tag{12.113}$$

Note that the planet's sphere of influence radius is assumed to be much smaller than the planet's heliocentric orbit radius. As the spacecraft accelerates or decelerates during the fly-by maneuver, it does so by exchanging angular momentum with the planet. Thus we find that momentum change Δh to be

$$\Delta h = h(t_3) - h(t_2) \neq 0 \tag{12.114}$$

Since the total angular momentum of the solar system is constant, the planets change in angular momentum will be $-\Delta h$. While the planet will change its heliocentric orbit velocity during a fly-by, this effect can be ignored. The reason for this is the huge mass imbalance between the spacecraft and the planets. The change in the planetary orbit velocity is minuscule and has no practical effect on its trajectory. Let us define m to be the spacecraft mass and v to be the spacecraft velocity component normal to the heliocentric orbit radius. The total angular momentum H of both Venus and the spacecraft about the Sun is given by

$$H = m_{\varphi} r_{\varphi} v_{\varphi} + m r_{\varphi} v \tag{12.115}$$

Again we make the simplifying assumption that the planets sphere of influence radius is negligible compared to the heliocentric orbit radius. During the planetary fly-by, the total change in angular momentum must be zero.

$$\Delta H = 0 = m_{\varphi} r_{\varphi} \Delta v_{\varphi} + m r_{\varphi} \Delta v \tag{12.116}$$

Thus, for a given spacecraft velocity change Δv, the corresponding change in the planet's heliocentric velocity is

$$\Delta v_{\varphi} = \frac{m}{m_{\varphi}} \Delta v \tag{12.117}$$

Since $m/m_{\varphi} \to 0$, the change in the planets velocity can be ignored for planetary fly-by maneuvers.

Problems

12.1 Assume a spacecraft is to enter an interplanetary mission to either Mars, Venus or Jupiter and requires an initial heliocentric departure velocity v_1 from the Earth sphere of influence. Compute the sensitivities of errors in the initial hyperbolic injection burn $\Delta\nu_0$ if the parking orbit is at an altitude of 250 km. Perform this calculation for each of the three planets and comment on the values.

Bibliography

[1] Battin, R. H., *An Introduction to the Mathematics and Methods of Astrodynamics*, AIAA Education Series, New York, 1987.

[2] Hohmann, W., *Die Erreichbarkeit der Himmelskörper*, R. Oldenbourg, Munich, Germany, 1925.

[3] McLaughlin, W. I., "Walter Hohmann's Roads in Space," *Journal of Space Mission Architecture*, , No. 2, Fall 2000, pp. 1–14.

[4] Bate, R., Mueller, D. D., and White, J. E., *Fundamentals of Astrodynamics*, Dover Publications, Inc., New York, NY, 1971.

Chapter Thirteen

Spacecraft Formation Flying

Spacecraft formation flying concepts have been studied since the beginning of the manned space program. The challenge at that time was to have two-spacecraft rendez-vous and dock onto each other. This was particularly crucial for the Apollo space program which had the final lunar spacecraft being assembled in orbit. During this maneuver orbit corrections are performed not to correct the Earth relative orbit itself, but rather to adjust and control the *relative* orbit between two vehicles. For the docking maneuver, the relative distance is decreased to zero in a very slow and controlled manner.

The modern day focus of spacecraft formation flying has now extended to maintain a formation of various spacecraft. For example, the U.S. Air Force is studying concepts of having a cluster of identical satellites form a sparse aperture radar dish in space. Having multiple satellites flying at a specific geometry avoids the significant technical and financial challenge of attempting to build a radar dish of the equivalent size. These satellite formations can have diameters ranging from several dozens of meters to several kilometers. Attempting to build, control and navigate a light-weight radar dish structure that could span several kilometers would be very challenging and not cost effective. Instead, having a multitude of satellites form a virtual radar dish has the advantage of avoiding the structural flexing issues of the large dish structure and the associated pointing difficulties.

A conceptual difference between the formation flying problems that result in two or more vehicles docking and the spacecraft formation flying problem of maintaining the relative orbit of a cluster of satellites is that the later is significantly more sensitive to relative orbit modeling errors. If the satellites involved are being navigated to a rendez-vous, then the formation flying period of the two vehicles is relatively limited compared to the lifetime of the vehicle itself. Typically, the rendez-vous and docking maneuvers occur over 1-2 orbits. Thus, from a control perspective, if the relative orbit description contains some minor

simplifying assumptions, then this will have a minimal impact on the control performance. The feedback control laws are robust enough to compensate for such modeling errors and will guide the spacecraft involved to a safe docking. Also, as the two vehicles approach each other, the relative distance becomes smaller and smaller. Thus any errors introduced into the relative motion description by making linearizing assumptions become negligible during the final docking phase.

However, for the task of maintaining a spacecraft relative orbit formation, where a cluster of satellites are supposed to continuously orbit each other, making linearizing assumption can potentially lead to a substantially higher fuel cost. The reason is that this formation is supposed to be maintained over the *entire* life span of the satellites. If a relative orbit is designed using a very simplified orbit model, then the formation station keeping control law will need to continuously compensate for these modeling errors and burn fuel. Depending on the severity of the modeling errors, this fuel consumption could drastically reduce the lifetime of the spacecraft formation. It is precisely this sensitivity to the orbital dynamics that makes this type of formation flying problem very interesting from the celestial mechanics point of view.

There are two types of spacecraft formations that are being considered here. One case has the satellite cluster consisting of spacecraft of different type and built. This results in each craft having a different ballistic coefficient. Thus, the orbit of each vehicle will decay at a different rate due to the drag difference between the orbits. The main challenge for the station keeping control law of such spacecraft formations is to assure that all the orbits of each satellite decay on average at an equal rate. While this is a challenging control task, it is not that interesting from a dynamics or celestial mechanics point of view since these uncontrolled relative orbits are not closed relative orbits (i.e. they don't repeat each orbit). The second type of spacecraft formations consist of a cluster of satellites of equal type and built. Here each satellite ideally has the same ballistic coefficient. Thus each orbit will decay nominally at the same rate. For this case it is possible to analytically find closed relative orbits. These relative orbits describe a fixed geometry as seen by the rotating spacecraft reference frame. This frame will be defined more carefully in the next section.

This chapter develops the relative orbit descriptions of two or more satellites for both circular and elliptic reference orbits. All satellite constellations are assumed to be comprised of spacecraft of equal type and built. Thus, the relative drag effect will only have a secondary effect on the relative orbit. The dominant dynamical effect studied in this chapter will be the gravitational attraction of the planet. In particular, the effect of both a spherical or oblate body are considered. Finally, some relative orbit control laws will be presented that are able to establish and maintain a desired relative orbit among the spacecraft. While the relative orbit equations of motion developed here could be used to develop rendez-vous or docking control laws, this type of spacecraft formation flying is not specifically discussed in this paper. An excellent survey of the spacecraft rendez-vous problem is provided by Carter in Reference 1.

13.1 General Relative Orbit Description

This section develops the relative orbit equations of motion and presents methods to establish closed relative orbits. Both the Cartesian coordinates and the orbit element description will be used. The spacecraft formation flying nomenclature used in this chapter is as follows. The simplest type of spacecraft formation flying geometry is the leader-follower type of formation flying shown in Figure 13.1. Here the two spacecraft are essentially in identical orbits, but are separated only by having different anomalies. If this orbit is circular, then the spacecraft separation will remained fixed since both vehicles are always moving at the same orbital speed. If the orbit is elliptical, then the spacecraft separation will contract and expand, depending on whether the formation is approaching the orbit apoapses or periapses.

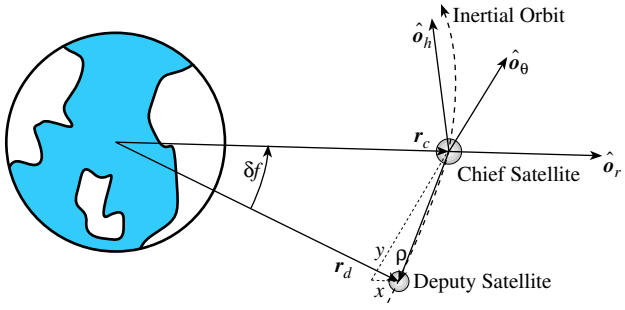


Figure 13.1: Illustration of a Leader-Follower Type of a Two-Spacecraft Formation

The satellite about which all other satellites are orbiting is referred to as the chief satellite. The remaining satellites, referred to as the deputy satellites, are to fly in formation with the chief. Note that it is not necessary that the chief position actually be occupied by a physical satellite. Sometimes this chief position is simply used as an orbiting reference point about which the deputy satellites orbit.

The inertial chief position is expressed through the vector $\boldsymbol{r}_c(t)$, while the deputy satellite position is given by $\boldsymbol{r}_d(t)$. To express how the relative orbit geometry is seen by the chief, we introduce the Hill coordinate frame.[2] Its origin is at the chief satellite position and its orientation is given by the vector triad $\{\hat{\boldsymbol{o}}_r, \hat{\boldsymbol{o}}_\theta, \hat{\boldsymbol{o}}_h\}$ shown in Figures 13.1 and 13.2. The vector $\hat{\boldsymbol{o}}_r$ is in the orbit radius direction, while $\hat{\boldsymbol{o}}_h$ is parallel to the orbit momentum vector in the orbit normal direction. The vector $\hat{\boldsymbol{o}}_\theta$ then completes the right-handed coordinates system. Mathematically, these $\mathcal{O}$ frame orientation vectors are expressed as

$$\hat{\boldsymbol{o}}_r = \frac{\boldsymbol{r}_c}{r_c} \tag{13.1a}$$

$$\hat{\boldsymbol{o}}_\theta = \hat{\boldsymbol{o}}_h \times \hat{\boldsymbol{o}}_r \tag{13.1b}$$

$$\hat{\boldsymbol{o}}_h = \frac{\boldsymbol{h}}{h} \tag{13.1c}$$

with $\boldsymbol{h} = \boldsymbol{r}_c \times \dot{\boldsymbol{r}}_c$. Note that if the inertial chief orbit is circular, then $\hat{\boldsymbol{o}}_\theta$ is parallel to the satellite velocity vector. This rotating reference frame is sometimes also referred to as the Hill frame.

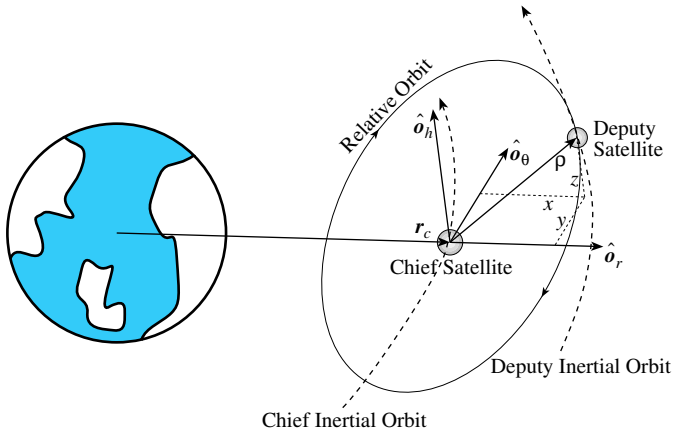


Figure 13.2: Illustration of a General Type of Spacecraft Formation with Out-Of-Orbit Plane Relative Motion

A dynamically more challenging type of general spacecraft formation flying than the leader-follower type is shown in Figure 13.2. The relative orbit position vector $\boldsymbol{\rho}$ is expressed in $\mathcal{O}$ frame components as

$$\boldsymbol{\rho} = (x, y, z)^T \tag{13.2}$$

Here the various spacecraft are on slightly different orbits that will satisfy some specific constraints. These constraints ensure that the relative orbit is bounded and that the spacecraft will not drift apart. With these types of orbit, the chief satellite (or chief position) is the relative orbit interior point about which all the other deputy satellites are orbiting.

13.2 Cartesian Coordinate Description

In this section we chose to describe the relative orbit in terms of the Cartesian coordinate vector $\boldsymbol{\rho} = (x, y, z)^T$. The vector components are taken in the rotating chief Hill frame. The advantage of using Hill frame coordinates is that the physical relative orbit dimensions are immediately apparent from these coordinates. The (x, y) coordinates define the relative orbit motion in the chief orbit plane. The z coordinate defines any motion out of the chief orbit plane.

13.2.1 Clohessy-Wiltshire Equations

To derive the relative equations of motion using Cartesian coordinates in the rotating Hill frame, we write the deputy satellite position vector as

$$\boldsymbol{r}_d = \boldsymbol{r}_c + \boldsymbol{\rho} = (r_c + x)\hat{\boldsymbol{o}}_r + y\hat{\boldsymbol{o}}_\theta + z\hat{\boldsymbol{o}}_h \tag{13.3}$$

where r_c is the current orbit radius of the chief satellite. The angular velocity vector of the rotating Hill frame $\mathcal{O}$ frame relative to the inertial $\mathcal{N}$ frame is given by

$$\boldsymbol{\omega}_{\mathcal{O}/\mathcal{N}} = \dot{f}\hat{\boldsymbol{\omega}}_h \tag{13.4}$$

with f being the chief frame true anomaly. Taking two derivatives with respect to the inertial frame, the deputy satellite acceleration vector is given by

$$\begin{aligned}\ddot{\boldsymbol{r}}_d = &\left(\ddot{r}_c + \ddot{x} - 2\dot{y}\dot{f} - \ddot{f}y - \dot{f}^2(r_c + x)\right)\hat{\boldsymbol{o}}_r \\ &+ \left(\ddot{y} + 2\dot{f}(\dot{r}_c + \dot{x}) + \ddot{f}(r_c + x) - \dot{f}^2 y\right)\hat{\boldsymbol{o}}_\theta + \ddot{z}\hat{\boldsymbol{o}}_h\end{aligned} \tag{13.5}$$

This kinematic expression can be simplified by making use of the following identities. The chief orbit angular momentum magnitude is given by $h = r_c^2\dot{f}$. Since h is constant for Keplerian motion, taking the first time derivative of h yields

$$\dot{h} = 0 = 2r_c\dot{r}_c\dot{f} + r_c^2\ddot{f} \tag{13.6}$$

This orbit element constraint can be used to solve for the true anomaly acceleration.

$$\ddot{f} = -2\frac{\dot{r}_c}{r_c}\dot{f} \tag{13.7}$$

Further, we write the chief satellite position as $\boldsymbol{r}_c = r_c\hat{\boldsymbol{o}}_r$. Taking two time derivatives with respect to the inertial frame and using the orbit equations of motion, the chief acceleration vector is expressed as

$$\ddot{\boldsymbol{r}}_c = \left(\ddot{r}_c - r_c\dot{f}^2\right)\hat{\boldsymbol{o}}_r = -\frac{\mu}{r_c^3}\boldsymbol{r}_c = -\frac{\mu}{r_c^2}\hat{\boldsymbol{o}}_r \tag{13.8}$$

Equating vector components in Eq. (13.8), the chief orbit radius acceleration is expressed as:

$$\ddot{r}_c = r_c\dot{f}^2 - \frac{\mu}{r_c^2} = r_c\dot{f}^2\left(1 - \frac{r_c}{p}\right) \tag{13.9}$$

Substituting Eqs. (13.7) and (13.9) into Eq. (13.5), the deputy acceleration vector expression is reduced to

$$\begin{aligned}\ddot{\boldsymbol{r}}_d = &\left(\ddot{x} - 2\dot{f}\left(\dot{y} - y\frac{\dot{r}_c}{r_c}\right) - x\dot{f}^2 - \frac{\mu}{r_c^2}\right)\hat{\boldsymbol{o}}_r \\ &+ \left(\ddot{y} + 2\dot{f}\left(\dot{x} - x\frac{\dot{r}_c}{r_c}\right) - y\dot{f}^2\right)\hat{\boldsymbol{o}}_\theta + \ddot{z}\hat{\boldsymbol{o}}_h\end{aligned} \tag{13.10}$$

Next, we substitute the kinematic acceleration expression in Eq. (13.10) into the orbit equations of motion. The deputy satellite orbital equations of motion are given by

$$\ddot{\boldsymbol{r}}_d = -\frac{\mu}{r_d^3}\boldsymbol{r}_d = -\frac{\mu}{r_d^3}\overset{\mathcal{O}}{\begin{pmatrix} r_c + x \\ y \\ z \end{pmatrix}} \tag{13.11}$$

with $r_d = \sqrt{(r_c + x)^2 + y^2 + z^2}$. Equating Eqs. (13.10) and (13.11), the exact nonlinear relative equations of motion are given by

$$\ddot{x} - 2\dot{f}\left(\dot{y} - y\frac{\dot{r}_c}{r_c}\right) - x\dot{f}^2 - \frac{\mu}{r_c^2} = -\frac{\mu}{r_d^3}(r_c + x) \tag{13.12a}$$

$$\ddot{y} + 2\dot{f}\left(\dot{x} - x\frac{\dot{r}_c}{r_c}\right) - y\dot{f}^2 = -\frac{\mu}{r_d^3}y \tag{13.12b}$$

$$\ddot{z} = -\frac{\mu}{r_d^3}z \tag{13.12c}$$

The only assumption which has been made is that no disturbances are acting on the satellites and thus the Keplerian motion assumption in the orbital equations of motion in Eq. (13.11) are correct. The relative equations of motion in Eq. (13.12) are valid for arbitrarily large relative orbits and the chief orbit may be eccentric. If the relative orbit coordinates (x, y, z) are small compared to the chief orbit radius r_c, then Eq. (13.12) can be further simplified. The deputy orbit radius r_d is approximated as

$$r_d = r_c\sqrt{1 + 2\frac{x}{r_c} + \frac{x^2 + y^2 + z^2}{r_c^2}} \approx r_c\sqrt{1 + 2\frac{x}{r_c}} \tag{13.13}$$

This allows us to write

$$\frac{\mu}{r_d^3} \approx \frac{\mu}{r_c^3}\left(1 - 3\frac{x}{r_c}\right) \tag{13.14}$$

The term μ/r_c^3 can also be written in the following useful forms:

$$\frac{\mu}{r_c^3} = \frac{r_c}{p}\dot{f}^2 = \frac{\dot{f}^2}{1 + e\cos f} \tag{13.15}$$

Note that the orbit elements shown in Eq. (13.15) are chief orbit elements. Neglecting higher order terms, we are able to simplify the right hand side of Eq. (13.11) to

$$-\frac{\mu}{r_d^3}\overset{\mathcal{O}}{\begin{pmatrix} r_c + x \\ y \\ z \end{pmatrix}} \approx \frac{\mu}{r_c^3}\left(1 - 3\frac{x}{r_c}\right)\overset{\mathcal{O}}{\begin{pmatrix} r_c + x \\ y \\ z \end{pmatrix}} \approx \frac{\mu}{r_c^3}\overset{\mathcal{O}}{\begin{pmatrix} r_c - 2x \\ y \\ z \end{pmatrix}} \tag{13.16}$$

Substituting Eq. (13.16) into Eq. (13.12) and simplifying the resulting expressions yields the relative orbit equations of motion assuming that x, y, z are small compared to chief orbit radius r_c.

$$\ddot{x} - x\dot{f}^2\left(1 + 2\frac{r_c}{p}\right) - 2\dot{f}\left(\dot{y} - y\frac{\dot{r}_c}{r_c}\right) = 0 \tag{13.17a}$$

$$\ddot{y} + 2\dot{f}\left(\dot{x} - x\frac{\dot{r}_c}{r_c}\right) - y\dot{f}^2\left(1 - \frac{r_c}{p}\right) = 0 \tag{13.17b}$$

$$\ddot{z} + \frac{r_c}{p}\dot{f}^2 z = 0 \tag{13.17c}$$

Using Eqs. (13.7) and (13.15), along with the true latitude $\theta = \omega + f$, the general relative equations of motion are rewritten in the common form:[3]

$$\ddot{x} - x\left(\dot{\theta}^2 + 2\frac{\mu}{r_c^3}\right) - y\ddot{\theta} - 2\dot{y}\dot{\theta} = 0 \tag{13.18a}$$

$$\ddot{y} + x\ddot{\theta} + 2\dot{x}\dot{\theta} - y\left(\dot{\theta}^2 - \frac{\mu}{r_c^3}\right) = 0 \tag{13.18b}$$

$$\ddot{z} + \frac{\mu}{r_c^3} z = 0 \tag{13.18c}$$

If the chief satellite orbit is assumed to be circular, then $e = 0$, $p = r_c$, and the chief orbit radius r_c is constant. Since for a circular orbit the mean orbital rate n is equal to the true anomaly rate $\dot{f}$, the relative equations of motion reduce to the simple form known as the *Clohessy-Wiltshire* (CW) equations.[2, 4]

$$\ddot{x} - 2n\dot{y} - 3n^2 x = 0 \tag{13.19a}$$

$$\ddot{y} + 2n\dot{x} = 0 \tag{13.19b}$$

$$\ddot{z} + n^2 z = 0 \tag{13.19c}$$

Note that these equations of motion are only valid if the chief orbit is circular and the relative orbit coordinates (x, y, z) are small compared to the chief orbit radius r_c. The simple form of the differential equations in Eq. (13.19) allows them to be analytically integrated to find closed form solutions to the relative equations of motion. For example, the differential equations of motion for the relative orbit out-of-plane motion, shown in Eq. (13.17c), is that of a simple spring-mass system which has a known solution. This development of the analytic relative equations of motion solution is shown in the section 13.2.2.

The general relative equations of motion, shown in Eq. (13.17), take on a very elegant form if written in a non-dimensional form. Let us define the non-dimensional relative orbit coordinates (u, v, w) as

$$u = \frac{x}{r_c} \qquad v = \frac{y}{r_c} \qquad w = \frac{z}{r_c} \tag{13.20}$$

Instead of differentiating with respect to time, we now differentiate with respect to the chief orbit true anomaly f. This type of differentiation is written here as

$$()' \equiv \frac{d()}{df} \tag{13.21}$$

To obtain the non-dimensional relative equations of motion, the following identities relating time derivatives of (x, y, z) coordinates to corresponding non-dimensional derivatives of (u, v, w) will be used:

$$\frac{\dot{x}}{r_c} = u'\dot{f} + u\frac{\dot{r}_c}{r_c} \qquad \frac{\ddot{x}}{r_c} = u''\dot{f}^2 + u\dot{f}^2\left(1 - \frac{r_c}{p}\right) \tag{13.22a}$$

$$\frac{\dot{y}}{r_c} = v'\dot{f} + v\frac{\dot{r}_c}{r_c} \qquad \frac{\ddot{y}}{r_c} = v''\dot{f}^2 + v\dot{f}^2\left(1 - \frac{r_c}{p}\right) \tag{13.22b}$$

$$\frac{\dot{z}}{r_c} = w'\dot{f} + w\frac{\dot{r}_c}{r_c} \qquad \frac{\ddot{z}}{r_c} = w''\dot{f}^2 + w\dot{f}^2\left(1 - \frac{r_c}{p}\right) \tag{13.22c}$$

Dividing the dimensional equations of motion in Eq. (13.17) by the chief orbit radius r_c, substituting the identities in Eq. (13.22) and simplifying leads to the following elegant non-dimensional relative equations of motion:

$$u'' - 2v' - \frac{3u}{1 + e\cos f} = 0 \tag{13.23a}$$

$$v'' + 2u' = 0 \tag{13.23b}$$

$$w'' + w = 0 \tag{13.23c}$$

The above relative equations of motion are valid for eccentric chief orbits, as long as $(u, v, w) \ll 1$. Comparing Eq. (13.19) to Eq. (13.23), it is clear that the form of the non-dimensional equations of motion is very close to that of the Clohessy-Wiltshire equations. The only algebraic difference is the additional fraction in the non-dimensional radial equations of motion in Eq. (13.23a).

13.2.2 Closed Relative Orbits in the Hill Reference Frame

Starting with the Clohessy-Wiltshire equations in Eq. (13.19), we would like to find constraints for the relative orbit coordinates (x, y, z) which will guarantee that the relative orbit geometry will remain bounded. The underlying assumption here is that the chief orbit is circular and perturbations to the Keplerian motion can be ignored. The relative equations of motion in terms of (x, y, z) are repeated here for convenience:

$$\ddot{x} - 2n\dot{y} - 3n^2x = 0 \tag{13.24a}$$

$$\ddot{y} + 2n\dot{x} = 0 \tag{13.24b}$$

$$\ddot{z} + n^2z = 0 \tag{13.24c}$$

The z component decouples here from the radial and along track directions and has the form of a simple un-forced oscillator differential equation. It's

general solution is given by

$$z(t) = B_0 \cos(nt + \beta) \tag{13.25}$$

where B_0 and β are integration constants which are determined through the initial conditions. Further, Eq. (13.24b) is of a perfect integrable form leading to

$$\dot{y} = -2nx + d \tag{13.26}$$

where the integration constant $d = \dot{y}_0 + 2nx_0$ is defined through the initial conditions. Substituting Eq. (13.26) into Eq. (13.24a), the equation of motion in the x direction is written as the forced oscillator differential equation

$$\ddot{x} + n^2 x = 2nd \tag{13.27}$$

Solving this differential equation, the radial position component is given by

$$x(t) = A_0 \cos(nt + \alpha) + \frac{2d}{n} \tag{13.28}$$

where A_0 and α are determined through the initial conditions. Defining the scalar offset in the orbit radial direction as

$$x_{off} = \frac{2d}{n} \tag{13.29}$$

the $x(t)$ equation can be written as

$$x(t) = A_0 \cos(nt + \alpha) + x_{off} \tag{13.30}$$

Substituting Eq. (13.28) into Eq. (13.26), the first order differential equation for $y(t)$ is written as

$$\dot{y} = -2nA_0 \cos(nt + \alpha) - 3d \tag{13.31}$$

Integrating this differential equation and using y_{off} as the integration constant, $y(t)$ is written as

$$y(t) = -2A_0 \sin(nt + \alpha) - 3dt + y_{off} \tag{13.32}$$

Thus the analytical solution to the CW equations in Hill frame components is summarized as

$$x(t) = A_0 \cos(nt + \alpha) + x_{off} \tag{13.33a}$$

$$y(t) = -2A_0 \sin(nt + \alpha) - \frac{3}{2} nt x_{off} + y_{off} \tag{13.33b}$$

$$z(t) = B_0 \cos(nt + \beta) \tag{13.33c}$$

Note the the expression for $y(t)$ contains a secular term which will grow infinitely large as $t \to \infty$. All other terms in Eq. (13.33) are either sin or cos functions,

or a constant offset. Thus, to avoid the secular growth in $y(t)$ we must set x_{off} equal to zero. This is equivalent to setting the integration constant d, defined in Eq. (13.29), equal to zero. The bounded relative orbit constraint for a circular chief orbit is given in terms of Hill frame coordinates as:

$$d = \dot{y}_0 + 2nx_0 = 0 \tag{13.34}$$

Note that this requirement is identical to assuring that $x_{off} = 0$. The analytical solutions for bounded solutions of the CW equations are then given by

$$x(t) = A_0 \cos(nt + \alpha) \tag{13.35a}$$
$$y(t) = -2A_0 \sin(nt + \alpha) + y_{off} \tag{13.35b}$$
$$z(t) = B_0 \cos(nt + \beta) \tag{13.35c}$$

Example 13.1: For the sparse aperture type of spacecraft formation mission, the sensor requirements specify that the (y, z) projection of the relative orbit be circular. Studying Eq. (13.35), the (y, z) trajectory will only describe a circle if the along track offset y_{off} is to zero. Further, the along track sinusoidal amplitude A_0 and the out-of-plane sinusoidal amplitude B_0 must satisfy

$$B_0 = 2A_0$$

Note that the projection of the relative orbit onto the (x, y) plane always forms an ellipse which is twice as large in the along track direction than in the radial direction. Further, the phase angles α and β must satisfy

$$\alpha = \beta \qquad \text{or} \qquad \alpha = \beta + 180^{\circ}$$

Adding 180 degrees to the phase angle difference makes the relative orbit have a different relative inclination. This is illustrated in Figure 13.3. Both relative orbits are computed using $A_0 = 0.5$ km, $B_0 = 1.0$ km and $\alpha = 0$ degrees.

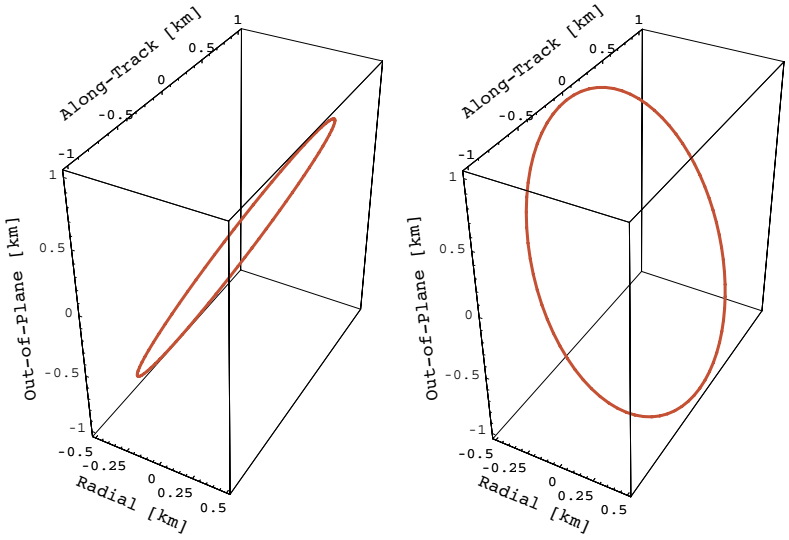
Therefore, assuming $B_0 = 2A_0$ and $\alpha = \beta$, then the relative orbit trajectories of interest are given in Hill frame vector components by

$$\boldsymbol{\rho} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = A_0 \begin{pmatrix} \cos(nt + \alpha) \\ -2\sin(nt + \alpha) \\ 2\cos(nt + \alpha) \end{pmatrix}$$

with the local velocity as seen in the Hill frame being

$$\frac{^{\mathcal{O}}d(\boldsymbol{\rho})}{dt} = \begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} = -A_0 n \begin{pmatrix} \sin(nt + \alpha) \\ 2\cos(nt + \alpha) \\ 2\sin(nt + \alpha) \end{pmatrix}$$

To map the relative vectors $\boldsymbol{\rho}$ and $\dot{\boldsymbol{\rho}}$ back to inertial position and velocity vectors of the deputy satellite, it is assumed that the inertial position and velocity vectors of the chief are given. Using the Hill frame orientation vectors



(i) Relative Orbit with $\beta = \alpha$

(ii) Relative Orbit with $\beta = \alpha + 180^o$

Figure 13.3: Relative Orbit Comparison in the Hill Reference Frame for Different Phase Angle Differences

defined in Eq. (13.1), the direction cosine matrix which relates inertial frame vector components to Hill frame components is computed through

$$[ON] = \begin{bmatrix} \hat{\boldsymbol{o}}_r^T \\ \hat{\boldsymbol{o}}_\theta^T \\ \hat{\boldsymbol{o}}_h^T \end{bmatrix}$$

The inertial deputy position vector $\boldsymbol{r}_d$ is then computed as

$$\boldsymbol{r}_d = \boldsymbol{r}_c + [ON]^T \boldsymbol{\rho}$$

where it is assumed that $\boldsymbol{r}_c$ is given in inertial components. The inertial deputy satellite velocity is found through

$$\dot{\boldsymbol{r}}_d = \dot{\boldsymbol{r}}_c + [ON]^T \left(\frac{^{\mathcal{O}}d(\boldsymbol{\rho})}{dt} + n\hat{\boldsymbol{o}}_h \times \boldsymbol{\rho} \right)$$

13.3 Orbit Element Difference Description

While using the Hill frame coordinates (x, y, z) is a common method to describe a relative orbit, they have the distinct disadvantage that their differential equations must be solved in order to obtain the relative orbit geometry. The relative

orbit is determined through the chief orbit motion and the relative orbit initial conditions

$$\boldsymbol{X} = (x_0, y_0, z_0, \dot{x}_0, \dot{y}_0, \dot{z}_0)^T \tag{13.36}$$

To find out where a deputy satellite would be at time t, the appropriate differential equations in either Eq. (13.17) or Eq. (13.19) need to be integrated forward to time t. Thus, the six initial conditions in Eq. (13.36) form six invariant quantities of the relative orbit motion. However, they are not convenient to determine the instantaneous geometry of the relative orbit motion. However, if the chief orbit is circular, then the elegant CW equations have an analytical solution shown in Eq. (13.35). This lead to the initial condition constraint in Eq. (13.34) which guarantees bounded relative motion. However, this constraint is only valid if the relative orbit dimension is small compared to the chief orbit radius (linearizing assumptions are valid), and if the chief orbit is circular ($e \to 0$).

Instead of using the six relative orbit invariants shown in Eq. (13.36) to define the relative orbit and obtain a bounded relative orbit constraint, we would like to investigate other relative orbit invariant parameters that would yield equivalent results without the need for the linearizing assumptions and near-circular chief orbit requirements.

To do this, we first review how the inertial orbits are described and solved for in a Keplerian two-body system. Let $\boldsymbol{r}$ be the inertial position vector of a spacecraft about a spherical planet, then the differential equations to be solved are given by

$$\ddot{\boldsymbol{r}} = -\frac{\mu}{r^3}\boldsymbol{r} \tag{13.37}$$

with the initial conditions being $\boldsymbol{r}(t_0) = \boldsymbol{r}_0$ and $\dot{\boldsymbol{r}}(t_0) = \dot{\boldsymbol{r}}_0$. These six initial conditions form the six invariant parameter of this dynamical system. However, as is shown in Chapter 8, the Keplerian motion of a satellite can also be defined through six orbit elements. For example, let us define the orbit element vector $\boldsymbol{e}$ to contain the parameters

$$\boldsymbol{e} = (a, e, i, \Omega, \omega, M_0)^T \tag{13.38}$$

where a is the semi-major axis, e is the eccentricity, i is the orbit inclination angle, Ω is the longitude of the ascending Node, ω is the argument of the pericenter and M_0 is the initial mean anomaly. Instead of solving a differential equation to find the current satellite states, the algebraic Kepler's equation must be numerically solved to find the current mean anomaly angle. Thus there is essentially only one state M that must be solved to find the satellite position. Compare this to using the $\boldsymbol{X}$ state vector. Here all six states are fast variables, meaning that they vary throughout the orbit. Using the orbit elements thus simplifies the orbit description and the satellite position computation. Further, note that even with disturbances present such as gravitational perturbations or

atmospheric or solar drag, these orbit elements will only change slowly. This is illustrated in Chapter 11.

Assuming for now that no disturbances are present, then the six orbit elements are invariant unless some control thrust is applied to the spacecraft. Instead of defining the relative orbit in terms of the six Hill frame coordinates in Eq. (13.36), let us propose to define the relative orbit in terms of the orbit element difference vector $\delta\boldsymbol{e}$

$$\delta\boldsymbol{e} = \boldsymbol{e}_d - \boldsymbol{e}_c = (\delta a, \delta e, \delta i, \delta\Omega, \delta\omega, \delta M_0)^T \tag{13.39}$$

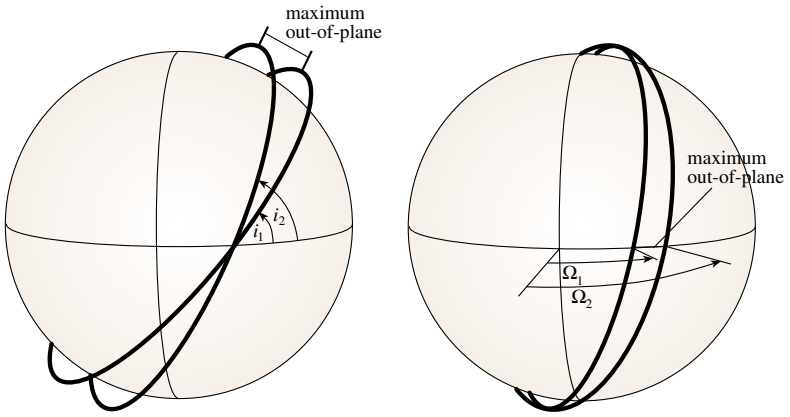
where $\boldsymbol{e}_d$ is the deputy satellite orbit element vector and $\boldsymbol{e}_c$ is the chief orbit element vector. Note that this relative orbit description using orbit element differences is not constraint to the particular orbit elements used here. Any complete set of orbit elements could be used. Given $\delta\boldsymbol{e}$ and $\boldsymbol{e}_c$, the deputy satellite position can be computed at any instance of time by solving Kepler's equation. As is the case with the inertial orbit description, we are able to avoid having to solve a differential equation. Note that the relative orbit description in Eq. (13.39) does not make any assumptions on how large the relative orbit is compared to the chief orbit radius, nor does it require that the chief orbit is circular.

Working with orbit element differences also provides some insight into the orbit geometry itself. Simply starting out with the Hill frame initial conditions in Eq. (13.36), the relative orbit geometry is only determined after solving the differential equations. However, by describing the relative orbit in terms of orbit element differences, it is possible to make certain statements regarding the relative orbit geometry. This concept is illustrated in Figure 13.4. Both the inclination angle and ascending node differences will affect the magnitude of the out-of-plane motion of the relative orbit. The inclination angle difference δi specifies how much out-of-plane motion the relative orbit will have as the the satellite cross the northern or southern most regions. The ascending node difference shown what the out-of-plane motion will be as the satellite crosses the equator. For example, if corresponding orbit element differences are computed for a given relative orbit and δi is found to be zero, then it can immediately be concluded that the relative orbit will have zero out-of-plane motion as the chief passes the outer latitude extremes.

13.3.1 Linear Mapping Between Hill Frame Coordinates and Orbit Element Differences

To map between the $\boldsymbol{X}$ and $\delta\boldsymbol{e}$ relative orbit coordinates, the nonlinear mapping between inertial coordinates and orbit elements shown in Chapter 8 could be used. However, if the relative orbit is small compared to the chief orbit radius, then it is possible to obtain a direct linear mapping $[A(\boldsymbol{e}_c)]$ between $\boldsymbol{X}$ and $\delta\boldsymbol{e}$.

$$\boldsymbol{X} = [A(\boldsymbol{e}_c)]\delta\boldsymbol{e} \tag{13.40}$$



(i) Inclination Angle Difference **(ii)** Ascending Node Difference

Figure 13.4: Relative Orbit Effect of Having Specific Orbit Element Differences

To avoid some numerical difficulties for near circular orbits, let us use the orbit element vector $\boldsymbol{e}$ is defined through

$$\boldsymbol{e} = (a, \theta, i, q_1, q_2, \Omega)^T \tag{13.41}$$

with θ being the true latitude angle (sum of argument of perigee and true anomaly), and q_1 and q_2 being defined through

$$q_1 = e\cos\omega \tag{13.42}$$

$$q_2 = e\sin\omega \tag{13.43}$$

Let us define the following three coordinates systems. Let $\mathcal{C}$ and $\mathcal{D}$ be the Hill coordinate frames of the chief and deputy satellites, respectively, and let $\mathcal{N}$ be the inertial frame. Then $[\mathcal{CN}] = [\mathcal{CN}(\Omega_c, i_c, \theta_c)]$ is the direction cosine matrix mapping vector components in the inertial frame to components in the chief Hill frame. To relate the orbit element difference vector $\delta\boldsymbol{e}$ to the corresponding LVLH Cartesian coordinate vector $\boldsymbol{X}$, we write the deputy spacecraft inertial position vector $\boldsymbol{r}_d$ in chief and deputy Hill frame components as

$${}^{\mathcal{C}}\boldsymbol{r}_d = {}^{\mathcal{C}}(r_c + x, y, z)^T \tag{13.44}$$

$${}^{\mathcal{D}}\boldsymbol{r}_d = {}^{\mathcal{D}}(r_d, 0, 0)^T \tag{13.45}$$

where r is the inertial orbit radius. The deputy position vector $\boldsymbol{r}_d$ is now mapped from the deputy Hill frame vector components to the chief Hill frame vector components using

$${}^{\mathcal{C}}\boldsymbol{r}_d = [\mathcal{CN}][\mathcal{ND}]\ {}^{\mathcal{D}}\boldsymbol{r}_d \tag{13.46}$$

To simplify the notation from here on, the subscript c is dropped and any parameter without a subscript is implied to be a chief orbit parameter. Taking the first variation of $[\mathcal{ND}]$ and r_d about the chief satellite motion leads to the first-order approximations

$$[\mathcal{ND}] \approx [\mathcal{NC}] + [\delta\mathcal{NC}] \tag{13.47}$$

$$r_d \approx r + \delta r \tag{13.48}$$

Eq. (13.46) is then expanded to yield

$${}^{\mathcal{C}}\boldsymbol{r}_d = ([I_{3\times 3}] + [\mathcal{CN}][\delta\mathcal{NC}]) \begin{pmatrix} r+\delta r \\ 0 \\ 0 \end{pmatrix} \tag{13.49}$$

Dropping second-order terms, the deputy position vector is written as

$${}^{\mathcal{C}}\boldsymbol{r}_d = \begin{pmatrix} r+\delta r \\ 0 \\ 0 \end{pmatrix} + r\ [\mathcal{CN}] \begin{pmatrix} \delta\mathcal{NC}_{11} \\ \delta\mathcal{NC}_{21} \\ \delta\mathcal{NC}_{31} \end{pmatrix} \tag{13.50}$$

with the matrix components $\delta\mathcal{NC}_{i1}$ given by

$$\delta\mathcal{NC}_{11} = \mathcal{NC}_{12}\ \delta\theta - \mathcal{NC}_{21}\ \delta\Omega + \mathcal{NC}_{31} \sin\Omega\ \delta i \tag{13.51}$$

$$\delta\mathcal{NC}_{21} = \mathcal{NC}_{22}\ \delta\theta + \mathcal{NC}_{11}\ \delta\Omega - \mathcal{NC}_{31} \cos\Omega\ \delta i \tag{13.52}$$

$$\delta\mathcal{NC}_{31} = \mathcal{NC}_{32}\ \delta\theta + \sin\theta\cos i\ \delta i \tag{13.53}$$

Substituting Eqs. (13.51) - (13.53) into Eq. (13.50), the deputy position vector is written in terms of orbit element differences as

$${}^{\mathcal{C}}\boldsymbol{r}_d = \begin{pmatrix} r+\delta r \\ 0 \\ 0 \end{pmatrix} + r \begin{pmatrix} 0 \\ \delta\theta + \delta\Omega\cos i \\ -\cos\theta\sin i\delta\Omega + \sin\theta\delta i \end{pmatrix} \tag{13.54}$$

To be able to write Eq. (13.54) in terms of the desired orbit elements and their differences, the orbit radius r must be expressed in terms of the elements given in Eq. (13.41).

$$r = \frac{a(1-q_1^2-q_2^2)}{1+q_1\cos\theta+q_2\sin\theta} \tag{13.55}$$

Thus, the variation of r is expressed as

$$\delta r = \frac{r}{a}\delta a + \frac{V_r}{V_t} r\ \delta\theta - \frac{r}{p}(2aq_1 + r\cos\theta)\delta q_1 - \frac{r}{p}(2aq_2 + r\sin\theta)\delta q_2 \tag{13.56}$$

where the chief radial and transverse velocity components V_r and V_t are defined as

$$V_r = \dot{r} = \frac{h}{p}(q_1\sin\theta - q_2\cos\theta) \tag{13.57}$$

$$V_t = r\dot{\theta} = \frac{h}{p}(1 + q_1\cos\theta + q_2\sin\theta) \tag{13.58}$$

with h being the chief orbit momentum magnitude and p being the semilatus rectum. Comparing the chief Hill frame components of the deputy position vector descriptions in Eqs. (13.44) and (13.54), the local Cartesian Hill frame coordinates x, y and z are expressed in terms of the orbit element differences as[5, 6]

$$x = \delta r \tag{13.59a}$$

$$y = r(\delta\theta + \cos i \ \delta\Omega) \tag{13.59b}$$

$$z = r(\sin\theta \ \delta i - \cos\theta \sin i \ \delta\Omega) \tag{13.59c}$$

At this point half of the desired mappings between orbit element differences and the corresponding LVLH Cartesian coordinates have been developed. To derive the linear relationship between the orbit element differences and the Cartesian coordinate rates $(\dot{x}, \dot{y}, \dot{z})$, a similar approach as has been used to derive Eqs. (13.59a) through (13.59c) could be used. In Reference 7, the deputy velocity vector is expressed in both the chief and deputy frame. The desired Cartesian coordinate rates are then extracted by comparing the two algebraic expressions.

However, it is also possible to obtain the Cartesian coordinate rate expressions in terms of orbit element differences by differentiating Eqs. (13.59a) through (13.59c) directly with respect to time. The only time-varying quantities in these three expressions are the chief true latitude θ and the difference between deputy and chief latitude $\delta\theta$. Only the latter quantity needs special consideration. Using the conservation of angular momentum h, we express the true latitude rate $\dot{\theta}$ as

$$\dot{\theta} = \frac{h}{r^2} \tag{13.60}$$

The variation of Eq. (13.60) yields

$$\delta\dot{\theta} = \frac{\delta h}{r^2} - 2\frac{h}{r^3}\delta r \tag{13.61}$$

Using the angular momentum expression $h = \sqrt{\mu p}$, the δh variation is expressed as

$$\delta h = \frac{h}{2p}\delta p \tag{13.62}$$

where δp is given by

$$\delta p = \frac{p}{a}\delta a - 2a(q_1\delta q_1 + q_2\delta q_2) \tag{13.63}$$

Thus the desired variation in the true latitude rate is expressed as

$$\delta\dot{\theta} = \frac{h}{r^2}\left(\frac{\delta p}{2p} - 2\frac{\delta r}{r}\right) \tag{13.64}$$

After differentiating Eqs. (13.59a)-(13.59c) and making use of Eq. (13.64), the Cartesian coordinate rates are expressed in terms of orbit element differences as[5, 6]

$$\begin{aligned}\dot{x} = &-\frac{V_r}{2a}\delta a + (\frac{1}{r} - \frac{1}{p})h\delta\theta \\ &+ (V_r a q_1 + h\sin\theta)\frac{\delta q_1}{p} + (V_r a q_2 - h\cos\theta)\frac{\delta q_2}{p}\end{aligned} \tag{13.65a}$$

$$\begin{aligned}\dot{y} = &-\frac{3V_t}{2a}\delta a - V_r\delta\theta + (3V_t a q_1 + 2h\cos\theta)\frac{\delta q_1}{p} \\ &+ (3V_t a q_2 + 2h\sin\theta)\frac{\delta q_2}{p} + V_r\cos i\ \delta\Omega\end{aligned} \tag{13.65b}$$

$$\dot{z} = (V_t\cos\theta + V_r\sin\theta)\delta i + (V_t\sin\theta - V_r\cos\theta)\sin i\delta\Omega \tag{13.65c}$$

Combined, Eqs. (13.59)and (13.65) define the linear mapping $[A(\boldsymbol{e}_c)]$ from $\delta\boldsymbol{e}$ to $\boldsymbol{X}$ relative orbit states. The inverse of the matrix $[A(\boldsymbol{e})]$ is developed in an analogous manner.[6] To simplify the expressions, the following non-dimensional parameters are introduced

$$\alpha = a/r \tag{13.66}$$

$$\nu = V_r/V_t \tag{13.67}$$

$$\rho = r/p \tag{13.68}$$

$$\kappa_1 = \alpha\left(\frac{1}{\rho} - 1\right) \tag{13.69}$$

$$\kappa_2 = \alpha\nu^2\frac{1}{\rho} \tag{13.70}$$

The non-zero matrix elements of $[A(\boldsymbol{e}_c)]^{-1}$ are given by:

$$A_{11}^{-1} = 2\alpha\Big(2 + 3\kappa_1 + 2\kappa_2\Big) \tag{13.71a}$$

$$A_{12}^{-1} = -2\alpha\nu\Big(1 + 2\kappa_1 + \kappa_2\Big) \tag{13.71b}$$

$$A_{14}^{-1} = \frac{2\alpha^2\nu p}{V_t} \tag{13.71c}$$

$$A_{15}^{-1} = \frac{2a}{V_t}\Big(1 + 2\kappa_1 + \kappa_2\Big) \tag{13.71d}$$

$$A_{22}^{-1} = \frac{1}{R} \tag{13.71e}$$

$$A_{23}^{-1} = \frac{\cot i}{R}(\cos\theta + \nu\sin\theta) \tag{13.71f}$$

$$A_{26}^{-1} = -\frac{\sin\theta\cot i}{V_t} \tag{13.71g}$$

$$A_{33}^{-1} = \frac{\sin\theta - \nu\cos\theta}{R} \tag{13.71h}$$

$$A_{36}^{-1} = \frac{\cos\theta}{V_t} \tag{13.71i}$$

$$A_{41}^{-1} = \frac{\rho}{R}(3\cos\theta + 2\nu\sin\theta) \tag{13.71j}$$

$$A_{42}^{-1} = -\frac{1}{R}\Big(\rho\nu^2\sin\theta + q_1\sin 2\theta - q_2\cos 2\theta\Big) \tag{13.71k}$$

$$A_{43}^{-1} = -\frac{q_2\cot i}{R}(\cos\theta + \nu\sin\theta) \tag{13.71l}$$

$$A_{44}^{-1} = \frac{\rho\sin\theta}{V_t} \tag{13.71m}$$

$$A_{45}^{-1} = \frac{\rho}{V_t}(2\cos\theta + \nu\sin\theta) \tag{13.71n}$$

$$A_{46}^{-1} = \frac{q_2\cot i\sin\theta}{V_t} \tag{13.71o}$$

$$A_{51}^{-1} = \frac{\rho}{R}(3\sin\theta - 2\nu\cos\theta) \tag{13.71p}$$

$$A_{52}^{-1} = \frac{1}{R}\Big(\rho\nu^2\cos\theta + q_2\sin 2\theta + q_1\cos 2\theta\Big) \tag{13.71q}$$

$$A_{53}^{-1} = \frac{q_1\cot i}{R}(\cos\theta + \nu\sin\theta) \tag{13.71r}$$

$$A_{54}^{-1} = -\frac{\rho\cos\theta}{V_t} \tag{13.71s}$$

$$A_{55}^{-1} = \frac{\rho}{V_t}(2\sin\theta - \nu\cos\theta) \tag{13.71t}$$

$$A_{56}^{-1} = -\frac{q_1\cot i\sin\theta}{V_t} \tag{13.71u}$$

$$A_{63}^{-1} = -\frac{\cos\theta + \nu\sin\theta}{R\sin i} \tag{13.71v}$$

$$A_{66}^{-1} = \frac{\sin\theta}{V_t\sin i} \tag{13.71w}$$

Similarly as was done with the relative equations of motion in terms of Hill frame Cartesian coordinates, it is possible to express the linear mapping $[A(\boldsymbol{e}_c)]$ in terms of the non-dimensional coordinates (u, v, w) defined in Eq. (13.20). Often this non-dimensional form provides for more convenient algebraic expressions.[5, 8, 9] Dividing Eq. (13.59) by the orbit radius r, the non-dimensional relative orbit coordinates (u, v, w) are expressed as:

$$\begin{aligned} u = \frac{x}{r} = \frac{\delta a}{a} + \nu\ \delta\theta - (2\alpha q_1 + \cos\theta)\rho\delta q_1 \\ - (2\alpha q_2 + \sin\theta)\rho\delta q_2 \end{aligned} \tag{13.72a}$$

$$v = \frac{y}{r} = \delta\theta + \cos i\ \delta\Omega \tag{13.72b}$$

$$w = \frac{z}{r} = \sin\theta\ \delta i - \cos\theta\sin i\ \delta\Omega \tag{13.72c}$$

Instead of differentiating (u, v, w) with respect to time, we choose to use the true

latitude angle θ as the time dependent variable. Let a prime symbol indicate a derivative with respect to θ. To differentiate the expressions in Eq. (13.72), only the $\delta\theta$ terms must be given special consideration. Note that

$$\frac{\partial(\delta\theta)}{\partial\theta}\frac{d\theta}{dt} = \delta\theta'\dot{\theta} = \delta\dot{\theta} \tag{13.73}$$

Using Eq. (13.64), the partial derivative of $\delta\theta$ with respect to the true latitude is given by:

$$\delta\theta' = \frac{\delta p}{2p} - 2u \tag{13.74}$$

Taking the partial derivative of Eqs. (13.72a)-(13.72c) while making use of Eq. (13.74) yields the following non-dimensional rates with respect to true latitude.

$$\begin{aligned} u' = &-\frac{3}{2}\nu\frac{\delta a}{a} + \left(\rho(q_1\cos\theta + q_2\sin\theta) - \nu^2\right)\delta\theta \\ &+ (3\nu\alpha q_1 + \sin\theta + \nu\cos\theta)\,\rho\;\delta q_1 \\ &+ (3\nu\alpha q_2 + \cos\theta - \nu\sin\theta)\,\rho\;\delta q_2 \end{aligned} \tag{13.75a}$$

$$\begin{aligned} v' = &-\frac{3}{2}\frac{\delta a}{a} - 2\nu\delta\theta + (2\cos\theta + 3\alpha q_1)\,\rho\;\delta q_1 \\ &+ (2\sin\theta + 3\alpha q_2)\,\rho\;\delta q_2 \end{aligned} \tag{13.75b}$$

$$w' = \cos\theta\delta i + \sin\theta\sin i\delta\Omega \tag{13.75c}$$

Note that these non-dimensional rate expressions are not necessarily simpler than their their dimensional counterparts. To map these rates with respect to true latitude into the corresponding dimensional (x, y, z) time rates, the following equations are used.

$$\dot{x} = V_t u' + V_r u \tag{13.76a}$$

$$\dot{y} = V_t v' + V_r v \tag{13.76b}$$

$$\dot{z} = V_t w' + V_r w \tag{13.76c}$$

13.3.2 Bounded Relative Motion Constraint

To find what conditions the orbit element differences must satisfy for the relative orbit to remain bounded, let us examine the orbit period of the Keplerian two-body problem. A bounded relative orbit is one that repeats itself after each chief orbit. The orbit period P is given in Eq. (8.68) as

$$P = 2\pi\sqrt{\frac{a^3}{\mu}} \tag{13.77}$$

If two satellite have different orbit periods, then they will drift apart and the relative orbit is considered to be unbounded. Since P only depends on the

semi-major axis a, two satellites will have the same orbit periods if

$$\delta a = 0 \tag{13.78}$$

is satisfied. The other five orbit element differences shown in Eq. (13.39) only define the relative orbit geometry, but will not cause the relative orbit to drift apart. Thus, while the Hill frame specialized bounded relative orbit constraint is $\dot{y}_0 + 2nx_0 = 0$, the orbit element difference equivalent constraint is simply $\delta a = 0$. However, while the Hill frame constraint is only valid assuming that the relative orbit size is small compared to the chief orbit radius and that the chief orbit is circular, the orbit element constraint is valid for any size relative orbit and any type of chief orbit eccentricity.

If we do assume that the relative orbit radius is small compared to the chief orbit radius, then we can use the linear mapping $\delta\boldsymbol{e} = [A(\boldsymbol{e}_c)]^{-1}\boldsymbol{X}$ to determine the general bounded relative orbit constraint for eccentric chief orbits. Using Eqs. (13.71a) - (13.71d), the bounded relative orbit constraint in Eq. (13.78) is written as

$$\begin{aligned}\delta a = 0 = 2\alpha(2 + 3\kappa_1 + 2\kappa_2)x(t) + 2\alpha\nu(1 - 2\kappa_1 + \kappa_2)y(t) \\ + \frac{2\alpha^2\nu p}{V_t}\dot{x}(t) + \frac{2a}{V_t}(1 + 2\kappa_1 + \kappa_2)\dot{y}(t)\end{aligned} \tag{13.79}$$

Note that Hill frame coordinates must satisfy this constraint at all times for the relative orbit to be bounded (i.e. all orbits have the same period). This general constraint can be further simplified by expressing it at the initial time, where t_0 is defined as the time where the true anomaly f is equal to zero and the satellite is at the orbit periapses. Note that the orbit radius is now given by

$$r(t_0) = r_p = a(1 - e) \tag{13.80}$$

Further, the radial velocity V_r is given by

$$V_r(t_0) = \dot{r}(t_0) = \frac{h}{p}(q_1 \sin\omega - q_2 \cos\omega) = 0 \tag{13.81}$$

Thus, using Eqs. (13.67) and (13.70) we find that $\nu = 0$ and $\kappa_2 = 0$. The bounded relative orbit constraint equation is now written specifically for the initial time as

$$0 = 2\frac{a}{r_p}\left(2 + 3\frac{a}{r_p}\left(\frac{p}{r_p} - 1\right)\right)x_0 + 2\frac{a}{V_t(t_0)}\left(1 + 2\frac{a}{r_p}\left(\frac{p}{r_p} - 1\right)\right)\dot{y}_0 \tag{13.82}$$

Since $V_t(t_0) = r_p\dot{\theta}_p$ and making use of Eq. (13.80), this constraint is further reduced to the simpler form

$$(2 + e)x_0 + \frac{1}{\dot{\theta}_p}(1 + e)\dot{y}_0 = 0 \tag{13.83}$$

Expressing the true latitude rate $\dot{\theta}$ at perigee as

$$\dot{\theta}_p = \frac{h}{r_p^2} = \frac{\sqrt{\mu p}}{a^2(1-e)^2} = n\sqrt{\frac{1+e}{(1-e)^3}} \tag{13.84}$$

the constraint is written in its final form as[10]

$$\frac{\dot{y}_0}{x_0} = \frac{-n(2+e)}{\sqrt{(1+e)(1-e)^3}} \tag{13.85}$$

Let us linearize this constraint about a small eccentricity. In this case terms which are linear in e are retained and higher order terms in e are dropped. Since we are already beginning with a linear mapping between orbit element difference and Cartesian Hill frame coordinates where terms of order ρ/r are dropped, this implies that $e > \rho/r$ and higher order terms of e are less than or equal to ρ/r. The bounded relative motion constraint on the initial Cartesian coordinates is then given by

$$\dot{y}_0 + (2+3e)nx_0 = 0 \tag{13.86}$$

The find the initial Cartesian coordinates constraint for bounded relative motion at the chief orbit apoapses, we set $r(t_0) = r_a = a(1+e)$ and follow the same steps. The resulting constraint for chief orbits with a general eccentricity is

$$\frac{\dot{y}_0}{x_0} = \frac{-n(2-e)}{\sqrt{(1-e)(1+e)^3}} \tag{13.87}$$

while the constraint for chief orbits with a small eccentricity is given by

$$\dot{y}_0 + (2-3e)nx_0 = 0 \tag{13.88}$$

Note that if the chief orbit is circular and $e = 0$, then this constraint reduces to the familiar form of $\dot{y}_0 + 2nx_0 = 0$ found in Eq. (13.34). The more general bounded relative orbit constraint in Eq. (13.85) is valid for eccentric chief orbits. However, its form requires that t_0 be defined to be at the orbit perigee point.

13.4 Relative Motion State Transition Matrix

The state transition matrix $[\Phi(t, t_0)]$ is defined in Eq. (11.195) as the sensitivity of the current state vector $\boldsymbol{X}(t)$ with respect to the state vector $\boldsymbol{X}(t_0)$ at the initial time.

$$[\Phi_{\boldsymbol{X}}(t, t_0)] = \left[\frac{\partial \boldsymbol{X}(t)}{\partial \boldsymbol{X}(t_0)}\right] \tag{13.89}$$

This matrix has many applications in orbital control theory and dynamical analysis of relative orbit motion. Let $\boldsymbol{X}(t)$ be the relative orbit position vector in Hill frame components of the deputy satellite relative to the chief position as

defined in Eq. (13.36). We are allowing here the chief orbit to be either circular or elliptical. However, no disturbance are considered and gravitational field is idealized as that of a spherical Earth. For the nonlinear dynamical system describing the relative motion of the satellites, the position vector at time t can be approximated using the state transition matrix $[\Phi_{\boldsymbol{X}}(t,t_0)]$ through

$$\boldsymbol{X}(t) \approx [\Phi_{\boldsymbol{X}}(t,t_0)]\boldsymbol{X}(t_0) \tag{13.90}$$

If the dynamical relative equations of motion were linear, then this would be a precise mapping between initial and current state vectors as shown in Eq. (11.172).

One brute force method to generate this state transition matrix $[\Phi_{\boldsymbol{X}}(t,t_0)]$ would be to solve the relative equations of motion analytically for $\boldsymbol{X}(t)$ and the take the required partial derivatives. For a circular chief orbit, the linearized equations of motion have been solved using the CW equations. The result is shown in Eq. (13.33). For general elliptical orbits, finding the analytical solution of the relative equations of motion is substantially more complicated.[11]

We will pursue a simpler solution making direct use of the linear mapping between the Hill frame relative Cartesian position vector and the corresponding orbit element differences shown in Eq. (13.40). In particular, at times t and t_0 we write express the relative orbit position vector as

$$\boldsymbol{X}(t) = [A(\boldsymbol{e}(t))]\delta\boldsymbol{e}(t) \tag{13.91}$$

$$\boldsymbol{X}(t_0) = [A(\boldsymbol{e}(t_0))]\delta\boldsymbol{e}(t_0) \tag{13.92}$$

Note that unless stated otherwise, all inertial orbit elements are assumed to be chief orbit elements. The state transition matrix of the orbit element difference vector $\delta\boldsymbol{e}(t)$ is defined analogously to $[\Phi_{\boldsymbol{X}}(t,t_0)]$ as

$$[\Phi_{\delta\boldsymbol{e}}(t,t_0)] = \left[\frac{\partial(\delta\boldsymbol{e}(t))}{\partial(\delta\boldsymbol{e}(t_0)}\right] \tag{13.93}$$

Let the orbit element difference vector be defined using Eq. (13.41) as

$$\delta\boldsymbol{e} = (\delta a, \delta\theta, \delta i, \delta q_1, \delta q_2, \delta\Omega)^T \tag{13.94}$$

where $\theta = \omega + f$ is the true latitude and q_1 and q_2 are defined in Eqs. (13.42) and (13.43). For the Keplerian motion assumed in this section, note that all these orbit element difference will remain *constant* except for the true anomaly difference $\delta\theta$. It does evolve nonlinearly in time for general elliptical chief orbits. However, for the remaining orbit element differences, the sensitivity at time t with respect to this state at time t_0 will simply be 1. This will provide a $[\Phi_{\delta\boldsymbol{e}}(t,t_0)]$ matrix which has a very simple structure. If perturbations are included, then the computation of $[\Phi_{\delta\boldsymbol{e}}(t,t_0)]$ becomes more involved. See References 6 and 12 for a detailed study of including J_2 perturbations and mean orbit elements into the state transition matrix calculations. Given the state

transition matrix $[\Phi_{\delta \boldsymbol{e}}(t, t_0)]$, the orbit element difference vector at time t is be approximated as

$$\delta \boldsymbol{e}(t) \approx [\Phi_{\delta \boldsymbol{e}}(t, t_0)] \delta \boldsymbol{e}(t_0) \tag{13.95}$$

To compute the desired state transition matrix $[\Phi_{\boldsymbol{X}}(t, t_0)]$ in terms of $[A(\boldsymbol{e})]$ and $[\Phi_{\delta \boldsymbol{e}}(t, t_0)]$, we substitute Eqs. (13.91) and (13.92) to find

$$[A(t)] \delta \boldsymbol{e}(t) = [\Phi_{\boldsymbol{X}}(t, t_0)][A(t_0)] \delta \boldsymbol{e}(t_0) \tag{13.96}$$

Substituting Eq. (13.95) into the above equation and solving for $[\Phi_{\boldsymbol{X}}(t, t_0)]$ yields

$$[\Phi_{\boldsymbol{X}}(t, t_0)] = [A(t)][\Phi_{\delta \boldsymbol{e}}(t, t_0)][A(t_0)]^{-1} \tag{13.97}$$

The non-zero components of both $[A(t)]$ and $[A(t_0)]$ were developed earlier. Note that this $[\Phi_{\boldsymbol{X}}(t, t_0)]$ computation is valid for both circular and non-circular chief orbits. The assumption made here is that the orbit perturbations about the Keplerian motion are negligible.

The only non-trivial term of the $[\Phi_{\delta \boldsymbol{e}}(t, t_0)]$ matrix computation is the true latitude difference

$$\delta \theta = \delta \omega + \delta f \tag{13.98}$$

Since the argument of perigee difference $\delta \omega$ will not change with time, we can state that

$$\delta \theta(t) = \delta \omega + \delta f(t) \tag{13.99}$$

$$\delta \theta(t_0) = \delta \omega + \delta f(t_0) \tag{13.100}$$

and thus focus our treatment on the search of $\delta f(t)$ as a function of $\delta \boldsymbol{e}(t_0)$. Kepler's equation is given by

$$M(t) = M_0 + \sqrt{\frac{\mu}{a^3}}(t - t_0) \tag{13.101}$$

For notational convenience, a subscript "0" will indicated that a state is taken at time t_0. No subscript means the state is taken at time t. Taking the first variation of Kepler's equation, we are able to relate differences in mean anomaly at times t and t_0 through

$$\delta M = \delta M_0 - \frac{3}{2} \frac{\delta a}{a} (M - M_0) \tag{13.102}$$

To express the mean anomaly differences in terms of other anomaly differences, we make use of the mean anomaly definition

$$M = E - e \sin E \tag{13.103}$$

and take its first variation to yield

$$\begin{aligned}\delta M &= \frac{\partial M}{\partial E}\delta E + \frac{\partial M}{\partial e}\delta e \\ &= (1 - e\cos E)\delta E - \sin E \delta e\end{aligned} \tag{13.104}$$

Using the mapping between eccentric anomaly E and true anomaly f

$$\tan\frac{f}{2} = \sqrt{\frac{1+e}{1-e}}\tan\frac{E}{2} \tag{13.105}$$

and taken its first variation, differences in E are then expressed as differences in f and e through

$$\delta E = \frac{\eta}{1+e\cos f}\delta f - \frac{\sin f}{1+e\cos f}\frac{\delta e}{\eta} \tag{13.106}$$

with $\eta = \sqrt{1-e^2}$. Substituting Eq. (13.106) into Eq. (13.104) and solving for δM using the orbit identities in Appendix E yields

$$\delta M = \frac{\eta}{(1+e\cos f)^2}\left(\eta^2\delta f - \sin f(2+e\cos f)\delta e\right) \tag{13.107}$$

Analogously, the initial mean anomaly difference δM_0 is expressed by taking advantage of the fact that only the anomalies will differ in time through

$$\delta M_0 = \frac{\eta}{(1+e\cos f_0)^2}\left(\eta^2\delta f_0 - \sin f_0(2+e\cos f_0)\delta e\right) \tag{13.108}$$

Substituting Eqs. (13.107) and (13.108) into Eq. (13.102) and solving for δf yields[9]

$$\begin{aligned}\delta f = &\underbrace{-\frac{3}{2}\frac{a\eta}{r^2}(M - M_0)}_{A}\delta a + \underbrace{\left(\frac{r}{r_0}\right)^2}_{B}\delta f_0 \\ &+ \underbrace{\left(\sin f(2+e\cos f) - \sin f_0(2+e\cos f_0)\left(\frac{r}{r_0}\right)^2\right)\frac{1}{\eta^2}}_{C}\delta e\end{aligned} \tag{13.109}$$

Let us write the true anomaly differences in the following compact form

$$\delta f = A\delta a + B\delta f_0 + C\delta e \tag{13.110}$$

In terms of the orbit elements used in Eq. (13.41), the scalar parameters are

defined as

$$A = -\frac{3}{2}\frac{a\eta}{r^2}(M - M_0) \tag{13.111}$$

$$B = \left(\frac{r}{r_0}\right)^2 \tag{13.112}$$

$$\begin{aligned} C = \frac{1}{\eta^2\sqrt{q_1^2+q_2^2}}\Big((\sin\theta q_1 - \cos\theta q_2)(2 + q_1\cos\theta + q_2\sin\theta) \\ - (\sin\theta_0 q_1 - \cos\theta_0 q_2)(2 + q_1\cos\theta_0 + q_2\sin\theta_0)\left(\frac{r}{r_0}\right)^2\Big) \end{aligned} \tag{13.113}$$

with the orbit radius r being defined in Eq. (13.55) and $\eta = \sqrt{1-q_1^2-q_2^2}$. Using the definitions of q_1 and q_2 in Eq. (13.42) and (13.43), the differences in eccentricity and argument of perigee as expressed as

$$\delta e = \frac{1}{\sqrt{q_1^2+q_2^2}}\left(q_1\delta q_1 + q_2\delta q_2\right) \tag{13.114}$$

$$\delta\omega = \frac{1}{q_1^2+q_2^2}\left(q_1\delta q_2 - q_2\delta q_1\right) \tag{13.115}$$

Substituting Eqs. (13.99), (13.100), (13.114) and (13.115) into Eq. (13.110), we are able to express true latitude differences at time t in terms of initial orbit element differences through

$$\begin{aligned} \delta\theta = A\delta a + B\delta\theta_0 + \underbrace{\left(C\frac{q_1}{\sqrt{q_1^2+q_2^2}} - (1-B)\frac{q_2}{q_1^2+q_2^2}\right)}_{C_1}\delta q_1 \\ + \underbrace{\left(C\frac{q_2}{\sqrt{q_1^2+q_2^2}} + (1-B)\frac{q_1}{q_1^2+q_2^2}\right)}_{C_2}\delta q_2 \end{aligned} \tag{13.116}$$

Since all orbit element differences except for the anomalies and latitude angles are the same at any time, we are now able to write the orbit element difference state transition matrix as

$$[\Phi_{\delta\boldsymbol{e}}(t,t_0)] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ A & B & 0 & C_1 & C_2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{13.117}$$

Using this matrix in Eq. (13.97), we are able to directly compute the state transition matrix $[\Phi_{\boldsymbol{X}}(t,t_0)]$ of the rotating Hill frame relative orbit position vector $\boldsymbol{X}$ at any time. The presented method could also be used to develop the state transition matrix using other orbit elements in an analogous manner.

13.5 Linearized Relative Orbit Motion

Note that Eq. 13.59 provides us a direct linear mapping between orbit element differences $\delta\boldsymbol{e}$ and the Hill frame Cartesian coordinate vector $\boldsymbol{\rho}$. The only linearizing assumption that was made is that the relative orbit radius ρ is small compared to the inertial chief orbit radius r. Since this mapping must hold at any instance of time, however, these linearized equations also approximate a solution for the relative orbit motion $\boldsymbol{\rho}$ in terms of the true anomaly angle f. To map between time and the true anomaly we must solve Kepler's equation. However, to be able to describe the relative orbit geometry in terms of the Hill frame Cartesian coordinates, the solution in terms of the true anomaly f is preferred. The reason for this is that by sweeping f through a complete revolution, the (x, y, z) coordinates found through these equations will yield the linearized relative orbit approximation that results due to a prescribed set of constant orbit element differences. Note that no differential equations are solved here to determine the relative orbit motion, and that the dominant relative orbit radial (x-direction), along-track (y-direction) and out-of-plane motion (z-direction) can be trivially extracted.

13.5.1 General Elliptic Orbits

However, when describing a relative orbit through orbit element differences, it is not convenient to describe the anomaly difference through $\delta\theta$ or δf. For elliptic chief orbits, the difference in true anomaly between two orbits will vary with throughout the orbit. To avoid this issue, the desired anomaly difference between two orbits is typically expressed in terms of a mean anomaly difference δM. This anomaly difference will remain constant, assuming unperturbed Keplerian motion, even if the chief orbit is elliptic. Using Eq. (13.107), differences in true anomaly are written in terms of differences in mean anomaly and differences in eccentricity as

$$\delta f = \frac{(1+e\cos f)^2}{\eta^3}\delta M + \frac{\sin f}{\eta^2}(2+e\cos f)\delta e \tag{13.118}$$

Let us define the orbit element difference vector $\delta\boldsymbol{e}$ to consist of

$$\delta\boldsymbol{e} = (a, \delta M, \delta i, \delta\omega, \delta e, \delta\Omega)^T \tag{13.119}$$

Note that all these orbit element differences are constants for Keplerian two-body motion. Further, while using q_1 and q_2 instead of e and ω allows us to avoid singularity issues for near-circular orbits, for the following relative orbit geometry discussion such singularities do not appear. In fact, describing the relative orbit path using δe and $\delta\omega$ instead of δq_1 and δq_2 yields a simpler and more elegant result. Using Eqs. (13.42) and (13.43), the differences in the q_i parameters are expressed as

$$\delta q_1 = \cos\omega\delta e - e\sin\omega\delta\omega \tag{13.120a}$$

$$\delta q_2 = \sin\omega\delta e + e\cos\omega\delta\omega \tag{13.120b}$$

After substituting Eqs. (13.118) and (13.120) into the linear mapping in Eq. (13.59) and simplifying the result, we are able to express the relative position coordinates (x, y, z) in terms of the orbit element differences in Eq. (13.119) through

$$x(f) \approx \frac{r}{a}\delta a + \frac{ae\sin f}{\eta}\delta M - a\cos f\delta e \tag{13.121a}$$

$$y(f) \approx \frac{r}{\eta^3}(1+e\cos f)^2\delta M + r\delta\omega + \frac{r\sin f}{\eta^2}(2+e\cos f)\delta e + r\cos i\delta\Omega \tag{13.121b}$$

$$z(f) \approx r(\sin\theta\delta i - \cos\theta\sin i\delta\Omega) \tag{13.121c}$$

Note that with this linearized mapping the difference in the argument of perigee $\delta\omega$ does not appear in the $x(f)$ expression. Further, these equations are valid for both circular and elliptic chief orbits. Only the δM and δe terms contribute periodic terms to the radial x solution. Due to the dependence of r on the true anomaly f, all orbit element difference terms in the along-track y motion contribute both static offsets as well as periodic terms. For the out-of-plane z motion both the δi and $\delta\Omega$ terms control the out-of-plane oscillations.

However, note that Eq. (13.121) does not explicitly contain any secular terms as is the case with the general solution to the CW equations in Eq. (13.33). For the classical two-body orbital motion, the only condition on two inertial orbits to have a closed relative orbit is that their orbit energies must be equal and thus $\delta a = 0$. This constraint is valid for both circular and elliptical chief orbits. Also, note that this constraint is the precise requirement of the Keplerian motion for bounded relative orbit paths; no linearizations have been made here. For Keplerian two-body motion, all the orbit element differences will naturally remain constant except for the mean anomaly difference. If δa is not zero between two orbits, then these orbits will drift apart due to having different orbit periods. In this case δM will not remain a constant but grow larger with time. The linearization in Eq. (13.121) can still be used to predict the relative orbit motion, but only until the relative orbit radius ρ is no longer small compared to the inertial chief orbit radius r. If perturbations such as the J_2 gravitational perturbations are included, then the appropriate orbit element differences must be treated as slowly time varying. Thus the potential secular drift of a relative orbit is hidden within the behavior of the orbit element differences themselves.

By dividing the dimensional (x, y, z) expressions in Eq. (13.121) by the chief orbit radius r and making use of Eq. (13.55), we obtain the non-dimensional relative orbit coordinates (u, v, w).

$$u(f) \approx \frac{\delta a}{a} + (1+e\cos f)\frac{e\sin f}{\eta^3}\delta M - \frac{(1+e\cos f)}{\eta^2}\cos f\delta e \tag{13.122a}$$

$$v(f) \approx (1+e\cos f)^2\frac{\delta M}{\eta^3} + \delta\omega + \frac{\sin f}{\eta^2}(2+e\cos f)\delta e + \cos i\delta\Omega \tag{13.122b}$$

$$w(f) \approx \sin\theta\delta i - \cos\theta\sin i\delta\Omega \tag{13.122c}$$

Since $(y, z) << r$, the non-dimensional coordinates (v, w) are the angular deputy satellite relative orbit position with respect to the chief orbit radius axis.

However, the present form of Eq. (13.122) is not convenient to determine the overall non-dimensional shape of the relative orbit. Reason is that there are several sin() and cos() functions being added here. Using the identities

$$\begin{aligned} A\sin t + B\cos t &= \sqrt{A^2+B^2}\cos\left(t - \tan^{-1}\left(\frac{A}{B}\right)\right) \\ &= -\sqrt{A^2+B^2}\sin\left(t - \tan^{-1}\left(\frac{B}{-A}\right)\right) \end{aligned} \tag{13.123}$$

as well as standard trigonometric identities, we are able to rewrite the linearized non-dimensional relative orbit motion as

$$\begin{aligned} u(f) \approx{}& \frac{\delta a}{a} + \frac{1}{\eta^2}\sqrt{\frac{e^2\delta M^2}{\eta^2} + \delta e^2}\cos(f - f_u) \\ &- \frac{e\delta e}{2\eta^2} + \frac{e}{2\eta^2}\sqrt{\frac{e^2\delta M^2}{\eta^2} + \delta e^2}\cos(2f - f_u) \end{aligned} \tag{13.124a}$$

$$\begin{aligned} v(f) \approx{}& \left(\left(1 + \frac{e^2}{2}\right)\frac{\delta M}{\eta^3} + \delta\omega + \cos i\delta\Omega\right) \\ &+ \frac{2}{\eta^2}\sqrt{\frac{e^2\delta M^2}{\eta^2} + \delta e^2}\cos(f - f_v) \\ &+ \frac{e}{2\eta^2}\sqrt{\frac{e^2\delta M^2}{\eta^2} + \delta e^2}\cos(2f - f_v) \end{aligned} \tag{13.124b}$$

$$w(f) \approx \sqrt{\delta i^2 + \sin^2 i\delta\Omega^2}\cos\left(\theta - \theta_w\right) \tag{13.124c}$$

with the phase angles f_u, f_v and θ_w being defined as

$$f_u = \tan^{-1}\left(\frac{e\delta M}{-\eta\delta e}\right) \tag{13.125a}$$

$$f_v = \tan^{-1}\left(\frac{\eta\delta e}{e\delta M}\right) = f_u - \frac{\pi}{2} \tag{13.125b}$$

$$\theta_w = \tan^{-1}\left(\frac{\delta i}{-\sin i\ \delta\Omega}\right) \tag{13.125c}$$

At these phase angles, the trigonometric terms will reach either their minimum or maximum value. Note that 180 degrees can be added or subtracted from these angles to yield the second extrema point of the trigonometric functions. To further reduce the expression in Eq. (13.124), let us introduce the small states δ_u and δ_w:

$$\delta_u = \sqrt{\frac{e^2\delta M^2}{\eta^2} + \delta e^2} \tag{13.126a}$$

$$\delta_w = \sqrt{\delta i^2 + \sin^2 i\delta\Omega^2} \tag{13.126b}$$

Using these δ_u and δ_w definitions as well as Eq. (13.125b), the linearized relative orbit motion is described through

$$u(f) \approx \frac{r}{a}\delta a - \frac{e\delta e}{2\eta^2} + \frac{\delta_u}{\eta^2}\left(\cos(f-f_u) + \frac{e}{2}\cos(2f-f_u)\right) \tag{13.127a}$$

$$\begin{aligned} v(f) \approx & \left(\left(1+\frac{e^2}{2}\right)\frac{\delta M}{\eta^3} + \delta\omega + \cos i\delta\Omega\right) \\ & - \frac{\delta_u}{\eta^2}\left(2\sin(f-f_u) + \frac{e}{2}\sin(2f-f_u)\right) \end{aligned} \tag{13.127b}$$

$$w(f) \approx \delta_w \cos(\theta - \theta_w) \tag{13.127c}$$

Note that the $\cos(2f)$ and $\sin(2f)$ terms are multiplied by the eccentricity e. Only if the chief orbit is very eccentric will these terms have a significant contribution to the overall relative orbit dimension. For the more typical case of having a chief orbit with a small eccentricity e, these terms only provide small perturbations to the dominant $\sin(f)$ and $\cos(f)$ terms. Using Eq. (13.127), it is trivial to determine the maximum radial, along-track and out-of-plane dimension of a relative orbit provided that the relative orbit geometry is prescribed through the set of orbit element differences $\{\delta a, \delta e, \delta i, \delta\Omega, \delta\omega, \delta M\}$. Note that this linearized relative orbit motion is valid for both circular and elliptic chief reference orbits. The only linearizing assumption made so far is that the relative orbit radius is small compared to the planet centric inertial orbit radius. However, note that we are only estimating the non-dimensional relative orbit shape. To obtain the true radial, along-track and out-of-plane motions, we need to multiply (u, v, w) by the chief orbit radius r. Since r is time dependent for an elliptic chief orbit, the points of maximum angular separation between deputy and chief satellites may not correspond to the point of maximum physical distance. To plot the dimensional linearized relative orbit motion, we use Eq. (13.121) instead. However, due to the ratio's of sin() and cos() terms, it is not trivial to obtain the maximum physical dimensions of the relative orbit.

Let us take a closer look at the out-of-plane motion. The true latitude angle θ_w, at which the maximum angular out-of-plane motion will occur, is given by Eq. (13.125c). As expected, if only a $\delta\Omega$ is prescribed, then the maximum w motion occurs during the equator crossing at $\theta = 0$ or 180 degrees. If only a δi is prescribed, then the maximum w motion occurs at $\theta = \pm 90$ degrees.

The maximum angular out-of-plane motion is given by the angle δ_w as shown in Figure 13.5. This angle δ_w is the tilt angle of the deputy orbit plane relative to the chief orbit plane. As such, it is the angle between the angular momentum vector of the chief orbit and the angular momentum vector of the deputy orbit. To prove that δ_w is indeed this angle, let us make use of the spherical law of cosines for angles. Using the spherical trigonometric law of cosines, we are able to relate the angles $\delta\Omega$, i, δi and δ_w through:[13]

$$\cos\delta_w = \cos i\cos(i+\delta i) + \sin i\sin(i+\delta i)\cos\delta\Omega \tag{13.128}$$

Assuming that $\delta\Omega$, δi and δ_w are small angles, we approximate $\sin x \approx x$ and

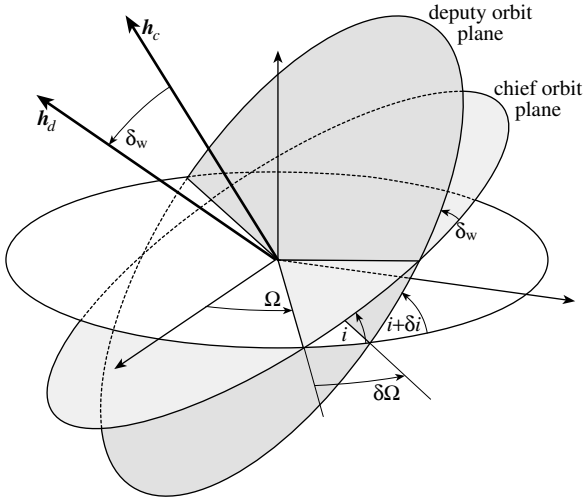


Figure 13.5: Illustration of Orbit Plane Orientation Difference between Chief and Deputy Satellites

$\cos x \approx 1 - x^2/2$ to solve for δ_w.

$$\delta_w = \sqrt{\delta i^2 + \sin^2 i \delta\Omega^2} \tag{13.129}$$

Using the angle δ_w, the out-of-plane motion u in Eq. (13.127c) is written in the compact form shown.[14, 15]

13.5.2 Chief Orbits with Small Eccentricity

In this section we assume that the chief orbit eccentricity e is a small quantity. In particular, we assume that e is small but greater than ρ/r, while powers of e are smaller than ρ/r. In this case we only retain terms which are linear in e and drop higher order terms of e. The orbit radius r is now approximated as

$$r = \frac{a\eta^2}{1 + e\cos f} \approx a(1 - e\cos f) \tag{13.130}$$

while $\eta^2 \approx 1$. The linearized dimensional relative orbit motion in Eq. (13.121) is written for the small eccentricity case as:

$$x(f) \approx (1 - e\cos f)\delta a + \frac{ae\sin f}{\eta}\delta M - a\cos f\delta e \tag{13.131a}$$

$$\begin{aligned} y(f) \approx{}& \frac{a}{\eta}(1 + e\cos f)\delta M + a(1 - e\cos f)\delta\omega \\ &+ a\sin f(2 - e\cos f)\delta e + a(1 - e\cos f)\cos i\delta\Omega \end{aligned} \tag{13.131b}$$

$$z(f) \approx a(1 - e\cos f)(\sin\theta\delta i - \cos\theta\sin i\delta\Omega) \tag{13.131c}$$

Making use of the trigonometric identity in Eq. (13.123), the (x, y, z) motion is written as

$$x(f) \approx \delta a + a\delta_x \cos(f - f_x) \tag{13.132a}$$

$$y(f) \approx a\left(\frac{\delta M}{\eta} + \delta\omega + \cos i\delta\Omega\right) - a\delta_y \sin(f - f_y) - \frac{ae}{2}\sin(2f)\delta e \tag{13.132b}$$

$$\begin{aligned} z(f) \approx{}& a\delta_z \cos(\theta - \theta_z) - \frac{ae}{2}\delta_z \cos(2f - f_z) \\ &- \frac{ae}{2}\left(\sin\omega\delta i - \cos\omega \sin i\delta\Omega\right) \end{aligned} \tag{13.132c}$$

with the small states δ_x, δ_y and δ_z defined as

$$\delta_x = \sqrt{\frac{e^2\delta M^2}{\eta^2} + \left(\delta e + \frac{\delta a}{a}\right)^2} \tag{13.133a}$$

$$\delta_y = \sqrt{4\delta e^2 + e^2\left(\frac{\delta M}{\eta} - \delta\omega - \cos i\delta\Omega\right)^2} \tag{13.133b}$$

$$\delta_z = \sqrt{\delta i^2 + \sin^2 i\delta\Omega^2} \tag{13.133c}$$

and the phase angles f_x, f_y, θ_z and f_z defined as

$$f_x = \tan^{-1}\left(\frac{e\delta M}{-\eta\left(\delta e + \frac{\delta a}{a}\right)}\right) \tag{13.134a}$$

$$f_y = \tan^{-1}\left(\frac{e\left(\frac{\delta M}{\eta} - \delta\omega - \cos i\delta\Omega\right)}{-2\delta e}\right) \tag{13.134b}$$

$$\theta_z = \tan^{-1}\left(\frac{\delta i}{-\sin i\delta\Omega}\right) \tag{13.134c}$$

$$f_z = \tan^{-1}\left(\frac{\cos\omega\delta i + \sin\omega\sin i\delta\Omega}{\sin\omega\delta i - \cos\omega\sin i\delta\Omega}\right) \tag{13.134d}$$

Note that the orbital radial motion $x(f)$ for the small eccentricity case is identical to the general orbit radial coordinate in Eq. (13.121a) if δa is zero. The semi-major axis different must be zero for bounded relative motion if no perturbations are present. With perturbations present, δa may be non-zero and the orbit radial coordinate will then be different between the linearizing approximations. The estimated along-track motion $y(f)$ and out-of-plane motion $z(f)$ will always be numerically different between the generally elliptic case and the small eccentricity case.

The dimensional form of the relative orbit motion in Eq. (13.131) is convenient to determine the amplitudes of the sinusoidal motion in either the along-track, orbit radial or out-of-plane motion. Note that since e is considered small, the double-orbit frequency terms $\sin(2f)$ are only a minor perturbation to the dominant orbit frequency sinusoidal terms.

13.5.3 Near-Circular Chief Orbit

If the chief orbit is circular or near-circular, then the linearized relative equations of motion are given through the famous Clohessy-Wiltshire or CW equations.[4] Assuming the bounded relative motion constraint $\dot{y}_0 + 2nx_0 = 0$ is satisfied, then the differential CW equations can now be solved for the analytical solution of the relative orbit motion shown in Eq. (13.33).

$$x(t) = A_0 \cos(nt + \alpha) \tag{13.135a}$$
$$y(t) = -2A_0 \sin(nt + \alpha) + y_{off} \tag{13.135b}$$
$$z(t) = B_0 \cos(nt + \beta) \tag{13.135c}$$

The integration constants A_0, B_0, α, β and y_{off} are determined through the relative orbit initial conditions. These equations have been extensively used to generate relative orbits if the chief orbit is circular. Let us now compare the predicted (x, y, z) motion in terms of the true anomaly in Eq. (13.132) to the CW solution in Eq. (13.33) if the chief orbit is assumed to be near-circular (i.e. $e < \rho/r$). In this case terms containing the eccentricity e are dropped, as compared to the small eccentricity case studied earlier where only higher order terms of e were dropped. Assuming that all $\delta\boldsymbol{e}$ components are small (i.e. the relative orbit radius is assumed to be small compared to the inertial orbit radius), and letting $e \to 0$, we find that $r \to a$ and $\eta \to 1$. Further, note that f_x and f_y approach 0. Using Eq. (13.132) the relative orbit motion $(x(f), y(f), z(f))$ is expressed for the near-circular chief orbit special case as

$$x(f) \approx \delta a - a \cos f \delta e \tag{13.136a}$$
$$y(f) \approx a(\delta\omega + \delta M + \cos i \delta\Omega) + 2a \sin f \delta e \tag{13.136b}$$
$$z(f) \approx a\sqrt{\delta i^2 + \sin^2 i\delta\Omega^2} \cos(\theta - \theta_z) \tag{13.136c}$$

Note that the maximum width of the oscillatory along-track motion y is given by $2a\delta e$. This result has been previously presented in References 15 and 16. Comparing Eqs. (13.33) and (13.136) and noting that $nt = f$ for this case, we are able to establish a direct relationship between the CW constants and the orbit element differences.

$$A_0 = -a\delta e \tag{13.137a}$$
$$B_0 = a\sqrt{\delta i^2 + \sin^2 i\delta\Omega^2} \tag{13.137b}$$
$$\alpha = 0 \tag{13.137c}$$
$$\beta = \omega - \theta_z \tag{13.137d}$$
$$y_{off} = a(\delta\omega + \delta M + \cos i\delta\Omega) \tag{13.137e}$$

Recall that Eqs. (13.33) require that the bounded relative motion constraint is satisfied. Thus the δa term is set to zero when comparing the two forms of the relation orbit motion expression.

Example 13.2: The following numerical simulations verify that the relative motion approximation in Eqs. (13.121), (13.132) and (13.136) do indeed predict the spacecraft formation geometry. These simulations also illustrate the accuracy at which these simplified linearized solutions are valid. Let the chief orbit be given by the orbit elements shown in Table 13.1.

Table 13.1: Chief Orbit Elements

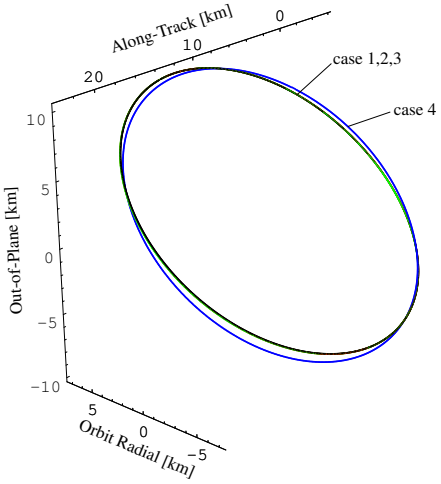
Orbit Elements	Value	Units
a	7555	km
e	0.03 or 0.13	
i	48.0	deg
Ω	20.0	deg
ω	10.0	deg
M_0	0.0	deg

The relative orbits are studied for two different chief eccentricities. For the relative orbits studied, the ratio ρ/r is about 0.003. The smaller of the two eccentricities considered is already an order of magnitude larger than this, while the second eccentricities is even larger again. The numerical simulations show that the small eccentricity assumption (i.e. retaining terms in e but dropping higher order terms in e) will still yield a reasonable relative orbit prediction for $e = 0.03$, even though it is larger than the small term ρ/r. The orbit element differences which define the relative orbit are given in Table 13.2. Since these simulations assume a two-body Keplerian motion of the satellites, the semi-major axis difference δa must be zero to achieve a bounded relative motion.

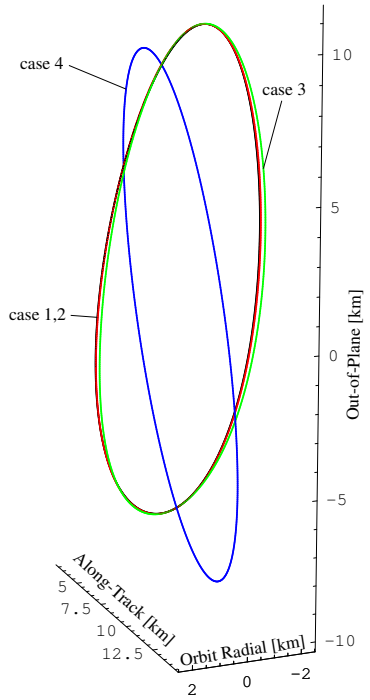
Table 13.2: Orbit Element Differences Defining the Spacecraft Formation Geometry

Orbit Elements	Value	Units
δa	0	km
δe	0.00095316	
δi	0.0060	deg
$\delta\Omega$	0.100	deg
$\delta\omega$	0.100	deg
δM_0	-0.100	deg

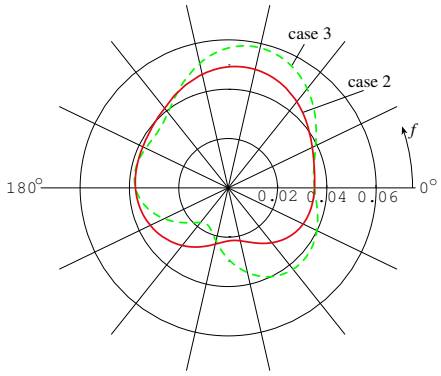
The following figures compare the relative orbit motion for four different cases. Case 1 is the relative motion that will result using the true nonlinear equations of motion. Case 2 uses the dimensional linearized analytical relative orbit solution in Eq. (13.121). The only assumption that has been made here is that ratio between the relative orbit radius ρ and the inertial chief orbit radius r is small and terms involving ρ/r have been dropped. Case 3 assumes that the chief orbit eccentricity is small, but not near zero. As such, higher order terms in e are dropped, while terms which depend linearly on e are



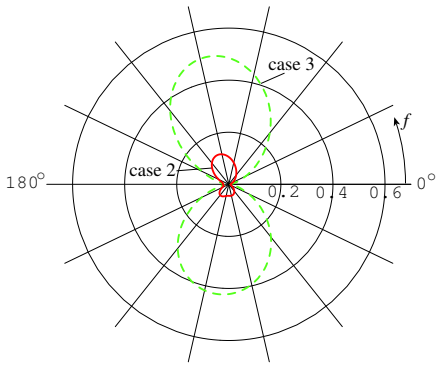
(i) Relative Orbits in Hill Frame for $e = 0.03$



(ii) Relative Orbits in Hill Frame for $e = 0.13$



(iii) RMS Relative Orbit Error in kilometers vs. Chief True Anomaly Angle f for $e = 0.03$



(iv) RMS Relative Orbit Error in kilometers vs. Chief True Anomaly Angle f for $e = 0.13$

Figure 13.6: Comparison of the Linearized Relative Orbit Solutions for Cases 1–4 with $e = 0.03$ and $e = 0.13$.

kept. The relative orbit motion is described through Eq. (13.132). Case 4 assumes that the chief orbit is near-circular and that e is very close to zero. Any terms involving the eccentricity e are dropped here to yield the classical CW equations in Eq. (13.136). Case 4 is not included here to suggest that a circular orbit assumption should be made when the chief orbit is clearly eccentric. The circular chief orbit assumption case is included to provide a relative comparison illustrating the extent of the eccentricity effect.

The resulting relative orbit motion is illustrated in Figure 13.6. Figures 13.6(i) and 13.6(ii) show the three-dimensional relative orbits for cases 1 through 4 as seen by the rotating Hill reference frame. The relative orbit radii vary between 10 and 20 kilometers. When $e = 0.03$, note that the relative orbits for cases 1–3 are virtually indistinguishable. Only the relative orbit prediction assuming a circular orbit (case 4) has a clearly visible distinct motion. Studying Figure 13.6(ii) with $e = 0.13$, the case 2 relative orbit is still indistinguishable on this scale from the true relative motion in case 1. With this larger eccentricity the relative motion predicted in case 3 (dropping higher order terms in e) does show some visible departure from the true relative motion. As expected, the circular chief orbit assumption (case 4) yields a very poor prediction of the relative orbit motion.

In Figures 13.6(iii) and 13.6(iv) the RMS relative orbit errors are shown in polar plots versus the chief orbit true anomaly. For the $e = 0.03$ simulations, the relative orbit errors for case 2 lie between 20 and 40 meters. Since the relative orbit radius is roughly 10 kilometers, this corresponds to a 0.2–0.4 percent relative motion error. The RMS relative motion error for case three is only marginally worse. As was discussed earlier, dropping the higher order e terms should begin to have a noticeably affect on the relative motion errors. For the $e = 0.13$ simulations, the relative motion errors for case 2 lie between 50 and 100 meters (roughly 0.5–1.0 percent errors). However, dropping the higher order e terms in case 3 has a very noticeable effect with the relative motion errors growing as large as 500 meters (about 5.0 percent error).

13.6 J_2-Invariant Relative Orbits

To motivate the discussion in this chapter, let us revisit one particular class of spacecraft formations where the satellite constellation is composed to form a rotating sparse aperture. These types of formations are typically considered in remote sensing missions where each satellite is an individual element of a large, virtual antenna formed by the formation. By sharing the individual measurements, the resolution of the spacecraft cluster is potentially much higher then the resolution of any individual craft. To minimize secular relative drift among the spacecraft, these missions typically are comprised of identical spacecraft to reduce the differential atmospheric drag. The spacecraft formation orbit will decay due atmospheric drag. However, all satellites orbits will decay are nominally the same rate. Thus, the atmospheric drag has only a secondary effect on the *relative orbit geometry*. The gravitational perturbations are typically the dominant factor producing the secular drift in this case. Ignoring these pertur-

bations leads to relative orbit designs which require more frequent corrections, and thus use more fuel.

Adding the J_2 perturbation to the classical Keplerian orbit motion causes three types of changes in the osculating orbit elements, short period and long period oscillations, and secular growth. The long period term is the period of the apsidal rotation. Over a short time this looks like a secular growth of $\mathcal{O}(J_2^2)$. The short period growth manifests itself as oscillations of the orbit elements, but doesn't cause the orbits to drift apart. The relative secular growth is the motion that needs to be avoided for relative orbits to be J_2 invariant. This growth is best described through *mean orbit elements* rather than the osculating elements. These orbit elements have the short and long period oscillations removed. A direct mapping between the osculating (instantaneous) orbit element and the mean (orbit averaged) orbit elements is provided by Brouwer in Reference 17. A popular modification to make this algorithm more robust near zero eccentricities and orbit inclinations was introduced by Lyddane in Reference 18. A numerical algorithm providing a first order approximation to this osculating to mean element mapping is provided in Appendix G. By studying the relative motion through the use of mean orbit elements,[19–21] we are able to ignore the orbit period specific oscillations and address the secular drift directly. It is not possible to set all of the individual orbit drifts equal to zero. However, instead we choose to set the *relative* mean orbit element drifts to zero to avoid *relative secular growth*.

The J_2 perturbations cause secular drift in the mean longitude of the ascending node $\overline{\Omega}$, the mean argument of perigee $\overline{\omega}$ and the mean anomaly $\overline{M}$. As shown in Eq. (11.87), the non-zero mean orbit element rates due to the J_2 gravitational perturbation are given by:

$$\frac{d\overline{\Omega}}{dt} = -\frac{3}{2} J_2 n \left(\frac{r_{eq}}{p}\right)^2 \cos i \tag{13.138a}$$

$$\frac{d\overline{\omega}}{dt} = \frac{3}{4} J_2 n \left(\frac{r_{eq}}{p}\right)^2 (5\cos^2 i - 1) \tag{13.138b}$$

$$\frac{d\overline{M}}{dt} = n + \frac{3}{4} J_2 n \left(\frac{r_{eq}}{p}\right)^2 \sqrt{1-e^2}(3\cos^2 i - 1) \tag{13.138c}$$

The magnitude of the secular drifts are determined by the semi-major axis a, eccentricity e and the inclination angle i.[22] If these quantities aren't carefully selected, then the relative drift rates will cause secular drift among the various spacecraft in the formation.[19]

13.6.1 Ideal Constraints

For Keplerian motion (i.e. no gravitational perturbations present), only the mean anomaly M is a time dependent orbit element. The rate at which M grows is given by the mean orbit rate n. For the relative motion between to

satellites to be bounded, the mean anomaly rates must be equal.

$$\delta\dot{M} = \dot{M}_d - \dot{M}_c = n_d - n_c = \sqrt{\frac{\mu}{a_d^3}} - \sqrt{\frac{\mu}{a_c^3}} = 0 \tag{13.139}$$

Since the mean anomaly rate only depends on a, the difference in mean anomaly rates $\delta\dot{M}$ can also be approximated to first order as

$$\delta\dot{M} \approx \frac{\partial \dot{M}_c}{\partial a_c}\delta a = -\frac{3}{2}\sqrt{\frac{\mu}{a_c^5}}\delta a = 0 \tag{13.140}$$

This leads to the same result as is shown in Eq. (13.78) which states that $\delta a = 0$ must be true for the relative orbit to be bounded in a Keplerian system. Note that here δa is defined as

$$\delta a = a_d - a_c \tag{13.141}$$

We would like to expand this method of finding bounded relative orbit constraints to orbit which include the J_2 gravitational perturbation effect. We attack this problem by working in the mean element space such that the mean orbit element drift rates shown in Eq. (13.138) are valid. For notational convenience the overbar notation to denote a parameter as being a mean element is dropped here. Unless stated otherwise, all shown orbit parameters are assumed to be mean elements with the short and long period motion removed. The following algebra is greatly simplified if we work with dimensionless variables. Therefore distances will be measured in Earth radii r_{eq} and time is normalized by the mean motion $n_{eq} = \sqrt{\mu/r_{eq}^3}$ of a satellite at one Earth radius. Let $\tau = tn_{eq}$ be the new time, then mean ascending node rate is written as

$$\frac{d\Omega}{dt} = \frac{d\Omega}{d\tau}\frac{d\tau}{dt} = \Omega' \; n_{eq} \tag{13.142}$$

where the notation $()' = d()/d\tau$ is introduced. Using $\eta = \sqrt{1-e^2}$, the non-dimensional ascending node rate can then be expressed as

$$\Omega' = -\frac{3}{2}J_2\sqrt{\frac{r_{eq}^7}{a^7}}\frac{\cos i}{\eta^4} \tag{13.143}$$

Let us define the non-dimensional semi-major axis measure L as

$$L = \sqrt{\frac{a}{r_{eq}}} \tag{13.144}$$

This allows us to write the there non-dimensional mean orbit element rates as

$$\Omega' = -\frac{3}{2}J_2\frac{\cos i}{L^7\eta^4} \tag{13.145a}$$

$$\omega' = \frac{3}{4}J_2\frac{(5\cos^2 i - 1)}{L^7\eta^4} \tag{13.145b}$$

$$M' = \frac{1}{L^3} + \frac{3}{4}J_2\frac{(3\cos^2 i - 1)}{L^7\eta^3} \tag{13.145c}$$

Since the mean angle quantities M, ω and Ω do not directly contribute to the secular growth caused by J_2, their values can be chosen at will. However, the values of L, η and i (and therefore implicitly a, e and i) must be carefully chosen to match the secular drift rates shown in Eq. (13.145). To keep the satellites from drifting apart over time, it would be desirable to match all three rates (Ω', ω', M') between the various satellites in a given formation. However, this can only be achieved by having the L, η and i being equal, which in return severely restricts the possible relative orbits. Therefore we chose the bounded relative orbit condition to define the relative average drift rate of the angle between the satellite position vectors be zero. This results in

$$\Omega'_d = \Omega'_c \tag{13.146}$$

$$\theta'_{M_d} = M'_d + \omega'_d = \theta'_{M_c} \tag{13.147}$$

where θ_M is the mean argument of latitude. Thus, the arguments of perigee and the mean anomalies are allowed to drift apart. In fact, they end up drifting apart at equal and opposite rates.[19] Imposing equal latitude rates instead of forcing equal argument of perigee and mean anomaly drift has little consequence on the general spacecraft formation geometry if the eccentricity is small. For the case of having a circular orbit (i.e. $e = 0$), then having the relative ω and M drift apart has no consequence at all. However, for relative orbits with a larger eccentricity, having the ω and M drift apart at equal and opposite rates causes the relative orbit to "balloon" out and in again as the argument of perigees drift apart from their nominal values. Combining Eqs. (13.145b) and (13.145c), the mean latitude rate θ'_M is expressed as

$$\theta'_M = \frac{1}{L^3} + \frac{3}{4} J_2 \frac{1}{L^7 \eta^4} \left[\eta \left(3\cos^2 i - 1 \right) + \left(5\cos^2 i - 1 \right) \right] \tag{13.148}$$

The drift rate θ'_{M_d} of a neighboring orbit can be written as a series expansion about the chief orbit element as

$$\theta'_{M_d} = \theta'_{M_c} + \frac{\partial \theta'_{M_c}}{\partial L} \delta L + \frac{\partial \theta'_{M_c}}{\partial \eta} \delta \eta + \frac{\partial \theta'_{M_c}}{\partial i} \delta i + H.O.T. \tag{13.149}$$

where we make use of the fact that $\theta'_M = \theta'_M(L, \eta, i)$ only. Let the difference in latitude rates be $\delta\theta'_M$, then a first order approximation of Eq. (13.149) is written as

$$\delta\theta'_M = \theta'_{M_d} - \theta'_{M_c} = \frac{\partial \theta'_{M_c}}{\partial L} \delta L + \frac{\partial \theta'_{M_c}}{\partial \eta} \delta \eta + \frac{\partial \theta'_{M_c}}{\partial i} \delta i \tag{13.150}$$

Similarly, the difference in nodal rate Ω' is expressed as

$$\delta\Omega' = \frac{\partial \Omega'_c}{\partial L} \delta L + \frac{\partial \Omega'_c}{\partial \eta} \delta \eta + \frac{\partial \Omega'_c}{\partial i} \delta i \tag{13.151}$$

For notational convenience, the sub-script "c" is dropped from here on. Unless noted otherwise, all orbit parameters are assumed to be chief orbit parameters.

To enforce equal drift rates θ'_M and Ω' between neighboring orbits, we must set $\delta\theta'_M$ and $\delta\Omega'$ equal to zero in Eqs. (13.150) and (13.151). Taking the appropriate partial derivatives of Eq. (13.61) and substituting them into Eq. (13.150), the condition that enforces equal latitude rates is rewritten as:

$$-\frac{3}{L^4}\delta L + J_2\frac{21}{4L^8\eta^4}\left[\eta(1-3\cos^2 i)+(1-5\cos^2 i)\right]\delta L + J_2\frac{3}{4L^7\eta^5}\left[3\eta(1-3\cos^2 i)+4(1-5\cos^2 i)\right]\delta\eta - J_2\frac{3}{2L^7\eta^4}(3\eta+5)\cos i\sin i\delta i = 0 \quad (13.152)$$

Note that only the term δL appears without being multiplied by the small parameter J_2. Thus δL must be itself of $\mathcal{O}(J_2)$ and the term involving $J_2\delta L$ is dropped as a higher order term. The first orbit element constraint is then simplified to

$$-\delta L + J_2\frac{1}{4L^3\eta^5}\left[3\eta(1-3\cos^2 i)+4(1-5\cos^2 i)\right]\delta\eta - J_2\frac{1}{2L^3\eta^4}(3\eta+5)\cos i\sin i\delta i = 0 \quad (13.153)$$

Taking the partial derivatives of Eq. (13.145a), the second condition for J_2 invariant orbits is written as

$$J_2\frac{3}{2L^7\eta^5}\left[\frac{7}{L}\cos i\delta L + 4\cos i\delta\eta + \eta\sin i\delta i\right] = 0 \quad (13.154)$$

Since $\delta L = \mathcal{O}(\epsilon)$ the δL term is dropped. Thus, the condition that results in equal nodal precession rates for two neighboring orbits is:

$$\delta\eta = -\frac{\eta}{4}\tan i\delta i \quad (13.155)$$

Observe that as the chief satellite approaches a polar orbit (i.e. i=90 degrees), the necessary change in eccentricity results in an eccentricity greater than unity (hyperbolic orbit) or less than zero. This issue will be discussed in more detail later on. Using the δi defined in Eq. (13.155), we are able to simplify the equal relative latitude rate condition in Eq. (13.153) to

$$\delta L = \underbrace{\frac{J_2}{4L^4\eta^5}(4+3\eta)\left(1+5\cos^2 i\right)}_{D} L\delta\eta \quad (13.156)$$

Combined, Eqs. (13.155) and (13.156) provide the two necessary conditions on the mean element differences between two neighboring orbits to yield a J_2 invariant relative orbit. When designing a relative orbit using the mean orbit element differences, either δi, δe or δa is chosen, and the other two element

differences are then prescribed through the two constraints. The remaining mean orbit element differences $\delta\Omega$, $\delta\omega$ and δM can be chosen at will without affecting the J_2 invariant conditions. Further, note that these two conditions are not precise answers to the nonlinear problem, but are only valid up to a first order approximation. Thus, relative orbits designed with these two conditions will still exhibit some small relative drift.

Example 13.3: Let us investigate the effect of dropping the terms containing $J_2\delta L$ when developing the two orbit element constraint equations in Eq. (13.155) and (13.156). The reason for this simplification is that in Eq. (13.152) δL is the only term appearing without being multiplied by J_2, and thus must itself be of order J_2. However, as the inclination angle approaches either 0 or 90 degrees, then the term in Eq. (13.152) which contains δi would also become very small. Thus, ignoring the $J_2\delta L$ terms in these cases could potentially contribute lead to significant numerical errors.

The following development will show that the error introduced by neglecting the $J_2\delta L$ terms in minimal. If the $J_2\delta L$ terms are retained, then the two J_2 invariant relative orbit conditions take on a more complicated form:

$$\delta\eta = -\frac{\eta\left(7J_2(\eta-2)(5+3\eta)\cos i\sin i + \eta(4L^4\eta^4 - 7J_2(1+\eta))\tan i\right)}{16L^4\eta^5 - 7J_2(4\eta^2+\eta-4) + 7J_2(\eta(11+12\eta)-20)\cos^2 i}\delta i$$

$$\delta L = \frac{J_2(4+3\eta)(1+5\cos^2 i)}{\eta(4L^4\eta^4 - 7J_2(1+\eta)) + 14J_2(\eta-2)(5+3\eta)\cos^2 i}L\delta\eta$$

Note that these orbit element constraints perfectly satisfy the first order conditions in Eqs. (13.152) and (13.154). If the higher order terms are dropped, then the previously presented J_2 invariant orbit constraints are retrieved. However, these more precise conditions on the mean orbit element are also more complex and analytically less trackable than their simplified cousins.

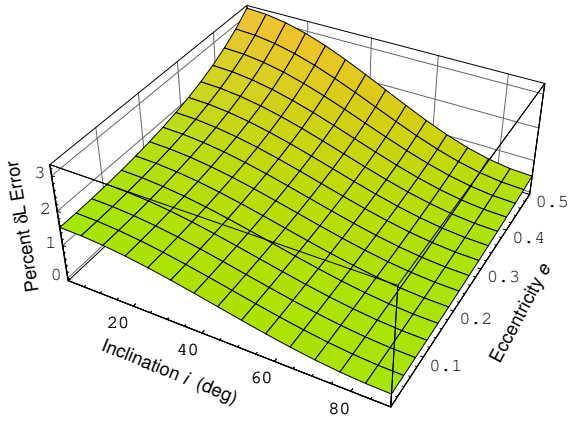


Figure 13.7: Percent Error in Computing δL by dropping the $\epsilon\delta L$ Terms

Figure 13.7 illustrates the percent error in computing the δL correction for a range of eccentricities and inclination angles. Here L is set to be 1.054. As

the figure shows, the numerical errors involved in using the simplified orbit constraint conditions are typically less than or equal to 1.5 percent. Only as the eccentricities grow larger do the numerical errors start to grow larger. It is interesting to note that dropping the $J_2\delta L$ term causes the largest numerical errors for near-zero inclination angles, while near-polar orbits show the least numerical errors. The numerical errors in computing $\delta\eta$ and δi are essentially equivalent. Thus, using the simplified J_2 orbit element constraints results in minimal numerical errors. For cases where the numerical errors are too large, the more precise expressions can be used.

For more physical insight into the J_2 invariant relative orbit constraints, it is convenient to map the differences in L into differences in the semi-major axis a. Recalling that $L = \sqrt{a/r_{eq}}$ (L is a non-dimensional variable), the variations in L and a are related through

$$\delta L = \frac{1}{2L}\frac{\delta a}{r_{eq}} \tag{13.157}$$

Substituting Eqs. (13.157) into Eq. (13.156), the constraint enforcing equal latitude rates between two orbits is rewritten as

$$\delta a = 2D\frac{a^2}{r_{eq}}\delta\eta \tag{13.158}$$

Combined, Eqs. (13.155) and (13.158) form the two necessary mean orbit element constraints expressed in terms of a difference in semi-major axis, eccentricity and inclination angle.

Note that it is numerically preferable to express the differences in eccentricity in terms of $\delta\eta$, and not in terms of the eccentricity δe itself. The reason for this is clear when we observe the variation of $\eta = \sqrt{1-e^2}$.

$$\delta e = -\frac{\eta}{e}\delta\eta \tag{13.159}$$

Using Eq. (13.159) clearly poses numerical difficulties whenever the orbits become circular. A finite change in η would erroneously appear as an infinite change in e. Thus, it is preferable to deal with $\delta\eta$ quantities and then use the precise mapping $\eta = \sqrt{(1-e^2)}$ to map these differences into corresponding δe quantities.

If J_2 is set to zero (i.e. pure Keplerian motion), then we are only left with the constraint that $\delta a = 0$. This makes sense intuitively, since the semi-major axis a determines the orbit period. For Keplerian motion, if the orbit periods are not equal, then the two spacecraft will drift apart. Thus, for Keplerian motion the initial conditions that result in relative motion orbits that do not drift apart are constrained to a five dimensional manifold , or in the (a, e, i) space to a two dimensional manifold, the surface of the sphere as illustrated in Figure 13.8. For a particular chief orbit with a, e and i, the neighboring orbit momenta elements must lie on this surface. However, once the J_2 perturbation

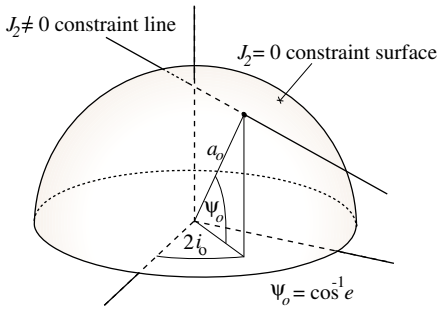


Figure 13.8: Drift Free Constraint Illustration In Momenta Space

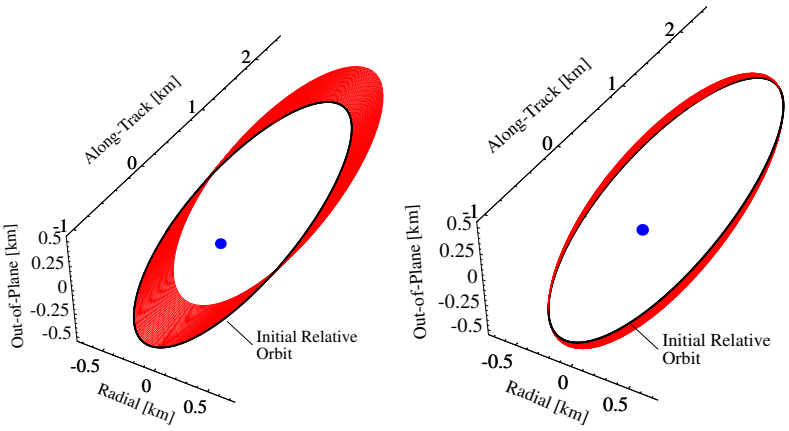
is included, the geometric constraint on the momenta elements to achieve drift free relative motion is a straight line which is not tangent to the sphere surface. Thus, the presence of gravitational perturbations changes the dimension of the constraint manifold from two to one.

Example 13.4: Let us illustrate the relative orbit drifts that may be introduced if the relative orbit is not setup correctly. The chief orbit elements using in this numerical simulation are shown in Table 13.3. The dynamical orbit model used in the numerical simulation includes the J_2 through J_5 gravitational perturbations. The relative orbit is described by choosing the following mean orbit element differences. To achieve some out-of-plane motion, an ascending node difference of $\delta\Omega = 0.005$ degrees is prescribed. The line of perigee and initial mean anomaly differences are set equal and opposite in sign as $\delta\omega = 0.01$ degrees and δM_0 = -0.01 degrees. Further, we chose to prescribe a change in eccentricity $\delta e = 0.0001$ to exaggerate the in-plane relative orbit. Using Eqs. (13.155) and (13.158), the corresponding changes in a and i must be δa = -0.351765 meters and $\delta i = 0.001035$ degrees. Note that both the required δa and δi to compensate for this δe are rather small.

Table 13.3: Mean Chief Satellite Orbit Elements.

Orbit Elements	Value	Units
a	7153	km
e	0.05	
i	48	deg
Ω	0.0	deg
ω	30.0	deg
M_0	0.0	deg

The relative orbits of two different simulation runs are shown in Figure 13.9 as seen in the rotating Hill frame. The plots show the data of 45 orbits, which correspond to roughly 3 days of simulation time. The initial relative orbit is shown as a solid black line, while the path of the remaining 45 orbits is shown



(i) Initial Relative Orbit Setup in Osculating Elements

(ii) Initial Relative Orbit Setup in Mean Elements

Figure 13.9: Relative Orbit Drift for a Non-Polar Master Orbit

as a red line. Both simulations use the same initial orbit element differences. However, in Figure 13.9(i) the initial orbit element differences, which determine the initial shape of the relative orbit, are chosen in osculating element space. Substantial relative orbit drift is apparent due to the perturbative influence of J_2. Figure 13.9(ii) illustrates the drastic improvements that may occur if the initial orbit geometry is setup in mean element space. Since the matching conditions in Eq. (13.158) and (13.155) are only up to first order, the relative orbit will not necessarily be perfectly J_2 invariant. While some periodic thrusting is still necessary, the frequency of these orbit corrections can be greatly reduced.

13.6.2 Energy Levels between J_2-Invariant Relative Orbits

It is interesting to study the energy levels of two neighboring orbits that are J_2 invariant using the necessary first order conditions established in Eqs. (13.155) and (13.158). For the system studied, the Hamiltonian H is the total energy. Including the J_2 term, the averaged energy in terms of normalized orbit elements is given by

$$H = -\frac{1}{2L^2} - J_2 \frac{1}{4L^6\eta^3}(3\cos^2 i - 1) \tag{13.160}$$

Where for Keplerian motion the energy level of an orbit only depends on the semi-major axis measure L, including the J_2 effect makes the energy expression depend on all three elements L, η and i. The difference in energy δH of a

neighboring orbit and a reference orbit is approximated as

$$\delta H \approx \frac{\partial H}{\partial L}\delta L + \frac{\partial H}{\partial \eta}\delta\eta + \frac{\partial H}{\partial i}\delta i \tag{13.161}$$

Computing the partial derivatives in Eq. (13.160) while keeping in mind that δL is of order $\mathcal{O}(J_2)$, we find that

$$\delta H = \frac{1}{L^3}\delta L + J_2\frac{3}{4L^6\eta^4}\left[(3\cos^2 i - 1)\delta\eta + 2\eta\sin i\cos i\delta i\right] \tag{13.162}$$

For two neighboring orbits to be J_2 invariant, the differences in L, η and i must satisfy the two conditions in Eqs. (13.156) and (13.155). Substituting these variational constraints, the energy difference between two J_2 invariant orbits is given by

$$\delta H = -J_2\frac{\tan i}{4L^6\eta^4}(1 + 5\cos^2 i)\delta i \tag{13.163}$$

Eq. (13.163) states that if the two orbits have a non-zero difference in inclination angle δi (or implicitly a difference in η or L), then the two orbits must have different energies. Only if all three elements L, η and i between two orbits are equal will the orbit energies themselves be equal. Note that this condition still allows the two orbits to have different mean M, Ω and ω between the orbits. Thus is it possible to have J_2 invariant relative orbits with zero energy difference between deputy and chief satellites. This energy difference must be non-zero however if the relative orbit is to have out-of-plane relative motion due to having a difference in inclination angles.

From this energy study an interesting conclusion can be made on the out-of-plane stability of the CW equations in Eq. (13.19c). The linearized relative equations of motion clearly indicate that the out-of-plane motion will take on a stable sinusoidal form. The bounded relative orbit constraint $\dot{y}_0 + 2nx_0 = 0$ does not even involve the out-of-plane coordinate z. This constraint was shown to be equivalent to saying that $\delta a =$ or that the orbit energies must be equal. Thus, solely considering the CW equations, it appears that the relative orbit motion is bounded for any out-of-plane motion. However, if any of this out-of-plane motion is due to a difference in inclination angles, then Eq. (13.163) clearly shows the relative energy difference can not be zero if the relative orbit is to be bounded. This illustrates that the CW equations are not well suited for performing any long-term stability study of the relative orbit.

13.6.3 Constraint Relaxation Near Polar Orbits

For near-polar orbits, where the inclination i approaches 90 degrees, the equal relative nodal rate condition in Eq. (13.155), given by

$$\delta\eta = -\frac{\eta}{4}\tan i\ \delta i$$

may pose some practical problems in designing J_2 invariant relative orbits. The issue here is that as i approaches 90 degrees, and the relative orbit design commands out-of-plane motion at the maximum latitude (i.e. δi is non-zero), then the corresponding change in eccentricity becomes unpractically large. The result is that the relative orbit becomes excessively large. Note that this near-polar issue only arises if a specific mean inclination angle difference is prescribed and the two J_2 constraints are then used to compute the necessary mean δa and δi. If a change in mean semi-major axis or eccentricity were required for a near-polar orbit, then the equal nodal rate condition in Eq. (13.155) would require a very small corresponding mean inclination angle difference. Thus, achieving out-of-plane motion the maximum latitude poses the greatest challenge in designing J_2 invariant relative orbits. If the out-of-plane motion should occur during the equator crossing, then this can be achieved by describing a difference in ascending nodes $\delta\Omega$. Since the three angular quantities $\delta\Omega$, $\delta\omega$ and δM can be chosen at will, no practical issues would arise here.

That the relative orbits become excessively large for near-polar orbits if a δi is prescribed was also shown in the previous relative energy discussion. Studying Eq. (13.163) is it clear that if the chief orbit is a polar orbit, a finite δi requires an infinite difference in orbit energy, an unrealistic condition. Thus, as the inclination approaches 90 degrees the size of the relative motion orbits increases.

The problem posed by attempting to design a J_2 invariant relative orbit for a near-polar chief orbit is illustrated in the following numerical simulation. The chief mean orbit elements used are shown in Figure 13.4.

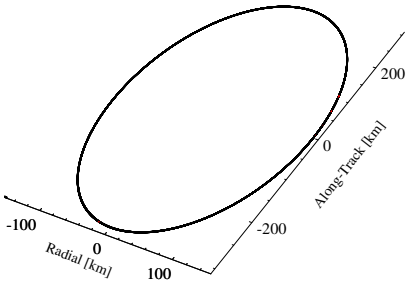
Table 13.4: Mean Chief Satellite Orbit Elements for Near-Polar Case Study.

Mean Orbit Elements	Value	Units
a	7153	km
e	0.05	
i	88	deg
h	0.0	deg
g	30.0	deg
l	0.0	deg

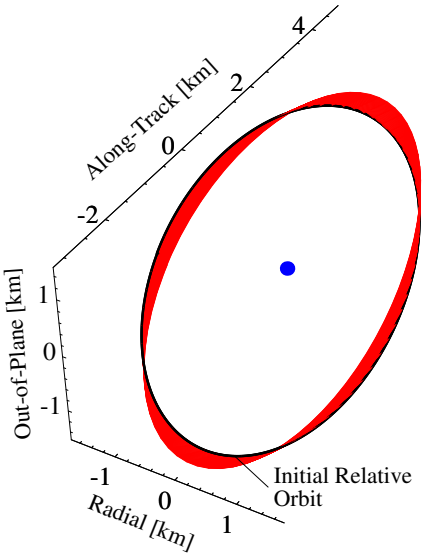
The numerical simulations are performed by integrating the nonlinear orbit equation

$$\ddot{\boldsymbol{r}} = -\frac{\mu}{|\boldsymbol{r}|^3}\boldsymbol{r} + \boldsymbol{f}(\boldsymbol{r}) \tag{13.164}$$

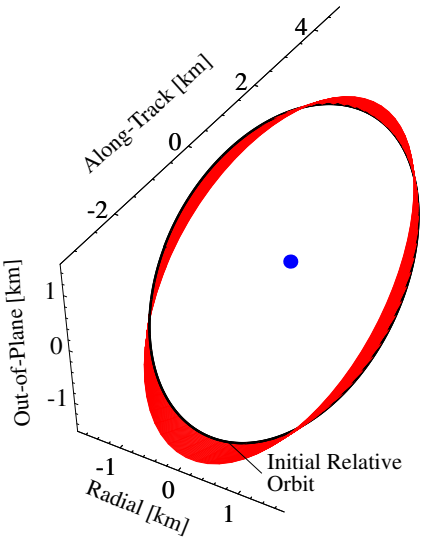
where the perturbative acceleration $\boldsymbol{f}(\boldsymbol{r})$ includes the zonal J_2 through J_5 effects. The relative orbit is described by choosing the mean orbit element differences $\delta\Omega = 0.0$ degrees (all out-of-plane motion produced through δi), $\delta\omega =$



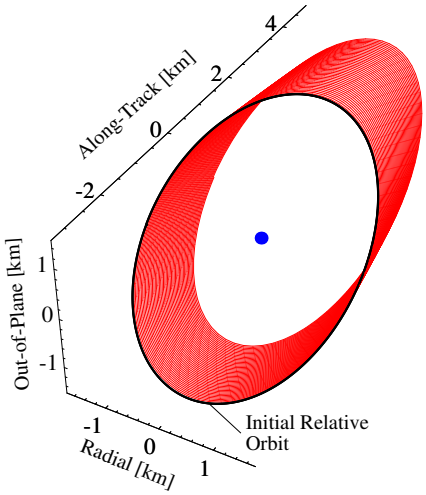
(i) Relative Orbit Enforcing Both J_2 Invariant Conditions



(ii) Relative Orbit Enforcing Only $\delta a = 2Da^2/r_{eq}\delta\eta$

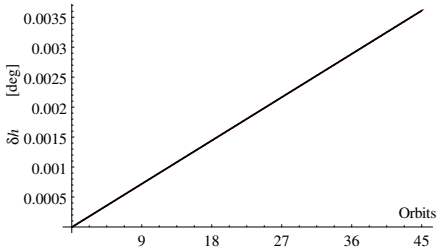


(iii) Inclination Angle Error δi (deg)

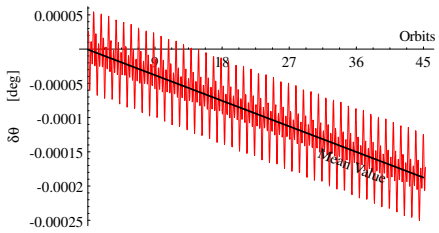


(iv) Relative Orbit Setup Performed in Osculating Orbit Elements

Figure 13.10: Relative Orbit Drift for a Near-Polar Chief Orbit



(i) Difference in Ascending Node Ω



(ii) Difference in Mean Latitude Angle θ_M

Figure 13.11: Mean and Osculating Orbit Element Differences for Case 2

0.1 degrees and δM = -0.1 degrees. Case 1 assumes the relative orbit geometry requires a δi of 0.01 degrees to achieve roughly 1 km of out-of-plane motion at maximum latitude. To achieve a desired δi of 0.01 degrees without inducing relative drift in the other orbit elements, the remaining two momenta elements differences must be $\delta e = 0.020648$ degrees and δa = -27.2122 meters. The resulting relative orbit is shown in Figure 13.10(i). Note that the necessary difference in eccentricity is very large, causing the relative orbit to become very large in the along track and radial direction. However, no apparent drift is visible for the 45 orbits plotted on the scale shown.

One method suggested in Reference 19 is to drop the equal relative nodal rate condition in Eq. (13.155) when a δi is prescribed for a near-polar chief orbit. The δi of 0.01 degrees is retained in case 2 shown in Figure 13.10(ii), but it is not used to prescribe a corresponding difference in eccentricity. Instead, a δe of 0.0001 is chosen and the corresponding δa of -0.24157 meters is computed through the equal relative latitude rate condition in Eq. (13.158). The resulting relative orbit does exhibit some drift since the ascending nodes are drifting apart. Over a year, the Δv required to compensate for this drift is roughly 56.8 m/s. However, for case 3 the equal latitude rate condition is also dropped (i.e. δa = 0 meters for the same δe), then the resulting orbit shown in Figure 13.10(iii) has some clear along-track drift. Case 4 has the same initial orbit element differences as the ones used in case e, but here the orbits were established using osculating orbit elements instead of mean orbit elements. The resulting relative orbit is shown in Figure 13.10(iv). This would be analogous to setting up the relative orbit initial conditions using the CW or Hills equations. Over the 45 orbits shown, clearly substantial drift would result. This emphasized the point that one should be working with mean orbit elements when design the relative orbits.

Figure 13.11 illustrates the relative nodal and latitude rate drifts for case 2. By dropping the equal nodal rate condition, the nodes clearly drift apart over time. The corresponding osculating relative ascending node variations are not visible due to the large drift. While the relative latitude drift is not perfectly

zero, it is kept very small. The fuel estimate to compensate for the $\delta\theta_M$ drift over one year is only 1.45 m/s, while it would be about 14.1 m/s if the equal latitude rate condition is dropped. However, as a comparison, to compensate for the relative ascending node drift it would take about a fuel cost of 56 m/s over year to compensate.

Thus, it is possible to design relative orbits with out-of-plane motion created by an inclination change and a chief orbit that is near-polar. However, the equal ascending node rate condition must be dropped here to obtain a relative orbit of practical value. Periodic maneuvers will be required to compensate for the $\delta\Omega$ drift. References 20 and 21 present continuous feedback and impulsive control schemes respectively in terms of the mean orbit elements. These methods will be discussed later in this chapter. For an orbit such as is presented in case 2, it would make sense to use the impulsive control scheme where the ascending node is correct during the polar region crossings using the out-of-plane burn:

$$\Delta v_{h_\Omega} = \frac{h \sin i}{r \sin\theta} \Delta\Omega \qquad \text{for } \theta = \pm 90 \text{ degrees} \tag{13.165}$$

Note that $\theta = \omega + f$ is the true latitude angle.

13.6.4 Near-Circular Chief Orbit

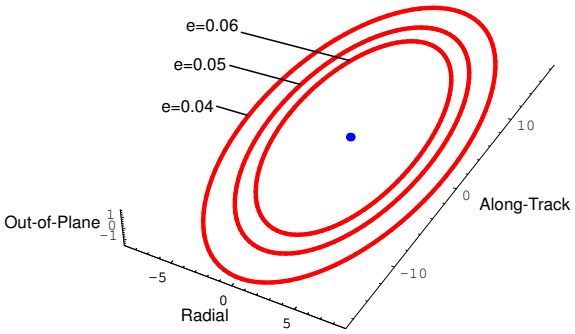
As the chief's orbit eccentricity becomes small, the eccentricity differences commanded by the equal nodal rate condition may cause the relative orbit to become very large in the along-track direction. This is clear from the linear mapping of differences in e to differences in η shown in Eq. (13.159) to be:

$$\delta e = \frac{\eta}{e} \delta\eta$$

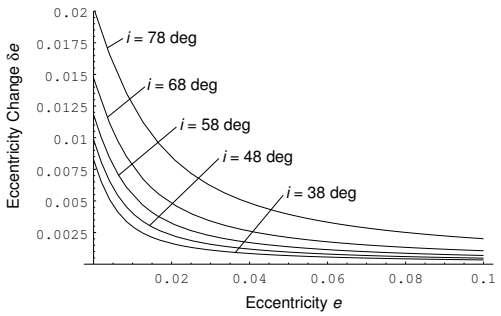
However, the change in e does not become infinitely large as $e \to 0$. The equal nodal rate condition in Eq. (13.155) shows a finite required difference in η as e goes to zero and η goes to one. Using the nonlinear relationship $\eta = \sqrt{1 - e^2}$, this finite $\delta\eta$ corresponds to a finite δe for a circular orbit. However, these eccentricity changes may still result in a relative orbit which is too large for practical use. Again, as was the case with near-polar chief orbits, if the out-of-plane motion can be produced by a change in node instead of a change in inclination angle, then having a chief orbit with a small eccentricity would not pose any practical difficulties.

A numerical simulation is performed to illustrate this behavior. The chief orbit elements are shown in Table 13.3. The relative orbit is established using the mean orbit element differences of $\delta\Omega = 0.01$ degrees, $\delta\omega = 0.01$ degrees and δM = -0.01 degrees. An inclination angle difference δi of 0.01 degrees is requested. The relative orbits were computed for the three mean chief eccentricities 0.04, 0.05 and 0.06.

Figure 13.12(i) compares the resulting three relative orbits. For the case where $e = 0.06$, the requested δi required a δe of 0.000799. The case where $e = 0.05$ resulted in a δe of 0.000957 and the case with $e = 0.04$ resulted in $\delta e =$



(i) Relative Orbit in LVLH Frame for various Eccentricities



(ii) Eccentricity Differences for Small Eccentricities

Figure 13.12: Small Eccentricity Case Study

0.001191. Clearly the smaller chief eccentricities result in a larger relative orbit in the along track direction.

This general behavior is also illustrated in Figure 13.12(ii) where the required δe for a δi of 0.01 degrees are displayed for various chief eccentricities e and inclination angles i. Due to the $\tan i$ term in the equal nodal rate condition, the effect of having small eccentricities is enhances for larger inclination angles. The δe here were computed using the nonlinear mapping between η and e. While the required eccentricity for the relative motion orbit does grow large as e approaches zero, it reaches a finite limit for a circular chief orbit case and does not become infinite.

This result is interesting in that it states that it is easier to compensate for out-of-plane motion induced by δi if the chief orbit has a larger eccentricity. The richer dynamics of having a more eccentric orbit makes it easier to compensate for the relative nodal drift condition.

13.6.5 Relative Argument of Perigee and Mean Anomaly Drift

To establish the J_2 invariant orbits, conditions are established which set the relative ascending node rate $\delta\dot{\Omega}$ and latitude rate $\delta\theta'_M$ equal to zero. While this guarantees that the angle between the chief and deputy position vector remains constant, it is possible that the argument of perigee and mean anomaly differences drift apart. The effect of this is that for chief orbits with non-zero eccentricity, the relative orbit geometry swells larger as $\delta\omega$ and δM drift apart and then shrinks again as they eventually approach each other. Since the relative latitude rate is equal to zero when the two presented J_2 constraint conditions are satisfied, then we know that

$$\delta\omega' = -\delta M' \tag{13.166}$$

To compute the relative drift in the argument of perigee, we take the partial derivative of Eq. (13.145b).

$$\delta\omega' = \frac{\partial\omega'}{\partial L}\delta L + \frac{\partial\omega'}{\partial\eta}\delta\eta + \frac{\partial\omega'}{\partial i}\delta i \tag{13.167}$$

After substituting the J_2 invariant conditions in Eqs. (13.158) and (13.155), the relative perigee drift rate is found to be

$$\delta\omega' = J_2\frac{3}{4L^7\eta^4}\left(\tan i(5\cos^2 i - 1) - 5\sin(2i)\right)\delta i \tag{13.168}$$

The following numerical simulation illustrates the effect of the perigee/mean anomaly drift has on the relative orbit geometry. The chief orbit elements are the same as are shown in Table 13.3 with an eccentricity set to be 0.05. A mean δi of 0.01 degrees is prescribed and the mean $\delta\Omega$ is set equal to 0.01 degrees. The argument of perigee and mean anomaly differences are set equal to

$$\delta\omega = -\delta M = 0.0, \quad 0.5 \text{ or } \quad 1.0 \quad \text{degrees}$$

The resulting three relative orbits are illustrated in the rotating Hill frame in Figure 13.13. As the argument of perigee and mean anomaly differences drift apart, the overall relative orbit geometry is expanded without changing the shape itself appreciably. Not that the presented orbit has a relatively large eccentricity of 0.05. If the eccentricity were closer to zero, then the effect of the perigee/mean anomaly drift on the relative orbit geometry would be even less. At the limiting case where the chief orbit becomes circular, the perigee/mean anomaly drift would have no effect on the relative geometry.

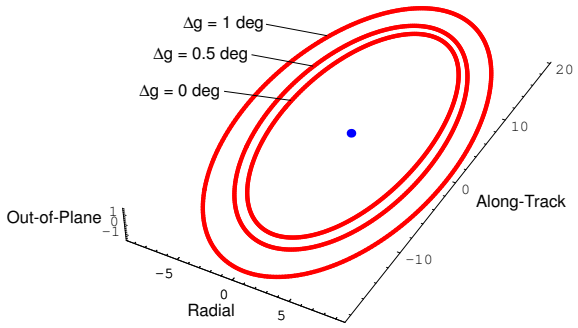


Figure 13.13: Relative Orbits in LVLH Frames for Three Different Argument of Perigee and Mean Anomaly Differences

While this drift in $\delta\omega$ and δM is an effect that may have to be periodically compensated for, the argument of perigee and mean anomaly drift occurs very slowly. For the presented numerical simulation, the $\delta\omega$ had only drifted 0.05 degrees after 45 revolutions (roughly three days). Thus, for $\delta\omega$ to drift the 1.0 degrees shown in Figure 13.13, it would take at approximately 60 days.

To correct such specific orbit element differences, Reference 21 developed an impulsive feedback control scheme with the mean orbit element errors as the feedback quantity. While this scheme is able to correct any types of orbit element errors, the ω and M correction are of interest to the present problem. Let Δv_{r_p} a orbit radial thrust performed at perigee, and Δv_{r_a} the orbit radial thrust performed at apogee. In order to correct a specific $\Delta\omega = -\Delta M$ error, the following control is used.

$$\Delta v_{r_p} = -\frac{na}{4}\left(\frac{(1+e)^2}{\eta} - 1\right)\Delta\omega \tag{13.169}$$

$$\Delta v_{r_a} = \frac{na}{4}\left(\frac{(1-e)^2}{\eta} - 1\right)\Delta\omega \tag{13.170}$$

The advantage of this impulsive firing scheme is that only the osculating ω and M are adjusted in a near-optimal manner. Reference 21 goes into further details describing how this scheme can also be used to correct for mean orbit element errors.

13.6.6 Fuel Consumption Prediction

As has been shown in the previous discussion, at times it may be beneficial to relax the two constraints on the mean orbit elements in order to obtain a relative orbit solution which is of practical value. We would like to present convenient formulas which allow us to predict the fuel cost in terms of Δv's that must be applied to cancel any J_2 induced drift if the orbit elements a, e and i do not perfectly match the conditions in Eqs. (13.155) and (13.158). To perform this analysis, it is convenient to use the dimensional mean orbit element drift equations.

$$\dot{\Omega} = -\frac{3}{2} J_2 \frac{r_e^2}{a^2} \frac{n}{\eta^4} \cos i \tag{13.171}$$

$$\dot{\omega} = \frac{3}{4} J_2 \frac{r_e^2}{a^2} \frac{n}{\eta^4} (5\cos^2 i - 1) \tag{13.172}$$

$$\dot{M} = n - \frac{3}{4} J_2 \frac{r_e^2}{a^2} \frac{n}{\eta^3} (1 - 3\cos^2 i) \tag{13.173}$$

The methodology to compute the fuel cost to combat the J_2 induced drift will be the same for all the cases. First, we will compute how much drift the momenta orbit element differences δa, δe and δi will cause over one orbit. Then, using impulsive control, we are able to provide an estimate of what Δv would be required to cancel the J_2 induced drift. Note that these fuel estimates will not be precise predictions, but rather they provide a convenient method to quickly assess how much fuel would be required to combat the J_2 perturbation if the orbit element differences are not set at their ideal J_2 invariant values.

Ascending Node Relative Drift Correction Cost Estimate

First, we find an estimate of the fuel required to control the J_2 induced ascending node drift. The derivative of Eq. (13.171) is used to compute the relative nodal drift $\delta\dot{\Omega}$. Note that advantage is taken here of the fact that the semi-major axis differences δa are assumed to be of order J_2 and are thus ignored here as higher order terms.

$$\delta\dot{\Omega} = \frac{3}{2} J_2 \frac{r_e^2}{a^2} \frac{n}{\eta^5} (\eta \sin i \ \delta i + 4\cos i \ \delta\eta) \tag{13.174}$$

The orbit period P of the chief satellite is given by

$$P = \frac{2\pi}{n} \tag{13.175}$$

The J_2 induced drift in the ascending node over one orbit period is then given by

$$\Delta\Omega_{orbit} = \delta\dot{\Omega} \cdot P = 3J_2\pi \frac{r_e^2}{a^2\eta^5} (\eta \sin i \ \delta i + 4\cos i \ \delta\eta) \tag{13.176}$$

Eq. (13.176) provides an estimate of the amount of ascending node correction that would be required per orbit. To compute what Δv would be required to perform these corrections, the impulsive control scheme developed in Reference 21 is used here. The impulses developed in this control law correct specific orbit element errors are based on Gauss' variational equations.[22] The ideal time to perform a node correction is during the polar crossings where $\theta = \pm 90$ degrees. Firing an impulse Δv_h in the orbit normal direction, the following node correction is achieved:

$$\Delta v_h = \frac{h \sin i}{r_h} \Delta h \tag{13.177}$$

Note that r_h is the orbit radius at $\theta = \pm 90$ degrees. After substituting Eq. (13.176) into Eq. (13.177), and performing several simplifications, the following fuel estimate is found to counter a J_2 induced nodal drift.

$$\Delta v_h = 3 J_2 \pi \frac{r_e^2}{r_h} \frac{n}{\eta^4} \sin i \, (\eta \sin i \; \delta i + 4 \cos i \; \delta \eta) \tag{13.178}$$

Note that this Δv estimate is the fuel required per orbit. To find a yearly fuel budget estimate, this figure needs to be multiplied by the number of orbits that occur in one year.

As expected, if the mean orbit element differences δi and $\delta \eta$ satisfy the equal nodal rate condition in Eq. (13.155), then the predicted fuel budget is zero. Note that the actual fuel budget would not be zero though. This is because several first order approximations were made in developing the two constraints in Eqs. (13.155) and (13.158).

Eq. (13.178) does provide a very convenient method to quickly estimate the fuel budget if the J_2 invariant conditions are not setup perfectly. Assume the relative orbit is designed using the linear CW equations. Here the chief orbit is circular and we set the inclination angle equal to 70 degrees and the semi-major axis equal to 7000 km. To obtain an out-of-plane motion of roughly one kilometer, a δi of 0.01 degrees is required. Using Eq. (13.178), this leads to an annual fuel budget estimate of 43.6 m/s solely to correct for the relative ascending node drift. A cost which could be avoided if the J_2 perturbation is taking into account when designing the relative orbit.

Argument of Perigee and Mean Anomaly Relative Drift Correction Cost Estimate

After having found a fuel budget estimate to correct the relative nodal drift, fuel budget estimates are now developed to correct for both the relative argument of perigee drift and mean anomaly drift. Taking the derivative of Eq. (13.172) and making use again of the fact that δa is of the order of J_2, the relative argument of perigee drift rate is expressed as

$$\delta \dot{\omega} = -\frac{3}{4} J_2 \frac{r_e^2}{a^2} \frac{n}{\eta^5} \left(5 \eta \sin(2i) \delta i + 4 \, (5 \cos^2 i - 1) \delta \eta \right) \tag{13.179}$$

Using Eq. (13.175), the perigee drift over one orbit is estimated to be

$$\Delta\omega_{orbit} = \delta\dot{g} \cdot P = -J_2 \frac{3\pi}{2\eta^4} \frac{r_e^2}{a^2} \left(5\eta \sin(2i)\delta i + 4\ (5\cos^2 i - 1)\delta\eta\right) \tag{13.180}$$

The mean anomaly drift over orbit is computed in an analogous manner. Note, however, that here δa appears without being multiplied by J_2 and is thus retained.

$$\Delta l_{orbit} = \delta\dot{l} \cdot T = -\frac{3\pi}{a}\delta a - \frac{9\pi}{2} J_2 \frac{r_e^2}{a^2} \frac{n}{\eta^4} \left(\eta \sin(2i)\delta i - (1 - 3\cos^2 i)\delta\eta\right) \tag{13.181}$$

Again, note that Eqs. (13.180) and (13.181) provide angular drift estimates for one orbit period. To compute the annual drift, these figures would be multiplied by the number of orbit period in a year.

To compute Δv's necessary to perform the required $\Delta\omega$ and ΔM corrections, the two impulse technique presented in Reference 21 is used. Here an orbit radial thrust is applied at both perigee and apogee to achieve the desired orbit element corrections in a near-optimal manner and without affecting the remaining orbit elements. Using this method, the two Δv are then computed through

$$\Delta v_{r_p} = -\frac{na}{4} \left(\frac{(1+e)^2}{\eta} \Delta\omega + \Delta M \right) \tag{13.182}$$

$$\Delta v_{r_a} = \frac{na}{4} \left(\frac{(1-e)^2}{\eta} \Delta\omega + \Delta M \right) \tag{13.183}$$

where $\Delta\omega$ and ΔM are computed through Eqs. (13.180) and (13.181) respectively. The total fuel estimate required to control either relative argument of perigee drift, relative mean anomaly drift or both is then computed as

$$\Delta v_{\omega,M} = |\Delta v_{r_p}| + |\Delta v_{r_a}| \tag{13.184}$$

Relative Mean Latitude Drift Correction Cost Estimate

While Eq. (13.184) is convenient to estimate the fuel budget to correct for g or l relative drifts, for the formation flying problem this is of lesser importance. What is more critical is what is the fuel budget to combat the latitude drift, i.e. the sum of both the relative perigee and mean anomaly drift. For nearly circular orbits the argument of perigee and mean anomaly can drift apart with negligible effect on the relative orbit geometry, as long the sum of their drifts is zero. In this section we will provide a fuel budget estimate to control the relative latitude drift. The amount of mean latitude drift rate is computed through

$$\delta\dot{\theta}_M = \delta\dot{\omega} + \delta\dot{M} \tag{13.185}$$

To estimate how much fuel is required to correct a latitude error, it is assumed that a Δv is applied to change the semi-major axis a (and thus the orbit

period) which will speed up or slow down the satellite such that it correct the $\delta\theta_M$ error over one orbit. At the end of the correction, a second such δa adjustment must be made to reinsert the satellite in the previous orbit. From Gauss' variational equation in Eq. (11.153a), the required Δv for a given δa is

$$\Delta v = \frac{h}{2a^2(1+e)}\delta a = \frac{n}{2}\sqrt{\frac{1-e}{1+e}}\delta a \tag{13.186}$$

if the Δv is applied at perigee. To relative the change in semi-major axis δa to the corresponding change in orbit period δP, we differentiate Eq. (13.175) and make use of $n = \sqrt{\mu/a^3}$.

$$\delta a = \frac{2a}{3}\frac{\delta P}{TP} \tag{13.187}$$

The final step is to relate the latitude drift amount $\delta\theta_M$ per orbit to the required orbit period change δP which will accomplish this correction. This is found through

$$\delta\theta_{M_{orbit}} = \delta\dot{\theta}_M \cdot P = n \cdot \delta P \tag{13.188}$$

Substituting Eqs. (13.187) and (13.188) into Eq. (13.186), a fuel budget estimate to correct the per orbit latitude drift is

$$\Delta v_{\theta_M} = \frac{a}{3}\sqrt{\frac{1-e}{1+e}}\delta\dot{\theta}_M \tag{13.189}$$

If the δa, δe and δi differences satisfy the conditions in Eqs. (13.155) and (13.158), then the latitude drift $\delta\dot{\theta}_M$ becomes zero, resulting in a zero fuel budget estimate.

13.7 Relative Orbit Control Methods

This section develops various relative orbit control laws. Typically, this feedback control laws operates on the orbit elements. Gauss' variational equations of motion, shown in Eq. (11.153), provide a convenient set of equations relating the effect of a control acceleration vector $\boldsymbol{u}$ to the osculating orbit element time derivatives.[22] They are repeated here for convenience:

$$\frac{da}{dt} = \frac{2a^2}{h}\left(e\sin f u_r + \frac{p}{r}u_\theta\right) \tag{13.190a}$$

$$\frac{de}{dt} = \frac{1}{h}\left(p\sin f u_r + \left((p+r)\cos f + re\right)u_\theta\right) \tag{13.190b}$$

$$\frac{di}{dt} = \frac{r\cos\theta}{h}u_h \tag{13.190c}$$

$$\frac{d\Omega}{dt} = \frac{r\sin\theta}{h\sin i}u_h \tag{13.190d}$$

$$\frac{d\omega}{dt} = \frac{1}{he}\left[-p\cos f u_r + (p+r)\sin f u_\theta\right] - \frac{r\sin\theta\cos i}{h\sin i}u_h \tag{13.190e}$$

$$\frac{dM}{dt} = n + \frac{\eta}{he}\left[(p\cos f - 2re)u_r - (p+r)\sin f u_\theta\right] \tag{13.190f}$$

where the control acceleration vector $\boldsymbol{u}$ is written in the deputy Hill frame components as

$$\boldsymbol{u} = \left(u_r, u_\theta, u_h\right)^T \tag{13.191}$$

with u_r pointing radially away from Earth, u_h being aligned with the orbit angular momentum vector and u_θ being orthogonal to the previous two directions. The parameter f is the true anomaly, r is the scalar orbit radius, p is the semilatus rectum and the true latitude angle is $\theta = \omega + f$.

13.7.1 Mean Orbit Element Continuous Feedback Control Laws

Since the relative orbit is being described in terms of relative differences in mean orbit elements when establishing J_2 invariant relative orbits, we examine a feedback law in terms of mean orbit elements instead of the more traditional approach of feeding back position and velocity vector errors. Doing so will allow us to control and correct specific orbit element errors. Not all orbit position errors are created equal. An error in the ascending node should be controlled at a different time in the orbit than an error in the inclination angle.

The mean angular velocity n is defined as

$$n = \sqrt{\frac{\mu}{a^3}} \tag{13.192}$$

Note that Gauss' variational equations in Eq. (13.190) were derived for Keplerian motion. In matrix form they are expressed as

$$\dot{\boldsymbol{e}}_{osc} = (0,0,0,0,0,n)^T + [B(\boldsymbol{e}_{osc})]\boldsymbol{u} \tag{13.193}$$

with $\boldsymbol{e}_{osc} = (a, e, i, \Omega, \omega, M)^T$ being the osculating orbit element vector and the 6×3 control influence matrix $[B]$ being defined as

$$[B(\boldsymbol{e})] = \begin{bmatrix} \frac{2a^2 e\sin f}{h} & \frac{2a^2 p}{hr} & 0 \\ \frac{p\sin f}{h} & \frac{(p+r)\cos f + re}{h} & 0 \\ 0 & 0 & \frac{r\cos\theta}{h} \\ 0 & 0 & \frac{r\sin\theta}{h\sin i} \\ -\frac{p\cos f}{he} & \frac{(p+r)\sin f}{he} & -\frac{r\sin\theta\cos i}{h\sin i} \\ \frac{\eta(p\cos f - 2re)}{he} & -\frac{\eta(p+r)\sin f}{he} & 0 \end{bmatrix} \tag{13.194}$$

Let the vector $\boldsymbol{e} = \left(a, e, i, \Omega, \omega, M\right)^T$ be the classical mean orbit element vector, and

$$\boldsymbol{e} = \boldsymbol{\xi}(\boldsymbol{e}_{osc}) \tag{13.195}$$

be an analytical transformation from the osculating orbit elements $\boldsymbol{e}_{osc}$ to the mean elements $\boldsymbol{e}$. In this study, a first order truncation of Brouwer's analytical satellite solution is used as shown in Appendix G.[17] Incorporating the J_2 influence, we write Gauss' variational equations for the mean motion as

$$\dot{\boldsymbol{e}} = [A(\boldsymbol{e})] + \left[\frac{\partial \boldsymbol{\xi}}{\partial \boldsymbol{e}_{osc}}\right]^T [B(\boldsymbol{e}_{osc})]\boldsymbol{u} \tag{13.196}$$

with the 6×1 plant matrix $[A(\boldsymbol{e})]$ being defined as

$$[A(\boldsymbol{e})] = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -\frac{3}{2}J_2\left(\frac{r_{eq}}{p}\right)^2 n\cos i \\ \frac{3}{4}J_2\left(\frac{r_{eq}}{p}\right)^2 n(5\cos^2 i - 1) \\ n + \frac{3}{4}J_2\left(\frac{r_{eq}}{p}\right)^2 \eta n(3\cos^2 i - 1) \end{bmatrix} \tag{13.197}$$

Studying Brouwer's transformation between osculating and mean orbit elements, it is evident that the matrix $[\partial \boldsymbol{\xi}/\partial \boldsymbol{e}_{osc}]$ is approximately a 6×6 identity matrix with the off-diagonal terms being of order J_2 or smaller. Therefore, for the purposes of developing a feedback control law, it is reasonable to approximate the mean orbit element rate equation as

$$\dot{\boldsymbol{e}} \approx [A(\boldsymbol{e})] + [B(\boldsymbol{e})]\boldsymbol{u} \tag{13.198}$$

The plant matrix $[A]$ in Eq. (13.197) rigorously describes the behavior of the mean orbit elements. As is clearly seen here, the J_2 perturbation has no secular effect on the elements a, e and i. The control influence matrix $[B]$, developed in Gauss' variational equations shown in Eq. (13.190), allows us to compute a change in osculating orbit elements due to a control acceleration vector $\boldsymbol{u}$. It is assumed that these osculating orbit element changes, as indicated in Eq. (13.198), are directly reflected in corresponding mean orbit element changes. For example, if a thrust is applied to change the osculating inclination angle by one degree, then the corresponding mean inclination angle is also changed by one degree. The errors introduced by this assumption will be of order J_2. Further, since the difference in osculating and mean orbit elements is relatively small for J_2 perturbations, the numerical difference in computing $[B]$ using osculating or mean orbit elements is typically negligible. In our use of Eq. (13.198) below, we assume that the $[B(\boldsymbol{e})]$ matrix is computed using mean orbit elements.

Let us assume that the relative orbit was set up such that the deputy satellite has a specific mean orbit element difference $\Delta\boldsymbol{e}$ relative to the chief mean orbit elements $\boldsymbol{e}_1$. At any instance, the desired deputy satellite location $\boldsymbol{e}_d$ is expressed in terms of mean orbit elements as

$$\boldsymbol{e}_d = \boldsymbol{e}_c + \Delta\boldsymbol{e} \tag{13.199}$$

Note that $\Delta \boldsymbol{e}$ is a fixed mean orbit element difference. Therefore it doesn't matter if the chief orbit was slightly perturbed by other influences such as atmospheric or solar drag. The relative orbit is always defined as a specific difference relative to the current chief mean orbit elements, in order to maintain a specific relative motion.

Given the true set of mean orbit elements $\hat{\boldsymbol{e}}_d$ of the deputy satellite, the relative orbit tracking error $\delta \boldsymbol{e}$ is expressed in terms of mean orbit elements as

$$\delta \boldsymbol{e} = \hat{\boldsymbol{e}}_d - \boldsymbol{e}_d \tag{13.200}$$

The Lyapunov control theory, presented in Chapter 7, is used here to develop a feedback control law. We define the Lyapunov function V as a positive definite measure of the mean orbit element tracking error $\delta \boldsymbol{e}$.

$$V(\delta \boldsymbol{e}) = \frac{1}{2} \delta \boldsymbol{e}^T \delta \boldsymbol{e} \tag{13.201}$$

Assuming the desired relative orbits are J_2 invariant (i.e. are natural, unforced solutions of the relative equations of motion), the derivative of $\boldsymbol{e}_d$ is

$$\dot{\boldsymbol{e}}_d = [A(\boldsymbol{e}_d)] \tag{13.202}$$

where no control is required to maintain this evolving orbit. Clearly non-J_2 perturbations are being treated as minor disturbances and are not considered in Eq. (13.202). Taking the derivative of V and substituting Eqs. (13.198) and (13.200), we find

$$\dot{V} = \delta \boldsymbol{e}^T \delta \dot{\boldsymbol{e}} = \delta \boldsymbol{e}^T \left([A(\hat{\boldsymbol{e}}_d)] - [A(\boldsymbol{e}_d)] + [B(\boldsymbol{e})] \boldsymbol{u} \right) \tag{13.203}$$

Setting $\dot{V}$ equal to the negative definite quantity

$$\dot{V} = -\delta \boldsymbol{e}^T [P] \delta \boldsymbol{e} \tag{13.204}$$

where $[P]$ is a positive definite feedback gain matrix, we arrive at the following control constraint for Lyapunov stability of the closed-loop departure motion dynamics.

$$[B] \boldsymbol{u} = -([A(\hat{\boldsymbol{e}}_d)] - [A(\boldsymbol{e}_d)]) - [P] \delta \boldsymbol{e} \tag{13.205}$$

Note that $[P]$ does not have to be a constant matrix. In fact, later on, we will make use of this fact to encourage certain orbit element corrections to occur during particular phases of the orbit. Using Eq. (13.194) to study the effectiveness of the control vector to influence a particular orbit element, one

choice is to give the feedback gain matrix $[P]$ the following diagonal form

$$P_{11} = P_{a0} + P_{a1} \cos^N \frac{f}{2} \tag{13.206a}$$

$$P_{22} = P_{e0} + P_{e1} \cos^N f \tag{13.206b}$$

$$P_{33} = P_{i0} + P_{i1} \cos^N \theta \tag{13.206c}$$

$$P_{44} = P_{\Omega 0} + P_{\Omega 1} \sin^N \theta \tag{13.206d}$$

$$P_{55} = P_{\omega 0} + P_{\omega 1} \sin^N f \tag{13.206e}$$

$$P_{66} = P_{M0} + P_{M1} \sin^N f \tag{13.206f}$$

with N being an even integer. The various feedback gains are now at a maximum whenever the corresponding orbit elements are the most controllable, and at a minimum or essentially zero when they are the least controllable. The size of N is chosen such that the P_{i1} gain influence drops off and rises sufficiently fast. Clearly there are an infinity of heuristic feedback gain logics that could be used here which belong to the stabilizing family. We could alternatively pose an optimization problem and optimize $[P(t)]$ to extremize some performance measure. For illustration purposes, we simply choose several stable controllers in this text.

One issue of writing the satellite equations of motion in first-order form in Eq. (13.198) becomes quickly apparent. Since the control vector only has three components, and we are attempting to control six orbit elements, we can't directly solve the control constraint equation in Eq. (13.205) for the control vector $\boldsymbol{u}$. Since the system of equations is over determined, we employ a least-square type inverse to solve for $\boldsymbol{u}$.

$$\boldsymbol{u} = -\left([B]^T[B]\right)^{-1}[B]^T\left(([A(\hat{\boldsymbol{e}}_d)] - [A(\boldsymbol{e}_d)]) + [P]\delta\boldsymbol{e}\right) \tag{13.207}$$

Due to the imprecise nature of the least-squares inverse, the resulting control law is no longer guaranteed to satisfy the stability constraint in Eq. (13.205). However, as numerical simulations show, this control law does successfully cancel mean element tracking errors and reestablish the desired relative orbit.

Other control methods could be employed to control the mean element tracking error defined in Eq. (13.200). The advantage of this method is the presence of the time varying 6×6 feedback gain matrix $[P]$. In particular, it allows us to selectively cancel particular orbit element errors at any time. A classical example is correcting for ascending node and inclination angle errors. Studying Eq. (13.190) or (13.194), it is evident that the feedback gain for $\delta\Omega$ should be large whenever $\theta = \pm 90$ degrees and near-zero whenever $\theta = 0, 180$ degrees. Near the equator it is known that the control effort required to correct for a $\delta\Omega$ would be very large. Therefore nodal corrections are best performed near the polar regions. Analogously, the inclination angle changes are best performed near the equator, with little or no inclination corrections being done near the polar region. Depending on the chief orbit elements, similar statements can be made for the remaining orbit elements. The result is that one can easily

design a variable gain control law which will wait for the satellite to be in an advantageous position within the orbit before correcting certain orbit element errors. Note that this approach enables one to simultaneously control the long term secular orbital dynamics (by considering orbit element control and using mean orbit elements) and to effectively time the control corrections during each orbit to "cooperate with the physics" of orbital dynamics.

The feedback law in Eq. (13.207) contains a term computing the difference in natural mean element rates between the actual mean orbit element vector $\hat{\boldsymbol{e}}_d$ of the deputy satellite and the desired mean orbit element vector $\boldsymbol{e}_d$. If the difference in actual and desired mean orbit elements of the deputy is small, as is typically the case with spacecraft formation flying, then it can be shown that this difference is very small and has a negligible influence on the control law. Linearizing this difference about the desired mean orbit element vector $\boldsymbol{e}_d$, we find

$$[A(\hat{\boldsymbol{e}}_d)] - [A(\boldsymbol{e}_d)] \simeq \left[\frac{\partial A}{\partial \boldsymbol{e}}\right]\Bigg|_{\boldsymbol{e}_d} \delta\boldsymbol{e} = [A^*(\boldsymbol{e}_d)]\delta\boldsymbol{e} \tag{13.208}$$

Using Eq. (13.208), we are able to write the linearized mean element error dynamics as

$$\delta\dot{\boldsymbol{e}} \simeq [A^*(\boldsymbol{e}_d)]\delta\boldsymbol{e} + [B(\boldsymbol{e})]\boldsymbol{u} \tag{13.209}$$

Note that the plant matrix is time dependent due to $\boldsymbol{e}_d$, and the control influence matrix is state dependent. Because $[A]$ only depends on the mean a, e and i parameters, the 6×6 matrix $[A^*]$ has block structure:

$$[A^*(\boldsymbol{e}_d)] = \begin{bmatrix} 0_{3\times3} & 0_{3\times3} \\ A^*_{21} & 0_{3\times3} \end{bmatrix} \tag{13.210}$$

Substituting Eq. (13.208) back into the control law in Eq. (13.207), we approximate $\boldsymbol{u}$ as

$$\boldsymbol{u} \simeq -\left([B]^T[B]\right)^{-1}[B]^T\left([A^*(\boldsymbol{e}_d)] + [P]\right)\delta\boldsymbol{e} \tag{13.211}$$

Taking the partial derivatives of Eq. (13.197) with respect to $\boldsymbol{e}$, the submatrix $[A^*_{21}]$ is found to be

$$[A^*_{21}] = \begin{bmatrix} \frac{21\epsilon}{4a}\cos i & -6\epsilon\frac{e}{\eta^2}\cos i & \frac{3}{2}\epsilon\sin i \\ -\frac{21\epsilon}{16a}(5\cos^2 i - 1) & 3\epsilon\frac{e}{\eta^2}(5\cos^2 i - 1) & -\frac{15}{4}\epsilon\sin(2i) \\ -\frac{3}{2a}\left[n + \frac{7}{8}\epsilon(3\cos^2 i - 1)\right] & 3\epsilon\frac{e}{\eta^2}(3\cos^2 i - 1) & -\frac{9}{4}\epsilon\sin(2i) \end{bmatrix} \tag{13.212}$$

with the small parameter ϵ being defined as

$$\epsilon = J_2\left(\frac{r_{eq}}{p}\right)^2 n \tag{13.213}$$

An approximate analysis of the $[A^*_{21}]$ matrix entry magnitudes in terms of metric units yields the following. Because both J_2 and n are of order 10^{-3}, and r_{eq}/p is of order 1, the parameter ϵ is of order 10^{-6}. Most entries of $[A^*_{21}]$ contain ϵ multiplied by either e, a small quantity of order 10^{-2} or smaller, or divided by a, a large quantity of order 10^3. These entries are then at least of order 10^{-8} or smaller. The largest entries contain only ϵ or n/a. Either one is of order 10^{-6}. Therefore, studying Eq. (13.211) shows that unless the feedback gain matrix $[P]$ is of order 10^{-5} or less, the $[A^*]$ matrix has a negligible influence on the control performance. In fact, if the feedback gain matrix $[P]$ is at least two or more magnitudes larger than the $[A^*]$ matrix, the $([A(\hat{\boldsymbol{e}}_d)] - [A(\boldsymbol{e}_d)])$ term can be dropped from the control law without any apparent performance loss.

Dropping the $([A(\hat{\boldsymbol{e}}_d)] - [A(\boldsymbol{e}_d)])$ term from the mean element feedback law, we are able to provide a rigorous stability proof for the special case where the feedback gain matrix $[P]$ is simply a positive constant scalar P.

$$\boldsymbol{u} = -P\left([B]^T[B]\right)^{-1}[B]^T\delta\boldsymbol{e} \tag{13.214}$$

Note that restraining the feedback gain to be a constant scalar would have a negative impact on the control performance, since it is no longer possible to use the celestial mechanics insight to guide when certain orbit elements should be corrected. But, this proof does provide some more analytical confidence in the control law and could be of use when only certain orbit elements have to be controlled.[23] We define a modified time dependent Lyapunov function $V(\delta\boldsymbol{e}, t)$ as[23]

$$V(\delta\boldsymbol{e}, t) = \frac{1}{2}(\alpha_1 + \alpha_2 e^{-\alpha_3 t})\delta\boldsymbol{e}^T\delta\boldsymbol{e} \tag{13.215}$$

with both α_1 and α_2 being positive constants. Due to having $V(\delta\boldsymbol{e}, t)$ explicitly depend on time, further steps are required in proving that V is positive or negative definite. This Lyapunov function $V(\delta\boldsymbol{e}, t)$ is positive definite since there exists a time-invariant positive definite $V_0(\delta\boldsymbol{e})$ such that[24]

$$V(\delta\boldsymbol{e}, t) \geq \frac{\alpha_1}{2}\delta\boldsymbol{e}^T\delta\boldsymbol{e} = V_0(\delta\boldsymbol{e}) \tag{13.216}$$

Further, this V is decrescent since there exists a time-invariant positive definite function $V_1(\delta\boldsymbol{e})$ such that[24]

$$V(\delta\boldsymbol{e}, t) \leq \frac{\alpha_1 + \alpha_2}{2}\delta\boldsymbol{e}^T\delta\boldsymbol{e} = V_1(\delta\boldsymbol{e}) \tag{13.217}$$

Since $V(\delta\boldsymbol{e}, t) \to \infty$ if $|\delta\boldsymbol{e}| \to \infty$ it is also radially unbounded. Taking the derivative of Eq. (13.215) and making use of $\delta\dot{\boldsymbol{e}} = [B(\boldsymbol{e})]\boldsymbol{u}$ and Eq. (13.214), we find

$$\begin{aligned}\dot{V}(\delta\boldsymbol{e}, t) = &- \alpha_2\alpha_3 e^{-\alpha_3 t}\delta\boldsymbol{e}^T\delta\boldsymbol{e} \\ &- (\alpha_1 + \alpha_2 e^{-\alpha_3 t})P\delta\boldsymbol{e}^T[B]([B]^T[B])^{-1}[B]^T\delta\boldsymbol{e}\end{aligned} \tag{13.218}$$

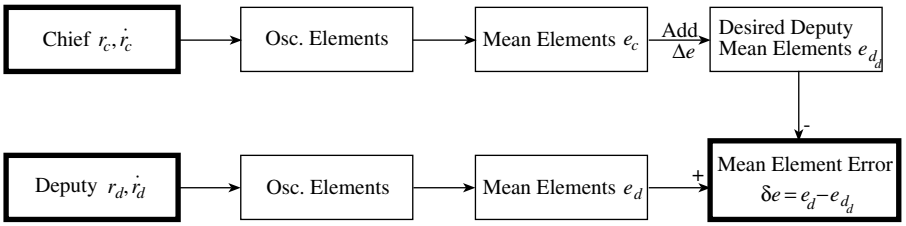


Figure 13.14: Mean Element Control Illustration

This time dependent function is negative definite since there exists a time-invariant negative definite function

$$\dot{V}_0(\delta \boldsymbol{e}) = -\alpha_2 \alpha_3 \delta \boldsymbol{e}^T \delta \boldsymbol{e} - \alpha_1 P \delta \boldsymbol{e}^T [B]([B]^T[B])^{-1}[B]^T \delta \boldsymbol{e} \tag{13.219}$$

such that[24]

$$\dot{V}(\delta \boldsymbol{e}, t) \leq \dot{V}_0(\delta \boldsymbol{e}) \tag{13.220}$$

Since $V(\delta \boldsymbol{e}, t)$ is positive definite, decrescent and radially unbounded, while $\dot{V}(\delta \boldsymbol{e}, t)$ is negative definite, the simplified control in Eq. (13.214) provides global uniform asymptotic stability under the assumption that the feedback gain P is large enough such that the term $([A(\hat{\boldsymbol{e}}_d)] - [A(\boldsymbol{e}_d)])$ can be dropped. Again, it should be noted thought that only having a scalar feedback gain P may provide un-acceptable fuel cost since the feedback control law may try to compensate for orbit element errors when it is very inefficient to do so.

A schematic layout of the mean element control is shown in Figure 13.14. Inertial position and velocity vectors are assumed to be available for both the chief and deputy satellite. After transforming both sets of vectors into corresponding mean orbit element vectors, the desired deputy mean elements are computed through a specified orbit element difference $\Delta \boldsymbol{e}$ relative to the chief satellite. The tracking error $\delta \boldsymbol{e}$ is then computed as the difference between the desired and actual deputy mean orbit elements. As mentioned earlier, the first order transformation used in this study to transform back and forth between osculating and mean orbit elements is not perfect. Taking a cartesian position and velocity vector, transforming first to mean elements and then back to cartesian coordinates can result in position differences in the dozens of meters. This is not a problem for typical orbit applications. However, for spacecraft formation flying, where the satellite relative orbit is to be controlled very precisely, this transformation error is significant. In the control strategy presented in Figure 13.14, both sets of mean elements are computed from inertial cartesian coordinates. While there is a minor error associated with this transformation, the error will be roughly the same for both sets of coordinates since the cartesian coordinates are relatively close to begin with. Because a *difference* in mean orbit elements is fed back, these transformation errors are found to approximately cancel each other and do not degrade the controller performance. Of course,

this transformation error could be further reduced by expanding the analytic orbit solution to higher order. However, even here it is beneficial to always deal with differences in orbit elements to achieve higher numerical accuracy.

13.7.2 Cartesian Coordinate Continuous Feedback Control Law

Traditional feedback laws depend on cartesian position and velocity error vector measurements. A nonlinear cartesian coordinate feedback law is presented which illustrates the steps necessary to track a prescribed relative orbit expressed in terms of mean orbit element differences. A related nonlinear feedback law is presented in Ref. 25.

The inertial equations of motion of the chief satellite $\boldsymbol{r}_1$ and deputy satellite $\boldsymbol{r}_2$ are

$$\ddot{\boldsymbol{r}}_c = \boldsymbol{f}(\boldsymbol{r}_c) \tag{13.221}$$

$$\ddot{\boldsymbol{r}}_d = \boldsymbol{f}(\boldsymbol{r}_d) + \boldsymbol{u} \tag{13.222}$$

where the chief satellite is assumed to be in a free, uncontrolled orbit and only the deputy satellite is being controlled to maintain the desired relative orbit. The vector function $\boldsymbol{f}(\boldsymbol{r})$ contains the gravitational acceleration. Expressing the inertial position vector in terms of inertial components $\boldsymbol{r} = (x, y, z)$ and including the J_2 perturbation, this function is defined as

$$\boldsymbol{f}(\boldsymbol{r}) = -\frac{\mu}{r^3}\left[\boldsymbol{r} - J_2\frac{3}{2}\left(\frac{r_{eq}}{r}\right)^2\begin{pmatrix}5x\left(\frac{z}{r}\right)^2 - x\\ 5y\left(\frac{z}{r}\right)^2 - y\\ 5z\left(\frac{z}{r}\right)^2 - 3z\end{pmatrix}\right] \tag{13.223}$$

where r is the scalar orbit radius. Let $\boldsymbol{r}_{d_d}$ be the desired inertial position vector of the deputy satellite for a J_2 invariant relative orbit. The position tracking error $\delta\boldsymbol{r}$ is then defined as

$$\delta\boldsymbol{r} = \boldsymbol{r}_d - \boldsymbol{r}_{d_d} \tag{13.224}$$

Using this error vector and its derivative, the positive definite Lyapunov function V is defined as

$$V(\delta\boldsymbol{r}, \delta\dot{\boldsymbol{r}}) = \frac{1}{2}\delta\dot{\boldsymbol{r}}^T\delta\dot{\boldsymbol{r}} + \frac{1}{2}\delta\boldsymbol{r}^T[K_1]\delta\boldsymbol{r} \tag{13.225}$$

where $[K_1]$ is a positive definite 3×3 position feedback gain matrix. Taking the derivative of V we find

$$\dot{V} = \delta\dot{\boldsymbol{r}}^T\left(\ddot{\boldsymbol{r}}_d - \ddot{\boldsymbol{r}}_{d_d} + [K_1]\delta\boldsymbol{r}\right) \tag{13.226}$$

Substituting Eq. (13.222) and making use of the fact that the desired relative orbit is J_2 invariant (i.e. control free), the Lyapunov rate is written as

$$\dot{V} = \delta\dot{\boldsymbol{r}}^T\left(\boldsymbol{f}(\boldsymbol{r}_d) - \boldsymbol{f}(\boldsymbol{r}_{d_d}) + \boldsymbol{u} + [K_1]\delta\boldsymbol{r}\right) \tag{13.227}$$

Enforcing $\dot{V}$ to be equal to the negative definite quantity

$$\dot{V} = -\delta\dot{\boldsymbol{r}}^T[K_2]\delta\dot{\boldsymbol{r}} \tag{13.228}$$

where $[K_2]$ is a positive definite 3×3 velocity feedback gain matrix, the asymptotically stabilizing control law $\boldsymbol{u}$ is found to be

$$\boldsymbol{u} = -\left(\boldsymbol{f}(\boldsymbol{r}_d) - \boldsymbol{f}(\boldsymbol{r}_{d_d})\right) - [K_1]\delta\boldsymbol{r} - [K_2]\delta\dot{\boldsymbol{r}} \tag{13.229}$$

Note that this control law controls the inertial deputy orbit directly. The orbit errors $\delta\boldsymbol{r}$ are only the difference between the desired and actual inertial deputy orbit. This control law can be used for formation flying control by having the desired deputy position $\boldsymbol{r}_{2_d}$ be defined relative to the chief orbit. As is, the feedback control law in (13.229) could be also be used to maintain the inertial orbit of a single satellite. The asymptotic stability property of this control law can be verified by checking the higher order derivatives of V on the set where $\dot{V}$ is zero (i.e. evaluated at $\delta\dot{\boldsymbol{r}} = 0$).[26] The first non-zero higher derivative of V on this set is found to be the third derivative

$$\dddot{V}(\delta\dot{\boldsymbol{r}} = 0) = -\delta\boldsymbol{r}^T[K_1]^T[K_2][K_1]\delta\boldsymbol{r} < 0 \tag{13.230}$$

which is negative definite in $\delta\boldsymbol{r}$. Thus the order of the first non-zero derivative is odd and the control law is asymptotically stabilizing.

Where the mean orbit element feedback law feeds back a difference in the natural orbit element rates, the cartesian coordinate feedback law in Eq. (13.229) feeds back a difference in gravitational accelerations. Linearizing this difference about the desired motion $\boldsymbol{r}_{d_d}(t)$ we find

$$\boldsymbol{f}(\boldsymbol{r}_d) - \boldsymbol{f}(\boldsymbol{r}_{d_d}) \simeq \left[\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{r}}\right]\Bigg|_{\boldsymbol{r}_{d_d}} \delta\boldsymbol{r} = [F(\boldsymbol{r}_{d_d})]\delta\boldsymbol{r} \tag{13.231}$$

Using Eq. (13.231), the closed-loop dynamics are now written in the linear form as

$$\delta\ddot{\boldsymbol{r}} \simeq [F(\boldsymbol{r}_{d_d})]\delta\boldsymbol{r} + \boldsymbol{u} \tag{13.232}$$

and the control law is linearized as

$$\boldsymbol{u} \simeq -([F(\boldsymbol{r}_{d_d})] + [K_1])\delta\boldsymbol{r} - [K_2]\delta\dot{\boldsymbol{r}} \tag{13.233}$$

The matrix $[F]$ can be written as $[F] = [F_{Kepler}] + [F_{J_2}]$ where $[F_{Kepler}]$ is the term due to the inverse square gravitational attraction and $[F_{J_2}]$ is the term due to the J_2 perturbation. Doing a similar dimensional study of $[F_{Kepler}]$ and $[F_{J_2}]$, as for $[A^*]$ earlier, the matrix $[F_{Kepler}]$ is found to be of order μ/r and $[F_{J_2}]$ of order $J_2\mu/r^3$. Since both J_2 and $1/r$ are roughly 10^{-3}, this means that $[F_{J_2}]$ is on the order of 10^{-9} smaller than $[F_{Kepler}]$. This means that excluding the J_2 term in the $\boldsymbol{f}(\boldsymbol{r})$ calculation will have a negligible effect on the performance. Therefore the largest component of $[F]$ is of order $\mu/r = 10^1$

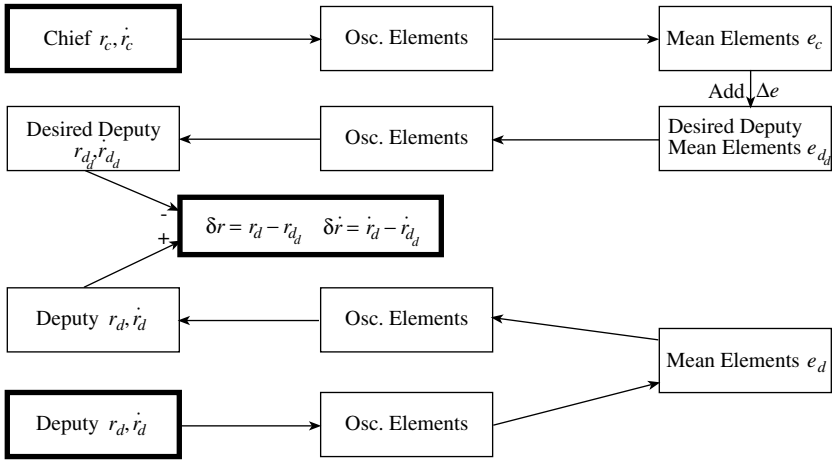


Figure 13.15: Tracking Error Computation Logic for Cartesian Coordinate Feedback Control

in metric units. For the cartesian feedback law, feeding back the difference in gravitational accelerations has a large influence on the performance. For example, if the gains are very small to allow the maneuver to take several orbit revolutions, then the control effort will still be large due to this gravitational acceleration difference term. This is in contrast to the mean orbit element feedback law where the maneuvers can easily be stretched over several orbit revolutions.

A critical detail in this cartesian coordinate feedback law is how to compute the desired deputy position and velocity vectors, because the relative orbit trajectory is described in terms of mean orbit element differences relative to the chief orbit. Figure 13.15 illustrates this process. After translating the chief cartesian coordinates into corresponding mean orbit elements, the desired deputy position and velocity vectors are computed by first adding the desired mean orbit element difference vector $\Delta \boldsymbol{e}$ and then transforming these desired elements back to cartesian space. However, if these desired inertial deputy states are differenced with the actual inertial deputy states, serious numerical difficulties may arise. The reason for this is the transformation error that occurs when mapping between osculating and mean orbit elements. The closed loop position errors will stop decaying once the accuracy of this transformation is reached. To avoid this limitation, we don't use the actual states of the deputy when computing the tracking error. Instead, we map these states first to mean orbit elements and then back to cartesian coordinates before differencing them with the desires states. With the difference between the chief and deputy position and velocity vectors being very small, the transformation error due to the forward and backward mapping will be essentially identical and cancel themselves when being differenced. This qualitative observation is consistent with our numerical experiments. The result is a nonlinear cartesian coordinate feedback law that

is able to establish the J_2 invariant orbit and overcome some of the limitations of having a first-order transformation between the osculating and mean orbit elements.

13.7.3 Impulsive Feedback Control Law

The previous two formation flying control law provide a continuous thrust vector to cancel any relative orbit errors. This section develops an impulsive control scheme to control the relative orbit. Instead of continuously controlling the relative orbit errors, the tracking errors will only be controlled at specific periods within the orbit. This impulsive feedback control law was born out of the quest to find a method to correct for the argument of perigee and mean anomaly drifts experienced by J_2 invariant orbits, while minimally impacting the remaining orbit elements. While the presented method is attractive to correct specific sets of orbit elements, it is also possible to use this method to correct for arbitrary relative orbit errors. Further, the results of this sections were used earlier in the development of control effort estimates for maintaining J_2 invariant orbits.

Gauss' variational equations are used again in this development to derive the required control law. Studying the $d\Omega/dt$ and di/dt expressions in Eq. (13.190), it is evident that the individual ascending node or inclination angles are adjusted best when the spacecraft passes through either the polar or the equatorial regions respectively. However, if both an inclination angle and nodal correction are to be performed, it is more fuel efficient to perform both corrections with one impulse only. Both elements are adjusted with an orbit normal impulsive Δv_h as shown in Eq. (13.190). The corresponding inclination angle and ascending node corrections are given by

$$\Delta i = \frac{r\cos\theta}{h}\Delta v_h \tag{13.234}$$

$$\Delta\Omega = \frac{r\sin\theta}{h\sin i}\Delta v_h \tag{13.235}$$

Dividing Eq. (13.235) by (13.234), the critical true latitude angle θ_c at which to perform this orbit normal thrusting maneuver is

$$\theta_c = \arctan\frac{\Delta\Omega\ \sin i}{\Delta i} \tag{13.236}$$

Squaring and summing Eqs. (13.234) and (13.235), the required Δv_h to perform the desired inclination correction Δi and ascending node correction $\Delta\Omega$ is

$$\Delta v_h = \frac{h}{r}\sqrt{\Delta i^2 + \Delta\Omega^2\sin i^2} \tag{13.237}$$

Note that applying this Δv_h only affects the orbit elements i, Ω and ω. This cross-coupling between the (i, Ω) correction and ω is the only coupling between osculating orbit element set corrections in this firing scheme. Note that while there always exists two possible critical true latitude angles θ_c from Eq. (13.236),

only the solution corresponding to a positive Δv_h is used in this control method. Thus (i, Ω) are only corrected at one point in the orbit.

Substituting the Δv_h in Eq. (13.235) into Eq. (13.190d), the $\Delta\Omega$ correction results in the following $\Delta\omega$ change:

$$\Delta\omega(\Delta v_h) = -\cos i \Delta\Omega \tag{13.238}$$

This secondary effect will be taken into account when specifying the impulse required to correct the argument of perigee.

The argument of perigee and the mean anomaly are also corrected together together as an orbit element pair, but with two impulsive maneuvers over one orbit. Each impulsive thrust is in the orbit radial direction only and is applied at both the orbit perigee and apogee. Let Δv_{r_p} be the radial impulse applied at perigee and Δv_{r_a} be the impulse at apogee. Computed over one orbit, and taking into account that an ascending node correction $\Delta\Omega$ could be occurring (which causes an additional change in ω), the Δv_{r_p} and Δv_{r_a} impulses cause the following osculating orbit element changes.

$$\Delta\omega = \frac{1}{he}(-p(\Delta v_{r_p} - \Delta v_{r_a}) - \Delta\Omega\cos i) \tag{13.239}$$

$$\Delta M = \frac{\eta}{he}\left((p - 2r_p e)\Delta v_{r_p} - (p + 2r_a e)\Delta v_{r_a}\right) \tag{13.240}$$

with $\eta = \sqrt{1 - e^2}$. To solve these two equations for the radial Δv's, the following identities are useful

$$p - 2r_p e = p\frac{1 - e}{1 + e} \tag{13.241a}$$

$$p - 2r_a e = p\frac{1 + e}{1 - e} \tag{13.241b}$$

along with $h/p = na/\eta$. Substituting these expressions into Eqs. (13.239) and (13.240) we find

$$\Delta v_{r_p} - \Delta v_{r_a} = -(\Delta\omega + \Delta\Omega\cos i)\frac{nae}{\eta} \tag{13.242}$$

$$(1 - e)^2\Delta v_{r_p} - (1 + e)^2\Delta v_{r_a} = nae\Delta M \tag{13.243}$$

Solving these two equations for the required radial impulses to achieve a desired $\Delta\omega$ and ΔM we find

$$\Delta v_{r_p} = -\frac{na}{4}\left(\frac{(1 + e)^2}{\eta}(\Delta\omega + \Delta\Omega\cos i) + \Delta M\right) \tag{13.244}$$

$$\Delta v_{r_a} = \frac{na}{4}\left(\frac{(1 - e)^2}{\eta}(\Delta\omega + \Delta\Omega\cos i) + \Delta M\right) \tag{13.245}$$

Note that if a $\Delta\Omega$ correction is performed during this orbit, then its effect is immediately taken into account in the above two equations.

The argument of perigee and mean anomaly corrections, provided by Eqs. (13.244) and (13.245), are convenient to compensate for the natural secular drift in these orbit elements that will occur with the J_2-invariant orbit presented in Reference 19. Only ω and M of the six orbit elements will not have an equal relative drift rate, but rather their sum will. This relative drift difference is not very large, but depending on the tolerances of the relative orbit it will have to be compensated for periodically. Further, the smaller the eccentricity of the orbit, the less effect the relative drift of ω and M will have on the orbit geometry. However, Eqs. (13.244) and (13.245) provide an impulsive control method which is able to directly readjust the argument of perigee and mean anomaly while minimally affecting the other osculating orbit elements.

The remaining two orbit elements to be corrected are the semi-major axis a and the eccentricity e. As is the case with the argument of perigee and mean anomaly corrections, the semi-major axis and eccentricity are adjusted together through two impulsive maneuvers over one orbit. However, these impulsive thrusts are fired in the tangential u_θ direction. One impulsive correction Δv_{θ_p} is fired at perigee and the other impulse Δv_{θ_a} is fired at apogee. With this firing sequence a and e are adjusted efficiently and without disturbing the other osculating orbit elements. From Eq. (13.190), the a and e corrections over one orbit are

$$\Delta a = \frac{2a^2}{h}\left(\frac{p}{r_p}\Delta v_{\theta_p} + \frac{p}{r_a}\Delta v_{\theta_a}\right) \tag{13.246}$$

$$\Delta e = \frac{1}{h}\Big((p + r_p + r_p e)\Delta v_{\theta_p} + (-p - r_a + r_a e)\Delta v_{\theta_a}\Big) \tag{13.247}$$

Note that in deriving Eqs. (13.246) and (13.247) it is assumed that the orbit corrections Δa and Δe are relatively small. Otherwise a and e could not be held constant during the two maneuvers. To solve these two equations for the tangential Δv's, the following identities are used.

$$p + r_p + r_p e = 2p \tag{13.248}$$

$$-p - r_a + r_a e = -2p \tag{13.249}$$

Eqs. (13.246) and (13.247) are now rewritten as

$$(1+e)\Delta v_{\theta_p} + (1-e)\Delta v_{\theta_a} = \frac{h^2}{2a^2}\Delta a \tag{13.250}$$

$$\Delta v_{\theta_p} - \Delta v_{\theta_a} = \frac{h}{2p}\Delta e \tag{13.251}$$

Using $h/a = na\eta$, with $\eta = \sqrt{1-e^2}$, the required tangential impulses are found to be

$$\Delta v_{\theta_p} = \frac{na\eta}{4}\left(\frac{\Delta a}{a} + \frac{\Delta e}{1+e}\right) \tag{13.252}$$

$$\Delta v_{\theta_a} = \frac{na\eta}{4}\left(\frac{\Delta a}{a} - \frac{\Delta e}{1-e}\right) \tag{13.253}$$

Note that in both the (ω, M) and (a, e) corrections, the sequence of impulsive maneuvers over an orbit is irrelevant. The first maneuver may occur at either perigee or apogee.

To implement these impulsive Δv's, the mean orbit element errors are established at some arbitrary point in the orbit, and are then held constant during the orbit while appropriate Δv's are applied as discussed earlier. This impulsive firing scheme assumes that all the mean orbit element errors will remain constant over an orbit. If the a, e and i elements do not satisfy the J_2 invariant conditions, then Ω, ω and M will experience some J_2 induced secular relative drift. However, this drift is relatively small over an orbit and can be ignored. The impulsive feedback control will correct, or at least substantially reduce, any remaining mean orbit element errors during the following orbit. The exception is if the deputy semi-major axis is substantially different from that of the chief. In this case the different orbit periods will cause the mean anomaly to exhibit substantial relative drift over one orbit. In this case it cannot be assumed that ΔM is constant over an orbit. Thus, the (ω, M) corrections do not begin until the second orbit. Doing this allows the a, e and i variables to be corrected during the first obit, which will set the orbit periods equal between deputy and chief satellite. During further orbits, any remaining relative mean anomaly errors will remain constant over an orbit. If the (ω, M) corrections are applied during the first orbit with a large semi-major axis error present, then the impulsive feedback control law still corrects the relative orbit. However, the fuel cost typically increases since incorrect (ω, M) corrections are performed during the first orbit.

Since it is advantageous to describe the relative orbit in terms of orbit element differences of the deputy satellite relative to the chief satellite, this impulsive firing sequence is a convenient method to correct orbit errors from the desired orbit element differences. If only one or two elements are to be adjusted, then this control solution is essentially optimal. If several orbit elements are to be corrected, then preliminary studies have shown this method to still yield a near-optimal solution with a fuel cost increase of only a few percent over the multi-impulse optimal solution. The advantage of this method is that through its simplicity and low computational overhead, it lends itself well to be implemented in an autonomous manner. Little ground support would be required for a cluster of spacecraft to maintain their formation as long as they are able to sense their inertial orbits themselves. This could be achieved through GPS measurements or direct line of sight measurements between the various satellites. Feeding back mean orbit element errors has the benefit that any short period oscillations are ignored.

Further, it is convenient to be able to adjust only certain orbit elements, leaving the remaining elements virtually untouched. For relative orbits designed using the J_2 orbit element constraints, the resulting relative orbit will be J_2 invariant in an angular sense. This means that the neighboring orbits will have equal nodal and mean latitude drift rates. However, the argument of perigee and mean anomaly will still drift apart at equal and opposite rates. The consequence of this drift is that the relative orbit will go through cycles of symmetrically

growing and shrinking as the chief satellite completes one orbit. This effect is more noticeable for satellite clusters with larger eccentricities. For a cluster with nominally zero eccentricity, having the argument of perigee and mean anomaly grow apart at equal and opposite rates has no affect on the overall relative orbit geometry. Further, this impulsive firing scheme could also be used as the initial conditions for an optimizer solving for the true minimum fuel orbit correction. Often indirect optimizing methods are sensitive to initial conditions, and the presented impulsive feedback law could provide reasonable initial guess as to the structure of the optimal control solution.

13.7.4 Hybrid Feedback Control Law

The use of Eq. (13.40) is investigated here to create a hybrid continuous feedback control law in terms of Cartesian Hill frame coordinates, while describing the desired relative orbit geometry through a desired set of orbit element differences $\delta\boldsymbol{e}^*$. Any desired states are denoted here with a superscript asterisk. The advantage of this type of hybrid control law is that the actual relative orbit is expressed in terms of coordinates in which it would actually be measured (i.e. the chief frame local Hill coordinates), while the desired relative orbit is conveniently expressed as a set of orbit element differences.

Let $\boldsymbol{x} = (x, y, z)^T$ be the deputy position vector and $\boldsymbol{v} = (\dot{x}, \dot{y}, \dot{z})^T$ be the deputy velocity vector expressed in the chief Hill frame coordinates. The general linearized relative equations of motion for a Keplerian system, given in Eq. (13.18), are expressed here as[3]

$$\dot{\boldsymbol{x}} = \boldsymbol{v} \tag{13.254}$$

$$\begin{aligned}\dot{\boldsymbol{v}} = &\underbrace{\begin{bmatrix} 2\frac{\mu}{R^3} + \dot{\theta}^2 & \ddot{\theta} & 0 \\ -\ddot{\theta} & \dot{\theta}^2 - \frac{\mu}{R^3} & 0 \\ 0 & 0 & -\frac{\mu}{R^3} \end{bmatrix}}_{[A_1]} \boldsymbol{x} \\ &+ \underbrace{\begin{bmatrix} 0 & 2\dot{\theta} & 0 \\ -2\dot{\theta} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{[A_2]} \boldsymbol{v} + \underbrace{\begin{pmatrix} u_x \\ u_y \\ u_z \end{pmatrix}}_{\boldsymbol{u}}\end{aligned} \tag{13.255}$$

with θ being the true latitude. These relative equations of motion are valid for both circular and elliptic chief orbits. The true latitude acceleration is computed through

$$\ddot{\theta} = -2\frac{\mu}{R^3}(q_1 \sin\theta - q_2 \cos\theta) \tag{13.256}$$

with q_1 and q_2 being defined in Eqs. (13.42) and (13.43).

Let us define the relative orbit tracking errors as

$$\Delta\boldsymbol{x} = \boldsymbol{x} - \boldsymbol{x}^* \tag{13.257}$$

$$\Delta\boldsymbol{v} = \boldsymbol{v} - \boldsymbol{v}^* \tag{13.258}$$

with the desired position and velocity vectors computed using

$$\boldsymbol{X}^* = \begin{pmatrix} \boldsymbol{x}^* \\ \boldsymbol{v}^* \end{pmatrix} = [A(\boldsymbol{e})]\delta\boldsymbol{e}^* \tag{13.259}$$

Note that if the desired orbit element differences call for a fixed mean anomaly difference, as is done in References 19, 20 and 21, then the vector $\delta\boldsymbol{e}^*$ is not constant, but rather $\delta\theta$ must be computed at each instant by solving Kepler's equation. Further, note that $\Delta\dot{\boldsymbol{x}} = \Delta\boldsymbol{v}$.

Let us define the control law $\boldsymbol{u}$ as

$$\boldsymbol{u} = \dot{\boldsymbol{v}}^* - A_1\boldsymbol{x} - A_2\boldsymbol{v} - [K]\Delta\boldsymbol{x} - [P]\Delta\boldsymbol{v} \tag{13.260}$$

with $[K]$ and $[P]$ being positive definite matrices. To prove that $\boldsymbol{u}$ is asymptotically stabilizing, a positive definite Lyapunov function V is defined as

$$V(\Delta\boldsymbol{x}, \Delta\boldsymbol{y}) = \frac{1}{2}\Delta\boldsymbol{v}^T\Delta\boldsymbol{v} + \frac{1}{2}\Delta\boldsymbol{x}^T[K]\Delta\boldsymbol{x} \tag{13.261}$$

Substituting Eqs. (13.255) and (13.258), the derivative of V along the state trajectory must be negative semi-negative

$$\dot{V} = \Delta\boldsymbol{v}^T(\Delta\dot{\boldsymbol{v}} + [K]\Delta\boldsymbol{x}) = -\Delta\boldsymbol{v}^T[P]\Delta\boldsymbol{v} \tag{13.262}$$

which guarantees that $\boldsymbol{u}$ is globally stabilizing. To prove that the control law is also asymptotically stabilizing, the higher order time derivatives of V are investigated. The second derivative of V is zero when evaluated on the set where $\dot{V} = 0$. The third derivative

$$\dddot{V}(\Delta\boldsymbol{v} = 0) = -2\Delta\boldsymbol{x}^T[K][P][K]\Delta\boldsymbol{x} \tag{13.263}$$

is negative definite in the state vector $\Delta\boldsymbol{x}$. Since this first non-zero derivative is an odd derivative, the control $\boldsymbol{u}$ is asymptotically stabilizing.[26]

Note that $\dot{\boldsymbol{v}}^* - [A_1]\boldsymbol{x}^* - [A_2]\boldsymbol{v}^*$ is zero if the desired relative motion is a natural solution to the linearized equations of motion shown in Eq. (13.255). Assuming that our chosen $\dot{\boldsymbol{v}}^*$ abides by

$$\dot{\boldsymbol{v}}^* = [A_1]\boldsymbol{x}^* + [A_2]\boldsymbol{v}^* \tag{13.264}$$

the control law $\boldsymbol{u}$ is written as

$$\boldsymbol{u} = -\begin{bmatrix} A_1 + K & A_2 + P \end{bmatrix}\left(\begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{v} \end{pmatrix} - [A(\boldsymbol{e})]\delta\boldsymbol{e}^*\right) \tag{13.265}$$

Note, however, that the desired relative motion may not necessarily be a natural solution. The control law in Eq. (13.260) is also valid for forced relative orbits. Studying this form of control law in Eq. (13.265), the hybrid nature of $\boldsymbol{u}$ is evident in that the desired relative orbit is prescribed through a set of orbit element differences, while the actual motion is expressed in terms of the chief Hill frame Cartesian components. The advantage here is that we are able to

express the actual and desired relative motion in coordinates that best suit their task. The continuous feedback control law in Eq. (13.229), in contrast, feeds back tracking errors in terms of the inertial deputy position vector. The hybrid control law in Eq. (13.265) takes advantage of the fact that the deputy satellite position is controlled relative to the chief position by expressing the tracking errors in terms of the relative Hill frame coordinates.

Since the $[A_2]$ matrix is skew-symmetric, it could be dropped from the control expression in Eq. (13.265). The Lyapunov-based stability proof remains the same and asymptotic stability is still guaranteed. However, computing $\dot{V}$ the term $\Delta\boldsymbol{v}^T[A_2]\Delta\boldsymbol{v}$ is dropped since it is always zero. The modified control expression is then

$$\boldsymbol{u} = -\begin{bmatrix} A_1 + K & P \end{bmatrix}\left(\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{v}\end{pmatrix} - [A(\boldsymbol{e})]\delta\boldsymbol{e}^*\right) \tag{13.266}$$

This control would no longer feedback-linearize the closed-loop dynamics, but it still guarantees asymptotic stability.

Note that while the control expression in Eq. (13.265) takes advantage of the linear mapping $[A(\boldsymbol{e})]$ between orbit element differences and their corresponding Hill Cartesian coordinates, the control expression in Eq. (13.260) does not rely on this mapping. In fact, the relative orbit tracking errors $\Delta\boldsymbol{x}$ and $\Delta\boldsymbol{v}$ could be computed using the complete nonlinear mapping between orbit elements and local Cartesian coordinates. Further, it is possible to incorporate the J_2 effect here by using Brouwer's theory to compute the relative orbit errors in mean element space and then map the error vector back to osculating space for control purposes.

Problems

13.1 Write a program that will display the orbit of satellite as seen by the rotating Hill reference frame of another satellite.

13.2 Let the chief orbit be determined through the orbit elements a = 7500km, e = 0.01, i = 45 degrees, Ω = 0.0 degrees, ω = 30 degrees and M_0 = 0 degrees. The deputy orbit has the same orbit elements except for the i = 45.1 degrees and ω = 29 degrees.

a) Use the nonlinear relative equations of motion in Eq. (13.12) and plot the relative orbit in the Hill frame.

b) Compute the relative orbit by using the inertial equations of motion in Eq. (13.11) and computing first the inertial deputy and chief orbits and then differencing them. Plot the result in the rotating Hill frame and compare to the previous answer.

13.3 Using the chief and deputy orbit elements in Problem 13.2, compute the relative orbit using both the nonlinear relative equations of motion in Eq. (13.18) and the linear CW equations in Eq. (13.19) for different chief orbit eccentricities. Start with a zero eccentricity and increase it until a critical value is found where the CW relative orbit calculation is off by 1 km.

13.4 ♣ Starting with the nonlinear relative equations of motion in Eq. (13.17), derive the non-dimensional relative equations of motion shown in Eq. (13.23). Include the derivation of the intermediate results shown in Eq. (13.22).

13.5 Create a program that will perform both the forward and inverse mapping between the a relative orbit Cartesian position vector $\boldsymbol{X}$ and corresponding orbit element difference vector $\delta\boldsymbol{e}$ shown in Eq. (13.40). Verify that the $[A(\boldsymbol{e}_c)][A(\boldsymbol{e}_c)]^{-1}$ does yield the identity matrix.

13.6 ♣ Derive the non-dimensional Cartesian rate computation in Eq. (13.75). Show all intermediate steps.

13.7 Use the orbit constraint condition in Eq. (13.85) to generate initial conditions that will yield a bounded relative orbit. Assume that the chief orbit is given by the orbit elements shown in Problem 13.2. Plot the resulting relative orbit in the rotating chief Hill frame.

13.8 Verify the two J_2-invariant orbit element constraints in Eqs. (13.155) and (13.156). Start with the J_2-invariant relative orbit definition in Eqs. (13.146) and (13.147) and show all intermediate steps.

13.9 ♣ Create a numerical simulation to compute the relative orbit shown in Example 13.4 by including the J_2 through J_5 perturbations. This program should compute the inertial orbit of both the deputy and chief satellite and then map the relative orbit into the rotating chief Hill reference frame. Use the mapping shown in Appendix G to map between the mean and osculating orbit elements.

a) Show how much the relative orbit still drifts despite the J_2-invariant conditions.

b) Of the J_2-J_5 gravitational perturbations, which one is the main cause for this drift.

c) Show how fast the argument of perigee and mean anomaly are drifting away from their initial values.

Bibliography

[1] Carter, T. E., "State Transition Matrix for Terminal Rendezvous Studies: Brief Survey and New Example," *Journal of Guidance, Navigation and Control*, Vol. 31, No. 1, 1998, pp. 148–155.

[2] Hill, G., "Researches in the Lunar Theory," *American Journal of Mathematics*, Vol. 1, 1878, pp. 5–26.

[3] Melton, R. G., "Time-Explicit Representation of Relative Motion Between Elliptical Orbits," *Journal of Guidance, Control, and Dynamics*, Vol. 23, No. 4, 2000, pp. 604–610.

[4] Clohessy, W. H. and Wiltshire, R. S., "Terminal Guidance System for Satellite Rendezvous," *Journal of the Aerospace Sciences*, Vol. 27, No. 9, Sept. 1960, pp. 653–658.

[5] Schaub, H. and Alfriend, K. T., "Hybrid Cartesian and Orbit Element Feedback Law for Formation Flying Spacecraft," *Journal of Guidance, Control and Dynamics*, 2001, submitted for publication.

[6] Alfriend, K. T., Schaub, H., and Gim, D.-W., "Gravitational Perturbations, Nonlinearity and Circular Orbit Assumption Effect on Formation Flying Control Strategies," *AAS Guidance and Control Conference*, Breckenridge, CO, February 2000, Paper No. AAS 00-012.

[7] Schaub, H. and Alfriend, K. T., "Hybrid Cartesian and Orbit Element Feedback Law for Formation Flying Spacecraft," *AIAA Guidance, Navigation and Control Conference*, Denver, CO, Aug. 2000, Paper No. 2000-4131.

[8] DeVries, J. P., "Elliptic Elements in Terms of Small Increments of Position and Velocity Components," *AIAA Journal*, Vol. 1, No. 9, Nov. 1963, pp. 2626–2629.

[9] Garrison, J. L., Gardner, T. G., and Axelrad, P., "Relative Motion in Highly Elliptic Orbits," *AAS/AIAA Space Flight Mechanics Meeting*, Albuquerque, NM, Feb. 1995, Paper No. 95-194.

[10] Inalhan, G. and How, J. P., "Relative Dynamics & Control of Spacecraft Formations in Eccentric Orbits," *AIAA Guidance, Navigation and Control Conference and Exhibit*, Denver, CO, August 2000, AIAA 2000-4433.

[11] Tschauner, J. and Hempel, P., "Rendezvous zu einem in Elliptischer Bahn Umlaufenden Ziel," *Astronautica Acta*, Vol. 11, 1965, pp. 104–109.

[12] Gim, D.-W. and Alfriend, K. T., "The State Transition Matrix of Relative Motion for the Perturbed Non-Circular Reference Orbit," *AAS/AIAA Space Flight Mechanics Meeting*, Santa Barbara, CA, Feb. 2001, Paper No. 01-222.

[13] Beyer, W. H., *Standard Mathematical Tables*, CRC Press, Inc., West Palm Beach, Fl., 1974.

[14] Hughes, S. P. and Mailhe, L. M., "A Preliminary Formation Flying Orbit Dynamics Analysis for Leonardo-BRDF," *IEEE Aerospace Conference*, Big Sky, Montana, March 11–17 2001.

[15] Hughes, S. P. and Hall, C. D., "Optimal Configurations of Rotating Spacecraft Formations," *Journal of the Astronautical Sciences*, Vol. 48, No. 2 and 3, April-Sept. 2000, pp. 225–247.

[16] Chichka, D. F., "Dynamics of Clustered Satellites via Orbital Elements," *AAS/AIAA Astrodynamics Specialist Conference*, Girdwood, Alaska, August 1999, Paper No. AAS 99-309.

[17] Brouwer, D., "Solution of the Problem of Artificial Satellite Theory Without Drag," *The Astronomical Journal*, Vol. 64, No. 1274, 1959, pp. 378–397.

[18] Lyddane, R. H., "Small Eccentricities or Inclinations in the Brouwer Theory of the Artificial Satellite," *The Astronomical Journal*, Vol. 68, No. 8, October 1963, pp. 555–558.

[19] Schaub, H. and Alfriend, K. T., "J_2 Invariant Reference Orbits for Spacecraft Formations," *Celestial Mechanics and Dynamical Astronomy*, Vol. 79, 2001, pp. 77–95.

[20] Schaub, H., Vadali, S. R., and Alfriend, K. T., "Spacecraft Formation Flying Control Using Mean Orbit Elements," *Journal of the Astronautical Sciences*, Vol. 48, No. 1, 2000, pp. 69–87.

[21] Schaub, H. and Alfriend, K. T., "Impulsive Feedback Control to Establish Specific Mean Orbit Elements of Spacecraft Formations," *Journal of Guidance, Control and Dynamics*, Vol. 24, No. 4, July–Aug. 2001, pp. 739–745.

[22] Battin, R. H., *An Introduction to the Mathematics and Methods of Astrodynamics*, AIAA Education Series, New York, 1987.

[23] Tan, Z., Bainum, P. M., and Strong, A., "The Implementation of Maintaining Constant Distance Between Satellites in Elliptic Orbits," *AAS Spaceflight Mechanics Meeting*, Clearwater, Florida, Jan. 2000, Paper No. 00-141.

[24] Slotine, J. E. and Li, W., *Applied Nonlinear Control*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1991.

[25] Queiroz, M. S. D., Kapila, V., and Yan, Q., "Nonlinear Control of Multiple Spacecraft Formation Flying," *Proceedings of AIAA Guidance, Navigation, and Control Conference*, Portland, OR, Aug. 1999, Paper No. AIAA 99–4270.

[26] Mukherjee, R. and Chen, D., "Asymptotic Stability Theorem for Autonomous Systems," *Journal of Guidance, Control, and Dynamics*, Vol. 16, Sept.–Oct. 1993, pp. 961–963.

Appendix A

Transport Theorem Derivation Using Linear Algebra

Previously the kinematic differential equation of the direction cosine matrix $[C]$ and the *Transport Theorem* were derived using vector algebra results. This appendix will derive the same results using linear algebra arguments. While not as easy to visualize as the vector algebra results, the advantage here is that the derivation illustrates that these results hold for *any N-dimensional space.*

Any NxN orthogonal matrix $[C]$ must satisfy the constraints

$$[C]^T[C] = I_{N\times N} \tag{A.267}$$

$$[C][C]^T = I_{N\times N} \tag{A.268}$$

where $I_{N\times N}$ is an NxN identity matrix. Taking the derivative of Eq. (A.268) we find that

$$[\dot{C}][C]^T + [C][\dot{C}]^T = 0_{N\times N} \tag{A.269}$$

Note that taking the derivative of a matrix here only involves a series of scalar derivatives when taking the time derivatives of the various matrix elements. Contrary to the vector algebra developments, we are not concerned with different reference frames here. Eq. (A.269) can be rearranged to the form

$$[\dot{C}][C]^T = -[C][\dot{C}]^T = -\left([\dot{C}][C]^T\right)^T = [Q] \tag{A.270}$$

where, by definition, the "angular-velocity-like" matrix $[Q]$ must be skew-symmetric. Therefore the derivative of aNxN orthogonal matrix can be written in the general form

$$[\dot{C}] = [Q][C] \tag{A.271}$$

To illustrate that for the rigid body dynamics case $[Q] = -[\tilde{\boldsymbol{\omega}}]$, we write the direction cosine matrix in terms of the $\mathcal{B}$ frame unit vector components. Let b_i be the *1*x*3* matrices whose elements are the $\mathcal{N}$ frame components of the unit vectors $\hat{\boldsymbol{b}}_i$. The *3*x*3* matrix $[C]$ is then written as

$$[C] = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \tag{A.272}$$

The time derivative of $[C]$ is then expressed as

$$[\dot{C}] = \begin{bmatrix} \dot{b}_1 \\ \dot{b}_2 \\ \dot{b}_3 \end{bmatrix} \tag{A.273}$$

Since the $\mathcal{B}$ frame unit vectors are fixed within the $\mathcal{B}$ frame, their differential equation is of the form given in Eq. (1.14).

$$\hat{\boldsymbol{b}}_i = \boldsymbol{\omega} \times \hat{\boldsymbol{b}}_i \tag{A.274}$$

Substituting Eq. (A.274) into Eq. (A.273) and making use of $\boldsymbol{\omega} = \omega_1 \hat{\boldsymbol{b}}_1 + \omega_2 \hat{\boldsymbol{b}}_2 + \omega_3 \hat{\boldsymbol{b}}_3$, the direction cosine matrix derivative is written as

$$[\dot{C}] = \begin{bmatrix} -\omega_2 b_3 + \omega_3 b_2 \\ \omega_1 b_3 - \omega_3 b_1 \\ -\omega_1 b_2 + \omega_2 b_1 \end{bmatrix} \tag{A.275}$$

Using the definition of the *3x3* tilde matrix in Eq. (3.23), this is written in the desired form.

$$[\dot{C}] = -\begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = -[\tilde{\boldsymbol{\omega}}][C] \tag{A.276}$$

Therefore, for the case where $[C]$ represents a rigid body attitude, the skew-symmetric matrix $[Q]$ is equal to

$$[Q] = -[\tilde{\boldsymbol{\omega}}] \tag{A.277}$$

Let a N-dimensional vector $\boldsymbol{v}$ have components taken in the $\mathcal{N}$ and $\mathcal{B}$ frame. The *Nx1* matrix v_n contains the $\mathcal{N}$ frame components of $\boldsymbol{v}$, and v_b contains the $\mathcal{B}$ frame components. These components are mapped from one reference frame to another through the corresponding direction cosine matrix $[C]$.

$$v_b = [C] v_n \tag{A.278}$$

Note that both v_b and v_n in Eq. (A.278) are not treated as vectors, but as a list of scalars (i.e. a matrix). Therefore, taking the derivative of v_b is equivalent to taking the derivative of the vector $\boldsymbol{v}$ as seen by the $\mathcal{B}$ frame.

$$\dot{v}_b \simeq \frac{{}^{\mathcal{B}}d}{dt}(\boldsymbol{v}) \tag{A.279}$$

$$\dot{v}_n \simeq \frac{{}^{\mathcal{N}}d}{dt}(\boldsymbol{v}) \tag{A.280}$$

Taking the derivative of Eq. (A.278), we find

$$\dot{v}_b = [C]\dot{v}_n + [\dot{C}] v_n \tag{A.281}$$

Substituting Eqs. (A.271) and (A.278), this is rewritten as

$$\dot{v}_b = [C]\dot{v}_n + [Q][C]v_n = [C]\dot{v}_n + [Q]v_b \tag{A.282}$$

Solving for $\dot{v}_n$ we find a generalization of the transport theorem developed earlier using matrix notation.

$$\dot{v}_n = [C]^T (\dot{v}_b - [Q]v_b) \tag{A.283}$$

To show that Eq. (A.283) is equivalent to the transport theorem for the rigid body case presented in Chapter 1, we first use the identity $[Q] = -[\tilde{\boldsymbol{\omega}}]$

$$\dot{v}_n = [C]^T (\dot{v}_b + [\tilde{\boldsymbol{\omega}}]v_b) \tag{A.284}$$

Using the relationships in Eqs. (A.279) and (A.280) and the equivalent vector operator to the tilde matrix operator, the transport theorem is written as a vector algebra expression as

$$\frac{{}^{\mathcal{N}}d}{dt}(\boldsymbol{v}) = \frac{{}^{\mathcal{B}}d}{dt}(\boldsymbol{v}) + \boldsymbol{\omega} \times \boldsymbol{v} \tag{A.285}$$

where $\boldsymbol{\omega} = \boldsymbol{\omega}_{\mathcal{B}/\mathcal{N}}$. Strictly speaking, Eqs. (A.284) and (A.285) are only equivalent if all the vector quantities in Eq. (A.285) have components taken in the $\mathcal{N}$ frame.

Appendix B

State Space Analysis

text here

Appendix C

Various Euler Angle Transformations

This appendix contains mappings between the 12 sets of Euler angles $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)^T$ and either the direction cosine matrix $[C]$ or the body angular velocity vector $\boldsymbol{\omega}$.

The direction cosine matrix is defined as

$$\{\hat{\boldsymbol{b}}\} = [C(\theta_1, \theta_2, \theta_3)]\{\hat{\boldsymbol{n}}\}$$

Thus, it maps vectors with components in the inertial frame into vector with components taken in the body frame. The short hand notation $c\theta_i = \cos\theta_i$ and $s\theta_i = \sin\theta_i$ is used here.

Direction Cosine Matrix in Terms of the 12 Euler Angle Sets

1-2-1	$\begin{bmatrix} c\theta_2 & s\theta_2 s\theta_1 & -s\theta_2 c\theta_1 \\ s\theta_3 s\theta_2 & -s\theta_3 c\theta_2 s\theta_1 + c\theta_3 c\theta_1 & s\theta_3 c\theta_2 c\theta_1 + c\theta_3 s\theta_1 \\ c\theta_3 s\theta_2 & -c\theta_3 c\theta_2 s\theta_1 - s\theta_3 c\theta_1 & c\theta_3 c\theta_2 c\theta_1 - s\theta_3 s\theta_1 \end{bmatrix}$
1-2-3	$\begin{bmatrix} c\theta_3 c\theta_2 & c\theta_3 s\theta_2 s\theta_1 + s\theta_3 c\theta_1 & c\theta_3 s\theta_2 c\theta_1 + s\theta_3 s\theta_1 \\ -s\theta_3 c\theta_2 & -s\theta_3 s\theta_2 s\theta_1 + c\theta_3 c\theta_1 & s\theta_3 s\theta_2 c\theta_1 + c\theta_3 s\theta_1 \\ s\theta_2 & -c\theta_2 s\theta_1 & c\theta_2 c\theta_1 \end{bmatrix}$
1-3-1	$\begin{bmatrix} c\theta_2 & s\theta_2 c\theta_1 & s\theta_2 s\theta_1 \\ -c\theta_3 s\theta_2 & c\theta_3 c\theta_2 c\theta_1 - s\theta_3 s\theta_1 & c\theta_3 c\theta_2 s\theta_1 + s\theta_3 c\theta_1 \\ s\theta_3 s\theta_2 & -s\theta_3 c\theta_2 c\theta_1 - c\theta_3 s\theta_1 & -s\theta_3 c\theta_2 s\theta_1 + c\theta_3 c\theta_1 \end{bmatrix}$
1-3-2	$\begin{bmatrix} c\theta_3 c\theta_2 & c\theta_3 s\theta_2 c\theta_1 + s\theta_3 s\theta_1 & c\theta_3 s\theta_2 s\theta_1 - s\theta_3 c\theta_1 \\ -s\theta_2 & c\theta_2 c\theta_1 & c\theta_2 s\theta_1 \\ s\theta_3 c\theta_2 & +s\theta_3 s\theta_2 c\theta_1 - c\theta_3 s\theta_1 & s\theta_3 s\theta_2 s\theta_1 + c\theta_3 c\theta_1 \end{bmatrix}$

2-1-2	$\begin{bmatrix} -s\theta_3c\theta_2s\theta_1 + c\theta_3c\theta_1 & s\theta_3s\theta_2 & -s\theta_3c\theta_2c\theta_1 - c\theta_3s\theta_1 \\ s\theta_2s\theta_1 & c\theta_2 & s\theta_2c\theta_1 \\ c\theta_3c\theta_2s\theta_1 + s\theta_3c\theta_1 & -c\theta_3s\theta_2 & c\theta_3c\theta_2c\theta_1 - s\theta_3s\theta_1 \end{bmatrix}$
2-1-3	$\begin{bmatrix} s\theta_3s\theta_2s\theta_1 + c\theta_3c\theta_1 & s\theta_3c\theta_2 & s\theta_3s\theta_2c\theta_1 - c\theta_3s\theta_1 \\ c\theta_3s\theta_2s\theta_1 - s\theta_3c\theta_1 & c\theta_3c\theta_2 & c\theta_3s\theta_2c\theta_1 + s\theta_3s\theta_1 \\ c\theta_2s\theta_1 & -s\theta_2 & c\theta_2c\theta_1 \end{bmatrix}$
2-3-1	$\begin{bmatrix} c\theta_2c\theta_1 & s\theta_2 & -c\theta_2s\theta_1 \\ -c\theta_3s\theta_2c\theta_1 + s\theta_3s\theta_1 & c\theta_3c\theta_2 & c\theta_3s\theta_2s\theta_1 + s\theta_3c\theta_1 \\ s\theta_3s\theta_2c\theta_1 + c\theta_3s\theta_1 & -s\theta_3c\theta_2 & -s\theta_3s\theta_2s\theta_1 + c\theta_3c\theta_1 \end{bmatrix}$
2-3-2	$\begin{bmatrix} c\theta_3c\theta_2c\theta_1 - s\theta_3s\theta_1 & c\theta_3s\theta_2 & -c\theta_3c\theta_2s\theta_1 - s\theta_3c\theta_1 \\ -s\theta_2c\theta_1 & c\theta_2 & s\theta_2s\theta_1 \\ s\theta_3c\theta_2c\theta_1 + c\theta_3s\theta_1 & s\theta_3s\theta_2 & -s\theta_3c\theta_2s\theta_1 + c\theta_3c\theta_1 \end{bmatrix}$

3-1-2	$\begin{bmatrix} -s\theta_3s\theta_2s\theta_1 + c\theta_3c\theta_1 & s\theta_3s\theta_2c\theta_1 + c\theta_3s\theta_1 & -s\theta_3c\theta_2 \\ -c\theta_2s\theta_1 & c\theta_2c\theta_1 & s\theta_2 \\ c\theta_3s\theta_2s\theta_1 + s\theta_3c\theta_1 & -c\theta_3s\theta_2c\theta_1 + s\theta_3s\theta_1 & c\theta_3c\theta_2 \end{bmatrix}$
3-1-3	$\begin{bmatrix} c\theta_3c\theta_1 - s\theta_3c\theta_2s\theta_1 & c\theta_3s\theta_1 + s\theta_3c\theta_2c\theta_1 & s\theta_3s\theta_2 \\ -s\theta_3c\theta_1 - c\theta_3c\theta_2s\theta_1 & -s\theta_3s\theta_1 + c\theta_3c\theta_2c\theta_1 & c\theta_3s\theta_2 \\ s\theta_2s\theta_1 & -s\theta_2c\theta_1 & c\theta_2 \end{bmatrix}$
3-2-1	$\begin{bmatrix} c\theta_2c\theta_1 & c\theta_2s\theta_1 & -s\theta_2 \\ s\theta_3s\theta_2c\theta_1 - c\theta_3s\theta_1 & s\theta_3s\theta_2s\theta_1 + c\theta_3c\theta_1 & s\theta_3c\theta_2 \\ c\theta_3s\theta_2c\theta_1 + s\theta_3s\theta_1 & c\theta_3s\theta_2s\theta_1 - s\theta_3c\theta_1 & c\theta_3c\theta_2 \end{bmatrix}$
3-2-3	$\begin{bmatrix} c\theta_3c\theta_2c\theta_1 - s\theta_3s\theta_1 & c\theta_3c\theta_2s\theta_1 + s\theta_3c\theta_1 & -c\theta_3s\theta_2 \\ -s\theta_3c\theta_2c\theta_1 - c\theta_3s\theta_1 & -s\theta_3c\theta_2s\theta_1 + c\theta_3c\theta_1 & s\theta_3s\theta_2 \\ s\theta_2c\theta_1 & s\theta_2s\theta_1 & c\theta_2 \end{bmatrix}$

The following list provides the various forward and inverse mappings between the body angular velocity vector $\boldsymbol{\omega}$ and the Euler angle rates $\dot{\boldsymbol{\theta}} = (\dot{\theta}_1, \dot{\theta}_2, \dot{\theta}_3)^T$. The matrix $[B]$ is defined as

$$\dot{\boldsymbol{\theta}} = [B(\boldsymbol{\theta})]\boldsymbol{\omega}$$

The short hand notation $c\theta_i = \cos\theta_i$ and $s\theta_i = \sin\theta_i$ is used again here.

Mapping Between Body Angular Velocity Vector and the Euler Angle Rates

	$[B(\boldsymbol{\theta})]$	$[B(\boldsymbol{\theta})]^{-1}$
1-2-1	$\frac{1}{s\theta_2}\begin{bmatrix} 0 & s\theta_3 & c\theta_3 \\ 0 & s\theta_2 c\theta_3 & -s\theta_2 s\theta_3 \\ s\theta_2 & -c\theta_2 s\theta_3 & -c\theta_2 c\theta_3 \end{bmatrix}$	$\begin{bmatrix} c\theta_2 & 0 & 1 \\ s\theta_2 s\theta_3 & c\theta_3 & 0 \\ s\theta_2 c\theta_3 & -s\theta_3 & 0 \end{bmatrix}$
1-2-3	$\frac{1}{c\theta_2}\begin{bmatrix} c\theta_3 & -s\theta_3 & 0 \\ c\theta_2 s\theta_3 & c\theta_2 c\theta_3 & 0 \\ -s\theta_2 c\theta_3 & s\theta_2 s\theta_3 & c\theta_2 \end{bmatrix}$	$\begin{bmatrix} c\theta_2 c\theta_3 & s\theta_3 & 0 \\ -c\theta_2 s\theta_3 & c\theta_3 & 0 \\ s\theta_2 & 0 & 1 \end{bmatrix}$
1-3-1	$\frac{1}{s\theta_2}\begin{bmatrix} 0 & -c\theta_3 & s\theta_3 \\ 0 & s\theta_2 s\theta_3 & s\theta_2 c\theta_3 \\ s\theta_2 & c\theta_2 c\theta_3 & -c\theta_2 s\theta_3 \end{bmatrix}$	$\begin{bmatrix} c\theta_2 & 0 & 1 \\ -s\theta_2 c\theta_3 & s\theta_3 & 0 \\ s\theta_2 s\theta_3 & c\theta_3 & 0 \end{bmatrix}$
1-3-2	$\frac{1}{c\theta_2}\begin{bmatrix} c\theta_3 & 0 & s\theta_3 \\ -c\theta_2 s\theta_3 & 0 & c\theta_2 c\theta_3 \\ s\theta_2 c\theta_3 & c\theta_2 & s\theta_2 s\theta_3 \end{bmatrix}$	$\begin{bmatrix} c\theta_2 c\theta_3 & -s\theta_3 & 0 \\ -s\theta_2 & 0 & 1 \\ c\theta_2 s\theta_3 & c\theta_3 & 0 \end{bmatrix}$
2-1-2	$\frac{1}{s\theta_2}\begin{bmatrix} s\theta_3 & 0 & -c\theta_3 \\ s\theta_2 c\theta_3 & 0 & s\theta_2 s\theta_3 \\ -c\theta_2 s\theta_3 & s\theta_2 & c\theta_2 c\theta_3 \end{bmatrix}$	$\begin{bmatrix} s\theta_2 s\theta_3 & c\theta_3 & 0 \\ c\theta_2 & 0 & 1 \\ -s\theta_2 c\theta_3 & s\theta_3 & 0 \end{bmatrix}$
2-1-3	$\frac{1}{c\theta_2}\begin{bmatrix} s\theta_3 & c\theta_3 & 0 \\ c\theta_2 c\theta_3 & -c\theta_2 s\theta_3 & 0 \\ s\theta_2 s\theta_3 & s\theta_2 c\theta_3 & c\theta_2 \end{bmatrix}$	$\begin{bmatrix} c\theta_2 s\theta_3 & c\theta_3 & 0 \\ c\theta_2 c\theta_3 & -s\theta_3 & 0 \\ -s\theta_2 & 0 & 1 \end{bmatrix}$
2-3-1	$\frac{1}{c\theta_2}\begin{bmatrix} 0 & c\theta_3 & -s\theta_3 \\ 0 & c\theta_2 \; s\theta_3 & c\theta_2 c\theta_3 \\ c\theta_2 & -s\theta_2 c\theta_3 & s\theta_2 s\theta_3 \end{bmatrix}$	$\begin{bmatrix} s\theta_2 & 0 & 1 \\ c\theta_2 c\theta_3 & s\theta_3 & 0 \\ -c\theta_2 s\theta_3 & c\theta_3 & 0 \end{bmatrix}$

	$[B(\boldsymbol{\theta})]$	$[B(\boldsymbol{\theta})]^{-1}$
2-3-2	$\frac{1}{s\theta_2}\begin{bmatrix} c\theta_3 & 0 & s\theta_3 \\ -s\theta_2 s\theta_3 & 0 & s\theta_2 c\theta_3 \\ -c\theta_2 c\theta_3 & s\theta_2 & -c\theta_2 s\theta_3 \end{bmatrix}$	$\begin{bmatrix} s\theta_2 c\theta_3 & -s\theta_3 & 0 \\ c\theta_2 & 0 & 1 \\ s\theta_2 s\theta_3 & c\theta_3 & 0 \end{bmatrix}$
3-1-2	$\frac{1}{c\theta_2}\begin{bmatrix} -s\theta_3 & 0 & c\theta_3 \\ c\theta_2 c\theta_3 & 0 & c\theta_2 s\theta_3 \\ s\theta_2 s\theta_3 & c\theta_2 & -s\theta_2 c\theta_3 \end{bmatrix}$	$\begin{bmatrix} -c\theta_2 s\theta_3 & c\theta_3 & 0 \\ s\theta_2 & 0 & 1 \\ c\theta_2 c\theta_3 & s\theta_3 & 0 \end{bmatrix}$
3-1-3	$\frac{1}{s\theta_2}\begin{bmatrix} s\theta_3 & c\theta_3 & 0 \\ s\theta_2 c\theta_3 & -s\theta_2 s\theta_3 & 0 \\ -c\theta_2 s\theta_3 & -c\theta_2 c\theta_3 & s\theta_2 \end{bmatrix}$	$\begin{bmatrix} s\theta_3 s\theta_2 & c\theta_3 & 0 \\ s\theta_2 c\theta_3 & -s\theta_3 & 0 \\ c\theta_2 & 0 & 1 \end{bmatrix}$
3-2-1	$\frac{1}{c\theta_2}\begin{bmatrix} 0 & s\theta_3 & c\theta_3 \\ 0 & c\theta_2 c\theta_3 & -c\theta_2 s\theta_3 \\ c\theta_2 & s\theta_2 s\theta_3 & s\theta_2 c\theta_3 \end{bmatrix}$	$\begin{bmatrix} -s\theta_2 & 0 & 1 \\ c\theta_2 s\theta_3 & c\theta_3 & 0 \\ c\theta_2 c\theta_3 & -s\theta_3 & 0 \end{bmatrix}$
3-2-3	$\frac{1}{s\theta_2}\begin{bmatrix} -c\theta_3 & s\theta_3 & 0 \\ s\theta_2 s\theta_3 & s\theta_2 c\theta_3 & 0 \\ c\theta_2 c\theta_3 & -c\theta_2 s\theta_3 & s\theta_2 \end{bmatrix}$	$\begin{bmatrix} -s\theta_2 c\theta_3 & s\theta_3 & 0 \\ s\theta_2 s\theta_3 & c\theta_3 & 0 \\ c\theta_2 & 0 & 1 \end{bmatrix}$

Appendix D

Various Proofs

This appendix contains various proofs and developments of identities used within the text. Typically proving these identities where they were used would have been distracting, so these proofs were added to this appendix.

In developing the MRP rates relative to a rotating orbit frame $\mathcal{O}$, the identity

$$[B(\boldsymbol{\sigma})][BO(\boldsymbol{\sigma})] = [B(\boldsymbol{\sigma})]^T \tag{D.286}$$

was used. This identity can be developed from the basic MRP definitions of the $[B]$ and $[BO]$ matrices. Using Eqs. (3.144) and (3.150) we find

$$\ldots = \left(\left(1-\sigma^2\right) I_{3\times 3} + 2[\tilde{\boldsymbol{\sigma}}] + 2\boldsymbol{\sigma}\boldsymbol{\sigma}^T\right)\left(I_{3\times 3} + \frac{8[\tilde{\boldsymbol{\sigma}}]^2 - 4(1-\sigma^2)[\tilde{\boldsymbol{\sigma}}]}{(1+\sigma^2)^2}\right) \tag{D.287}$$

Substituting the identity

$$[\tilde{\boldsymbol{\sigma}}]^2 = \boldsymbol{\sigma}\boldsymbol{\sigma}^T - \sigma^2 I_{3\times 3} \tag{D.288}$$

and expanding the matrix product, $[B(\boldsymbol{\sigma})][BO(\boldsymbol{\sigma})]$ is rewritten as

$$\begin{aligned}[B(\boldsymbol{\sigma})][BO(\boldsymbol{\sigma})] = \left(1-\sigma^2\right) I_{3\times 3} + 2[\tilde{\boldsymbol{\sigma}}] + 2\boldsymbol{\sigma}\boldsymbol{\sigma}^T + \frac{4}{(1+\sigma^2)^2}\Big(2\boldsymbol{\sigma}\boldsymbol{\sigma}^T(1-\sigma^2) \\ - 2\sigma^2(1-\sigma^2)I_{3\times 3} - (1-\sigma^2)^2[\tilde{\boldsymbol{\sigma}}] - 4\sigma^2[\tilde{\boldsymbol{\sigma}}] - 2(1-\sigma^2)[\tilde{\boldsymbol{\sigma}}]^2\Big)\end{aligned} \tag{D.289}$$

Using Eq. (D.288) again, this is then reduced to

$$[B(\boldsymbol{\sigma})][BO(\boldsymbol{\sigma})] = \left(1-\sigma^2\right) I_{3\times 3} + 2[\tilde{\boldsymbol{\sigma}}] + 2\boldsymbol{\sigma}\boldsymbol{\sigma}^T - \frac{4(1+\sigma^2)^2[\tilde{\boldsymbol{\sigma}}]}{(1+\sigma^2)^2} \tag{D.290}$$

$$= \left(1-\sigma^2\right) I_{3\times 3} - 2[\tilde{\boldsymbol{\sigma}}] + 2\boldsymbol{\sigma}\boldsymbol{\sigma}^T \tag{D.291}$$

$$= [B(\boldsymbol{\sigma})]^T \tag{D.292}$$

Appendix E

Conic Section Transformations

Various transformations exist between the conic section elements. The commonly used mappings were developed in the previous celestial mechanics chapters. This appendix provides a complete list of all possible transformations between the orbit elements a, b, p, r_a, r_p and e for both the elliptical and hyperbolic case, as well as various anomaly mappings and sensitivities.

Elliptic Orbit Elements

a	$\frac{p}{1-e^2}$	$\frac{r_a}{1+e}$	$\frac{r_a+r_p}{2}$	$\frac{r_a^2}{2r_a-p}$	$\frac{b^2+r_a^2}{2r_a}$
	$\frac{b}{\sqrt{1-e^2}}$	$\frac{r_p}{1-e}$	$\frac{b^2}{p}$	$\frac{r_p^2}{2r_p-p}$	$\frac{b^2+r_p^2}{2r_p}$
b	$\frac{p}{\sqrt{1-e^2}}$	$\sqrt{ap}$	$\sqrt{2ar_a-r_a^2}$	$r_a\sqrt{\frac{1-e}{1+e}}$	$r_a\sqrt{\frac{p}{2r_a-p}}$
	$a\sqrt{1-e^2}$	$\sqrt{r_a r_p}$	$\sqrt{2ar_p-r_p^2}$	$r_p\sqrt{\frac{1+e}{1-e}}$	$r_p\sqrt{\frac{p}{2r_p-p}}$
p	$a(1-e^2)$	$r_a(1-e)$	$\frac{2r_a r_p}{r_a+r_p}$	$\frac{2b^2 r_a}{b^2+r_a^2}$	$2r_a-\frac{r_a^2}{a}$
	$b\sqrt{1-e^2}$	$r_p(1+e)$	$\frac{b^2}{a}$	$\frac{2b^2 r_p}{b^2+r_p^2}$	$2r_p-\frac{r_p^2}{a}$

e	$\sqrt{1-\frac{p}{a}}$	$\frac{r_a}{a}-1$	$1-\frac{p}{r_a}$	$\sqrt{1-\left(\frac{b}{a}\right)^2}$	$\frac{r_a^2-b^2}{r_a^2+b^2}$
	$\frac{r_a-r_p}{r_a+r_p}$	$1-\frac{r_p}{a}$	$\frac{p}{r_p}-1$	$\sqrt{1-\left(\frac{p}{b}\right)^2}$	$\frac{b^2-r_p^2}{b^2+r_p^2}$
r_a	$a(1+e)$	$\frac{p}{1-e}$	$r_p\frac{1+e}{1-e}$	$a+\sqrt{a^2-b^2}$	$a\left(1+\sqrt{1-\frac{p}{a}}\right)$
	$2a-r_p$	$\frac{b^2}{r_p}$	$b\sqrt{\frac{1+e}{1-e}}$	$\frac{pr_p}{2r_p-p}$	$\frac{b^2}{p}\left(1+\sqrt{1-\left(\frac{p}{b}\right)^2}\right)$
r_p	$a(1-e)$	$\frac{p}{1+e}$	$r_a\frac{1-e}{1+e}$	$a-\sqrt{a^2-b^2}$	$a\left(1-\sqrt{1-\frac{p}{a}}\right)$
	$2a-r_a$	$\frac{b^2}{r_a}$	$b\sqrt{\frac{1-e}{1+e}}$	$\frac{pr_a}{2r_a-p}$	$\frac{b^2}{p}\left(1-\sqrt{1-\left(\frac{p}{b}\right)^2}\right)$

Elliptic Anomaly Mapping and Sensitivities:

$$\sin f = \frac{\sqrt{1-e^2}\sin E}{1-e\cos E} \qquad \sin E = \frac{\sqrt{1-e^2}\sin f}{1+e\cos f}$$

$$\cos f = \frac{\cos E - e}{1-e\cos E} \qquad \cos E = \frac{e+\cos f}{1+e\cos f}$$

$$\tan\frac{f}{2} = \sqrt{\frac{1+e}{1-e}}\tan\frac{E}{2} \qquad \tan\frac{E}{2} = \sqrt{\frac{1-e}{1+e}}\tan\frac{f}{2}$$

$$\frac{df}{dE} = \frac{\sqrt{1-e^2}}{1-e\cos E} = \frac{1+e\cos f}{\sqrt{1-e^2}} = \frac{b}{r}$$

$$\frac{dM}{dE} = 1-e\cos E = \frac{1-e^2}{1+e\cos f} = \frac{r}{a}$$

$$\frac{dM}{df} = \frac{(1-e\cos E)^2}{\sqrt{1-e^2}} = \frac{(1-e^2)^{3/2}}{(1+e\cos f)^2} = \frac{r^2}{ab}$$

Hyperbolic Orbit Parameter Transformations

Note that by convention the semi-axis a and b are chosen to be negative quantities for the hyperbolic case. Since $r_a \to \infty$ for a hyperbola, the transformations

using r_a are omitted from the following list.

a	$\dfrac{b}{\sqrt{e^2-1}}$	$\dfrac{p}{1-e^2}$	$\dfrac{r_p^2-b^2}{2r_p}$
	$-\dfrac{b^2}{p}$	$\dfrac{r_p}{1-e}$	$\dfrac{r_p^2}{2r_p-p}$
b	$a\sqrt{e^2-1}$	$\dfrac{-p}{\sqrt{e^2-1}}$	$-\sqrt{r_p(r_p-2a)}$
	$-\sqrt{-ap}$	$-r_p\sqrt{\dfrac{e+1}{e-1}}$	$-r_p\sqrt{\dfrac{p}{p-2r_p}}$
p	$a(1-e^2)$	$-b\sqrt{e^2-1}$	$r_p\left(2-\dfrac{r_p}{a}\right)$
	$r_p(1+e)$	$-\dfrac{b^2}{a}$	$\dfrac{2r_pb^2}{b^2-r_p^2}$
e	$1+\dfrac{r_p}{a}$	$\dfrac{b^2+r_p^2}{b^2-r_p^2}$	$\sqrt{1+\left(\dfrac{b}{a}\right)^2}$
	$\dfrac{p}{r_p}-1$	$\sqrt{1+\dfrac{p}{a}}$	$\sqrt{1+\left(\dfrac{p}{b}\right)^2}$
r_p	$a(1-e)$	$-b\sqrt{\dfrac{e-1}{e+1}}$	$a\left(1-\sqrt{1+\dfrac{p}{a}}\right)$
	$\dfrac{p}{1+e}$	$a-\sqrt{a^2-b^2}$	$\dfrac{b^2}{p}\left(\sqrt{1+\left(\dfrac{p}{b}\right)^2}-1\right)$

Hyperbolic Anomaly Mapping and Sensitivities:

$$\sin f = \frac{\sqrt{e^2-1}\sinh H}{e\cosh H - 1} \qquad \sinh H = \frac{\sqrt{e^2-1}\sin f}{e\cos f + 1}$$

$$\cos f = \frac{e-\cosh H}{e\cosh H - 1} \qquad \cosh H = \frac{e+\cos f}{e\cos f + 1}$$

$$\tan\frac{f}{2}=\sqrt{\frac{e+1}{e-1}}\tanh\frac{H}{2}\qquad\qquad\tanh\frac{H}{2}=\sqrt{\frac{e-1}{e+1}}\tan\frac{f}{2}$$

$$\begin{aligned}\frac{df}{dH}&=\frac{\sqrt{e^2-1}}{e\cosh H-1}=\frac{e\cos f+1}{\sqrt{e^2-1}}=-\frac{b}{r}\\ \frac{dN}{dH}&=e\cosh H-1=\frac{e^2-1}{e\cos f+1}=-\frac{r}{a}\\ \frac{dN}{df}&=\frac{(e\cosh H-1)^2}{\sqrt{e^2-1}}=\frac{(e^2-1)^{3/2}}{(e\cos f+1)^2}=\frac{r^2}{ab}\end{aligned}$$

Appendix F
MATLAB M-Files

A rigid body kinematics MATLAB toolbox is included with this textbook in the form of a series of M-Files. The operators perform transformations between various sets of attitude coordinates, form the composition of two successive rotations, compute the relative attitude vector between two orientations and computes the time derivative of the attitude parameter vector. The attitude coordinates covered in this toolbox include the direction cosine matrix $[C]$, the Gibbs or classical Rodrigues parameter vector $\boldsymbol{q}$, the modified Rodrigues parameter vector $\boldsymbol{\sigma}$, the principal rotation vector $\boldsymbol{\gamma}$ and the 12 sets of Euler angle vectors $\boldsymbol{\theta}_{ijk}$. The scalar indices i, j and k are either 1, 2 or 3. All transformations used in this toolbox are introduced in Chapter 3.

The function `DirCos...(q)` returns the 3×3 direction cosine matrix $[C]$ corresponding to the particular choice in attitude coordinates. Instead of ..., the user adds what type of attitude vector `q` is. For Euler parameters, an `EP` is added. If `q` is a Gibbs vector, then `Gibbs` is added. The MRP vector simply has `MRP` added, while the principal rotation vector has `PRV` added. If `q` is an Euler angle vector, then `Eulerijk` is added where i, j and k are replaced with the appropriate rotation sequence. Therefore, if `q` is a (3-2-1) Euler angle vector, then the corresponding direction cosine matrix is found through the command `DirCosEuler321(q)`. The attitude coordinate abbreviations introduced here are used throughout the MATLAB subroutines. A direction cosine matrix $[C]$ is translated back to the various attitude parameters using the command `C2...(C)`.

To translate between various 3 or 4 parameter attitude coordinate sets, the command `...2...(q)` is used, where the ... are replaced with the previously discussed attitude coordinate choice abbreviations. Whenever possible, direct transformations between the various sets are used to provide numerically efficient code.

The command `q = add...(q1,q2)` computes the composition of the two successive rotations `q1` and `q2`. Note that both `q1` and `q2` must be the type of attitude parameters. Let $\mathcal{N}$, $\mathcal{B}$ and $\mathcal{F}$ be three reference frames, then `q` is defined through the relationship

$$[FN(\boldsymbol{q})] = [FB(\boldsymbol{q}_2)][BN(\boldsymbol{q}_1)] \tag{F.293}$$

To compute the relative orientation vector `q2`, the attitude vector `q1` is "rotationally" subtracted from `q`. Using the direction cosine matrix notation, this

corresponds to

$$[FB(\boldsymbol{q}_2)] = [FN(\boldsymbol{q})][BN(\boldsymbol{q}_1)]^T \tag{F.294}$$

The command to find the relative orientation vector `q2` is `sub...(q,q1)`.

The attitude coordinate rate vector $\dot{\boldsymbol{q}}$ is related to $\boldsymbol{\omega}$ through a matrix $[B(\boldsymbol{q})]$.

$$\dot{\boldsymbol{q}} = [B(\boldsymbol{q})]\boldsymbol{\omega} \tag{F.295}$$

The MATLAB command `d...(q,w)` computes the time derivative of the attitude vector `q` for a given body angular velocity vector `w`. For example, if $\boldsymbol{q}$ is a (1-2-3) Euler angle vector, then the command `dEuler123(q)` would be invoked. Subroutines are also provided that compute just the $[B(\boldsymbol{q})]$ matrix and it's inverse. All attitude coordinates discussed have compact analytical inverse formulas for $[B(\boldsymbol{q})]$ as shown in Chapter 3. The $[B(\boldsymbol{q})]$ matrix is computed with the command `Bmat...(q)` and its inverse with `Binv...(q)`.

The following alphabetical list details the purpose of each MATLAB function provided in the rigid body kinematics toolbox.

`addEP(q1,q2)`	Sum the two Euler parameter vectors.
`addEulerijk(q1,q2)`	Sum the two (i-j-k) Euler angle vectors.
`addGibbs(q1,q2)`	Sum the two Gibbs vectors.
`addMRP(q1,q2)`	Sum the two MRP vectors.
`addPRV(q1,q2)`	Sum the two principal rotation vectors.
`BinvEP(q)`	Compute the inverse of $[B(\boldsymbol{\beta})]$.
`BinvEulerijk(q)`	Compute the inverse of $[B(\boldsymbol{\theta}_{ijk})]$.
`BinvGibbs(q)`	Compute the inverse of $[B(\boldsymbol{q})]$.
`BinvMRP(q)`	Compute the inverse of $[B(\boldsymbol{\sigma})]$.
`BinvPRV(q)`	Compute the inverse of $[B(\boldsymbol{\gamma})]$.
`BmatEP(q)`	Compute the matrix $[B(\boldsymbol{\beta})]$
`BmatEulerijk(q)`	Compute the matrix $[B(\boldsymbol{\theta}_{ijk})]$
`BmatGibbs(q)`	Compute the matrix $[B(\boldsymbol{q})]$
`BmatMRP(q)`	Compute the matrix $[B(\boldsymbol{\sigma})]$
`BmatPRV(q)`	Compute the matrix $[B(\boldsymbol{\gamma})]$
`C2EP(C)`	Extract the Euler parameters from $[C]$.
`C2Eulerijk(C)`	Extract the (i-j-k) Euler angles from $[C]$.
`C2Gibbs(C)`	Extract the Gibbs vector from $[C]$.
`C2MRP(C)`	Extract the MRP vector from $[C]$.
`C2PRV(C)`	Extract the principal rotation vector from $[C]$.
`dEP(q,w)`	Compute the Euler parameter time derivative.
`dEulerijk(q,w)`	Compute the (i-j-k) Euler angles time derivative.
`dGibbs(q,w)`	Compute the Gibbs vector time derivative.
`DirCosEP(q)`	Translate the Euler parameters into $[C]$.
`DirCosEulerijk(q)`	Translate the (i-j-k) Euler angles into $[C]$.
`DirCosGibbs(q)`	Translate the Gibbs vector into $[C]$.
`DirCosMRP(q)`	Translate the MRP vector into $[C]$.
`DirCosPRV(q)`	Translate the principal rotation vector into $[C]$.

`dMRP(q,w)`	Compute the MRP vector time derivative.
`dPRV(q,w)`	Compute the principal rotation vector time derivative.
`elem2PRV(q)`	Translates the $(\Phi, \hat{e}_1, \hat{e}_2, \hat{e}_3)$ into the principal rotation vector.
`EP2Eulerijk`	Translate Euler parameters into (i-j-k) Euler angles.
`EP2Gibbs`	Translate Euler parameters into a Gibbs vector.
`EP2MRP`	Translate Euler parameters into a MRP vector.
`EP2PRV`	Translate Euler parameters into a PRV vector.
`Euler1(theta)`	Returns the elementary rotation matrix about the first body axis.
`Euler2(theta)`	Returns the elementary rotation matrix about the second body axis.
`Euler3(theta)`	Returns the elementary rotation matrix about the third body axis.
`Eulerijk2EP(q)`	Translate the (i-j-k) Euler angles into Euler parameters.
`Eulerijk2Gibbs(q)`	Translate the (i-j-k) Euler angles into the Gibbs vector.
`Eulerijk2MRP(q)`	Translate the (i-j-k) Euler angles into MRPs.
`Eulerijk2PRV(q)`	Translate the (i-j-k) Euler angles into the principal rotation vector.
`Gibbs2EP(q)`	Translate the Gibbs vector into Euler parameters.
`Gibbs2Eulerijk(q)`	Translate the Gibbs vector into (i-j-k) Euler angles.
`Gibbs2MRP(q)`	Translate the Gibbs vector into MRPs.
`Gibbs2PRV(q)`	Translate the Gibbs vector into the principal rotation vector.
`MRP2EP(q)`	Translate the MRPs into Euler parameters.
`MRP2Eulerijk(q)`	Translate the MRPs into (i-j-k) Euler angles.
`MRP2Gibbs(q)`	Translate the MRPs into the Gibbs vector.
`MRP2PRV(q)`	Translate the MRPs into the principal rotation vector.
`MRPswitch(q,S)`	Switch the MRP vector such that $\|\boldsymbol{\sigma}\|^2 < S$.
`PRV2elem(q)`	Translates the principal rotation vector to $(\Phi, \hat{e}_1, \hat{e}_2, \hat{e}_3)$.
`PRV2EP(q)`	Translates the principal rotation vector to Euler parameters.
`PRV2Eulerijk(q)`	Translates the principal rotation vector to (i-j-k) Euler angles.
`PRV2Gibbs(q)`	Translates the principal rotation vector to the Gibbs vector.
`PRV2MRP(q)`	Translates the principal rotation vector to MRPs.
`subEP(q,q1)`	Compute the relative Euler parameter vector from `q1` to `q`.
`subEulerijk(q,q1)`	Compute the relative (i-j-k) Euler angles vector from `q1` to `q`.

`subGibbs(q,q1)`	Compute the relative Gibbs vector from `q1` to `q`.
`subMRP(q,q1)`	Compute the relative MRP vector from `q1` to `q`.
`subPRV(q,q1)`	Compute the relative PRV vector from `q1` to `q`.

Appendix G

First-Order Mapping Between Mean and Osculating Orbit Elements

A first-order mapping algorithm is outlined in this Appendix based on the theory developed by Brouwer in Reference 1 and Lyddane in Reference 2. The modifications suggested by Lyddane allow for a more robust mapping near zero eccentricities and inclination angles.

This mapping directly translates any osculating (instantaneous) orbit elements into mean (orbit averaged, with short and long period motion removed) orbit element equivalent values. Only first order J_2 terms are retained in this algorithm. Note that the forward and inverse transformation here is not perfect due to the first-order truncation of the infinite series. Small errors of order J_2 are to be expected.

Note that since a first-order truncation is performed of the infinite power series solution, the forward and inverse mapping function between the mean and osculating orbit elements only differs by a sign. Let the original orbit elements be given by $\boldsymbol{e} = (a, e, i, \Omega, \omega, M)$. Note that these elements could be either mean or osculating orbit elements. The transformed elements will be given through $\boldsymbol{e}' = (a', e', i', \Omega', \omega', M')$. With r_e being Earth's equatorial radius, the parameter γ_2 is either defined as

$$\gamma_2 = \frac{J_2}{2}\left(\frac{r_e}{a}\right)^2 \tag{G.296}$$

if the algorithm maps mean orbit elements to osculating orbit elements, or as

$$\gamma_2 = -\frac{J_2}{2}\left(\frac{r_e}{a}\right)^2 \tag{G.297}$$

if the algorithm maps osculating orbit elements to mean orbit elements.

Defining $\eta = \sqrt{1 - e^2}$, the parameter γ_2' is then defined as

$$\gamma_2' = \frac{\gamma_2}{\eta^4} \tag{G.298}$$

The mean anomaly M is translated into the corresponding eccentric anomaly E using Kepler's equation.

$$M = E - e\sin E \tag{G.299}$$

The true anomaly f is computed using

$$f = 2\tan^{-1}\left(\sqrt{\frac{1+e}{1-e}}\tan\left(\frac{E}{2}\right)\right) \tag{G.300}$$

The ratio a/r is computed using

$$\frac{a}{r} = \frac{1+e\cos f}{\eta^2} \tag{G.301}$$

with r being the current orbit radius.

The transformed semi-major axis a' (which could be either the mean or osculating state, depending on whether a is an osculating or mean element) is computed through

$$\begin{aligned} a' = a + a\gamma_2\Big((3\cos^2 i - 1)\left(\left(\frac{a}{r}\right)^3 - \frac{1}{\eta^3}\right) \\ + 3(1-\cos^2 i)\left(\frac{a}{r}\right)^3\cos(2\omega+2f)\Big) \end{aligned} \tag{G.302}$$

To following parameters are intermediate results used to transform the remaining orbit elements.

$$\delta e_1 = \frac{\gamma_2'}{8}e\eta^2\left(1 - 11\cos^2 i - 40\frac{\cos^4 i}{1-5\cos^2 i}\right)\cos(2\omega) \tag{G.303}$$

$$\begin{aligned} \delta e = \delta e_1 + \frac{\eta^2}{2}\Big\{\gamma_2\Big[\frac{3\cos^2 i - 1}{\eta^6}\Big(e\eta + \frac{e}{1+\eta} + 3\cos f \\ + 3e\cos^2 f + e^2\cos^3 f\Big) + 3\frac{1-\cos^2 i}{\eta^6}\big(e \\ + 3\cos f + 3e\cos^2 f + e^2\cos^3 f\big)\cos(2\omega+2f)\Big] \\ - \gamma_2'(1-\cos^2 i)\left(3\cos(2\omega+f) + \cos(2\omega+3f)\right)\Big\} \end{aligned} \tag{G.304}$$

$$\begin{aligned} \delta i = -\frac{e\delta e_1}{\eta^2\tan i} + \frac{\gamma_2'}{2}\cos i\sqrt{1-\cos^2 i}\Big(3\cos(2\omega+2f) \\ + 3e\cos(2\omega+f) + e\cos(2\omega+3f)\Big) \end{aligned} \tag{G.305}$$

$$
\begin{aligned}
M'+\omega'+\Omega' = M+\omega+\Omega &+ \frac{\gamma_2'}{8}\eta^3\Big(1-11\cos^2 i - 40\frac{\cos^4 i}{1-5\cos^2 i}\Big)\\
&- \frac{\gamma_2'}{16}\Big(2+e^2-11(2+3e^2)\cos^2 i\\
&- 40(2+5e^2)\frac{\cos^4 i}{1-5\cos^2 i} - 400e^2\frac{\cos^6 i}{(1-5\cos^2 i)^2}\Big)\\
&+ \frac{\gamma_2'}{4}\Big(-6*(1-5\cos^2 i)(f-M+e\sin f)\\
&+ (3-5\cos^2 i)(3\sin(2\omega+2f)+3e\sin(2\omega+f)\\
&+ e\sin(2\omega+3f))\Big)\\
&- \frac{\gamma_2'}{8}e^2\cos i\Big(11+80\frac{\cos^2 i}{1-5\cos^2 i}+200\frac{\cos^4 i}{(1-5\cos^2 i)^2}\Big)\\
&- \frac{\gamma_2'}{2}\cos i\Big(6(f-M+e\sin f)\\
&- 3\sin(2\omega+2f) - 3e\sin(2\omega+f) - e\sin(2\omega+3f)\Big)
\end{aligned}
\tag{G.306}
$$

$$
\begin{aligned}
(e\delta M) = &\frac{\gamma_2'}{8}e\eta^3\Big(1-11\cos^2 i - 40\frac{\cos^4 i}{1-5\cos^2 i}\Big)\\
&- \frac{\gamma_2'}{4}\eta^3\Big\{2(3\cos^2 i-1)\left(\left(\frac{a\eta}{r}\right)^2+\frac{a}{r}+1\right)\sin f\\
&+ 3(1-\cos^2 i)\Big[\left(-\left(\frac{a\eta}{r}\right)^2-\frac{a}{r}+1\right)\sin(2\omega+f)\\
&+ \left(\left(\frac{a\eta}{r}\right)^2+\frac{a}{r}+\frac{1}{3}\right)\sin(2\omega+3f)\Big]\Big\}
\end{aligned}
\tag{G.307}
$$

$$
\begin{aligned}
\delta\Omega = &-\frac{\gamma_2'}{8}e^2\cos i\Big(11+80\frac{\cos^2 i}{1-5\cos^2 i}+200\frac{\cos^4 i}{(1-5\cos^2 i)^2}\Big)\\
&- \frac{\gamma_2'}{2}\cos i\Big(6(f-M+e\sin f) - 3\sin(2\omega+2f)\\
&- 3e\sin(2\omega+f) - e\sin(2\omega+3f)\Big)
\end{aligned}
\tag{G.308}
$$

Now we are ready to compute the remaining transformed orbit elements. By defining

$$d_1 = (e+\delta e)\sin M + (e\delta M)\cos M \tag{G.309}$$

$$d_2 = (e+\delta e)\cos M - (e\delta M)\sin M \tag{G.310}$$

the mean anomaly M' is computed using

$$M' = \tan^{-1}\left(\frac{d_1}{d_2}\right) \tag{G.311}$$

while the eccentricity e' is computed using

$$e' = \sqrt{d_1^2 + d_2^2} \tag{G.312}$$

Similarly, we define

$$d_3 = \left(\sin\left(\frac{i}{2}\right) + \cos\left(\frac{i}{2}\right)\frac{\delta i}{2}\right)\sin\Omega + \sin\left(\frac{i}{2}\right)\delta\Omega\cos\Omega \tag{G.313}$$

$$d_4 = \left(\sin\left(\frac{i}{2}\right) + \cos\left(\frac{i}{2}\right)\frac{\delta i}{2}\right)\cos\Omega - \sin\left(\frac{i}{2}\right)\delta\Omega\sin\Omega \tag{G.314}$$

to compute the ascending node Ω' through

$$\Omega' = \tan^{-1}\left(\frac{d_3}{d_4}\right) \tag{G.315}$$

and the inclination angle i' through

$$i' = 2\sin^{-1}\left(\sqrt{d_3^2 + d_4^2}\right) \tag{G.316}$$

Finally, the argument of perigee ω' is computed through

$$\omega' = (M' + \omega' + \Omega') - M' - \Omega' \tag{G.317}$$

Note that when computing the inverse tangent functions in the algorithm above, care must be taken such that the resulting angle lies in the proper quadrant.

Bibliography

[1] Brouwer, D., "Solution of the Problem of Artificial Satellite Theory Without Drag," *The Astronautical Journal*, Vol. 64, No. 1274, 1959, pp. 378–397.

[2] Lyddane, R. H., "Small Eccentricities or Inclinations in the Brouwer Theory of the Artificial Satellite," *The Astronomical Journal*, Vol. 68, No. 8, October 1963, pp. 555–558.

Index