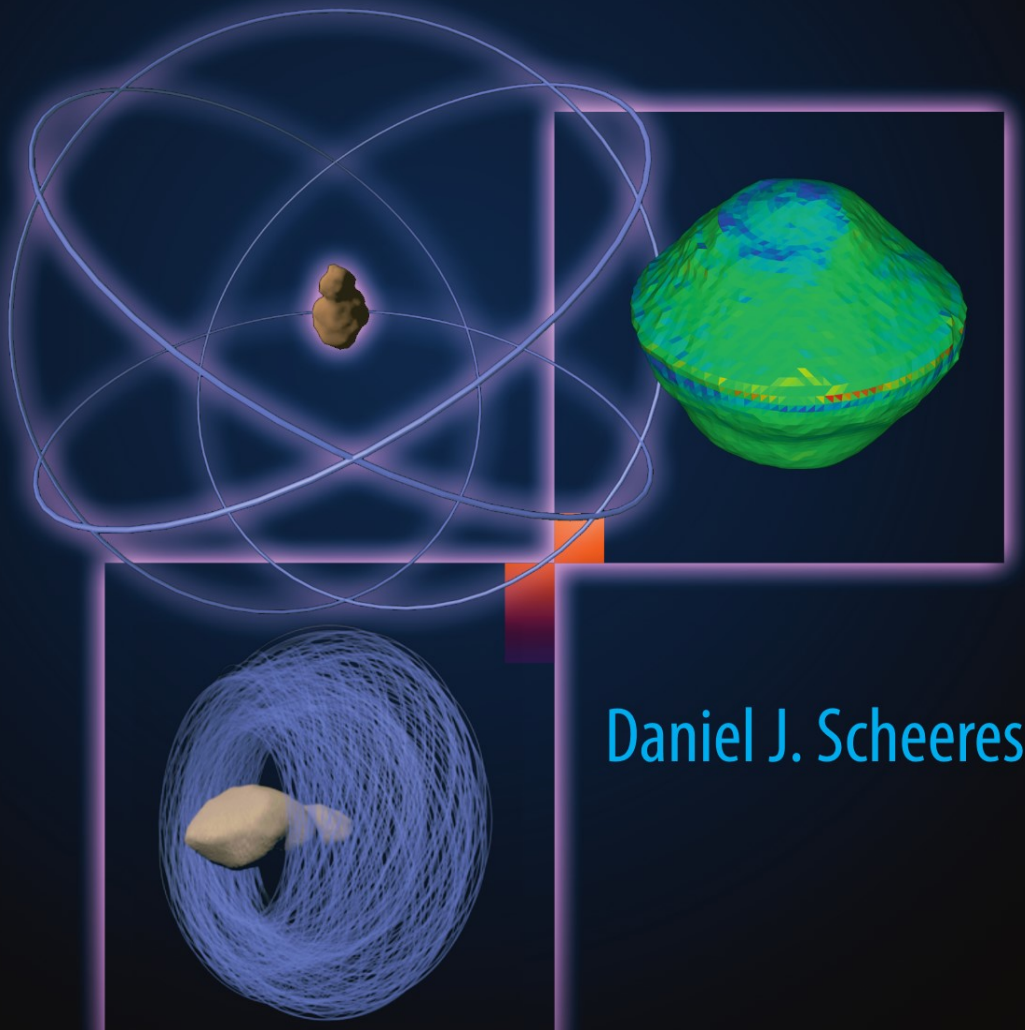
ORBITAL MOTION IN STRONGLY PERTURBED ENVIRONMENTS

Applications to Asteroid, Comet and Planetary Satellite Orbiters



Daniel J. Scheeres

Springer

PRAXIS

Orbital Motion in Strongly Perturbed Environments

Applications to Asteroid, Comet and Planetary Satellite Orbiters

Daniel J. Scheeres

Orbital Motion in Strongly Perturbed Environments

Applications to Asteroid, Comet and Planetary Satellite Orbiters

Springer

Published in association with
Praxis Publishing
Chichester, UK

PRAXIS

Dr. Daniel J. Scheeres
Colorado Center for Astrodynamics Research
Dept. of Aerospace Engineering Sciences
University of Colorado
Boulder
Colorado
U.S.A.

SPRINGER–PRAXIS BOOKS IN ASTRONAUTICAL ENGINEERING
SUBJECT *ADVISORY EDITOR*: John Mason, M.B.E., B.Sc., M.Sc., Ph.D.

ISBN 978-3-642-03255-4 e-ISBN 978-3-642-03256-1
DOI 10.1007/978-3-642-03256-1

Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011938789

Cover design: Jim Wilkie
Project copy editor: Mike Shardlow
Author-generated LaTex, processed by EDV-Beratung, Germany

Printed on acid-free paper

Springer is part of Springer Science + Business Media (www.springer.com)

This book is dedicated to the memory of
Steven J. Ostro (1946–2008),
my colleague, mentor and friend.

Contents

IV. Appendices

Preface

The *in situ* exploration of small bodies, most typically asteroids and comets but also including planetary satellites, is an exciting endeavor currently at the forefront of planetary science. The science one can do at such bodies is as fundamental as it is important, as these bodies provide windows into the past of the Solar System. Found within the asteroid population are various stages of planetary formation, albeit in shards and tumbled with each other. Found within the comet population are the pristine chemicals that dominated the proto-planetary disk prior to the formation of the modern Solar System. Found within planetary satellites are whole snapshots of the formational epoch of the Solar System, and also miniature worlds that have evolved towards their own unique ends. Thus, justification for the exploration of these bodies is well-founded and has motivated large portions of the planetary science community over the last decades.

More recently, these small bodies, especially Near-Earth asteroids, have also become of interest for human exploration of the Solar System. Motivated by an eventual human mission to Mars, much debate and discussion in the space exploration community has swirled around what the appropriate pathway towards this eventual goal may be. As of this writing, the current pathway towards Mars is seen to lie through an initial human exploration mission to a Near-Earth asteroid. The topics covered in this book are fundamental for the design, evaluation and navigation of such missions. Human exploration of asteroids can also be envisioned as a useful endeavor for the identification and exploitation of extra-terrestrial resources. Indeed, there has been much serious scholarship focused on how asteroid materials can be utilized for sustaining human presence in outer space. Finally, the mitigation of hazardous asteroids and comets on potential impact trajectories with the Earth relies fundamentally on our exploration and understanding of spacecraft mechanics at these bodies.

Thus there is a cornucopia of well-motivated scientific and exploration activities at asteroids, comets and planetary satellites. While there have been countless studies, proposals and papers describing these activities, there has not been such a clear focus on the practicalities of what one can do with space vehicles when they arrive at these bodies. It is surprising to note that there are many more studies of

how to plan a spacecraft's path to such small bodies than there are what should be done once one arrives. Whereas the theory for plotting a course to an asteroid, comet or planetary satellite is well understood and has been implemented many times in recent decades, the fundamental mechanics of motion in the vicinity of a small body is not fully understood and cannot be understood based solely on a simple application of Kepler's laws. Thus, the opinion is frequently expressed that it is *impossible* for a spacecraft to orbit an asteroid smaller than, say, a few tens of meters in size. But, on closer inspection, this turns out to be fully feasible in many circumstances. Similar examples abound and, while not always discussed specifically, will be addressed through the content of this book.

The goal of this book is to remove some of this mystery, and to lay out the fundamental mechanics of what one *could* do with a spacecraft when visiting a small body. While this book is not the definitive summation of all the work that has been done on this subject, it can at least serve as a background for further study and analysis.

The text is divided into three parts. Part I reviews the basics of small bodies in the Solar System, their orbits, their spin states, their sizes, their morphological properties, and the force environment about them. The emphasis is on describing those features of these bodies that are important for understanding orbital mechanics about them. Accordingly, the book does not discuss the many interesting scientific aspects of these bodies nor does it discuss motivations for their explorations – such motivations are taken as fact. Part II presents the background dynamical theory that is necessary if one wishes to fully explore the dynamics of motion about these bodies. For an expert in dynamical systems, these chapters may seem a bit naive, while to a novice in astrodynamics they may seem relatively advanced. The material presented in these chapters is there mainly because I have found these results useful in my own research on this problem, and thus feel that they must be presented. Part III applies and analyzes a range of different types and situations that may exist at a variety of small bodies of the Solar System. The list is certainly not exhaustive, as small bodies have a penchant for presenting hitherto unforeseen dynamical situations. Indeed, this is what makes them so exciting. The intent of these chapters is to provide case studies that can be used as a reference for other small bodies considered for exploration, and to serve as a springboard for investigating new situations or configurations that may arise.

There are many people to whom I owe a debt of gratitude for the development of this book and the topics covered herein. First, I must acknowledge the many students who have worked with me on these topics. In a very real sense, all of the students I have worked with on research have shaped me and my approach to these topics – oftentimes the connections between their research and how it impacts topics in this book are not clear, yet they exist nonetheless. Of special note are several students whose thesis research is clearly and explicitly called out in this book. I list them here in order of their graduation: Weiduo Hu, Benjamin Villac, Marci Possner (nee Paskowitz), Stephen Broschart, Ryan Park, Julie Bellerose, Eugene Fahnestock and Oier Peñagaricano-Muñoa. A special thank you is given to Aaron Rosengren for help in proof-reading the book.

Second, I must acknowledge my many colleagues from the scientific and engineering communities that have encouraged my research and continually asked the probing questions that force one to further refine and dig into this research. Primary among these is Steven Ostro, to whom I dedicate this book. It is not an understatement to say that Steve's enthusiasm for my research on these problems, his probing questions that went far beyond science, and his continual encouragement have served as the foundations of my success in my professional life. Many others have also encouraged or enabled my research on these problems. Two who have provided such support during crucial periods of my career are Bobby G. Williams and Donald K. Yeomans.

Finally, and most importantly, I must acknowledge my wife, Susan Postema Scheeres, my children Annaka, Samuel and Eleanor, and my parents, Jacob and Ann, for their continual encouragement, support and confidence in me.

Boulder, Colorado, August 2011 *Daniel J. Scheeres*

Part I
Modeling

1. Introduction and Background

1.1 Background

When studying motion in the vicinity of small bodies one must be cautious regarding what assumptions can be made, due to the variety and diversity of shapes, sizes and morphologies that these bodies present. One reliable assumption is that the total mass of these bodies is tiny as compared to planetary bodies, but huge as compared to visiting spacecraft. For an extreme example, even a boulder of size 1 meter with a density of 3.5 grams per cubic centimeter will still be an order of magnitude more massive than a typical interplanetary spacecraft, with the mass of the "asteroid" increasing by an order of magnitude with every doubling of the body's size. Thus, just as one can neglect the effect of small bodies on the planets and sun that they live near, in turn one can neglect the effect of a spacecraft when it orbits about these small bodies.

It is an understatement to say that all other physical assumptions are suspect and must be considered on a case-by-case basis. Almost all other overriding physical assumptions about these bodies could be violated, such is their diversity. Examples are easy to find, and include the assumption (prior to 1993) that asteroids were overwhelmingly single bodies whereas now it is known that large fractions of them are binaries and can even form stable triple systems. Previous to high-resolution imaging of asteroids, it was thought that small asteroids were monolithic bodies without any surface covering (called regolith), yet now every asteroid observed at high-resolution proves to be covered with pebbles, rocks and boulders. Recent photometric observations and surveys of asteroids have also overturned our conception of their spin states, and revealed interesting relationships between the sizes of asteroids and their spin rates, with small asteroids less than 100 meters in size able to spin with rotation periods on the order of minutes [133]. In contrast with our previous understanding, it may even be feasible that some of these fast-rotating bodies are in complex spin rates as well [134]. Driven by these realities, we will strive to make as few assumptions as possible in laying out our theory and applications.

It must also be admitted that the study of dynamics about small bodies requires a knowledge and understanding of higher-order theories of celestial mechanics ap-

D.J. Scheeres, *Orbital Motion in Strongly Perturbed Environments*, Springer Praxis Books,
DOI 10.1007/978-3-642-03256-1_1,

plied to space vehicles, a field frequently called "astrodynamics". Thus, a substantial portion of this book provides a review of fundamental theories and facts of astrodynamical systems, and strives to present them in a practical manner which facilitates their application to the real problems one encounters when planning and navigating space vehicles to visit small bodies.

1.2 A Brief History of Asteroid Exploration

We first present a brief review of the ways in which asteroids have been explored, remotely and *in situ*, since their discovery.

1.2.1 Optical Observations

The first asteroid discovered, Ceres, was observed by Giuseppe Piazzi on January 1, 1801, from the Palermo Observatory in Sicily. Since that time optical observation has been the primary, if not only, method of discovery for these bodies. In recent years the discovery rates of asteroids has surged, given the introduction of targeted surveys for Near-Earth asteroids (asteroids with a perihelion less than 1.3 AU) and general surveys in the Main Belt. To date the total number of asteroids detected ranges above 300,000, while only approximately half of these have precise orbits determined.

The vast majority of these observations only detect asteroids as points of light. The recent introduction of a new class of precision telescopes in orbit and on Earth have enabled the larger asteroids to be imaged and has led to the discovery of a number of binary asteroids [105]. In addition to detection, there are a few limited physical characteristics that can be deduced from these observations. The magnitude, or brightness, of an observed asteroid, once its orbit has been determined, can be used to estimate the body's total size (once an albedo is assumed). Also, the fluctuation of the reflected light, called a light curve, can enable the asteroid's shape to be constrained and its spin period to be estimated. Multiple observations of light curves can even determine the spin pole orientation of the asteroid. Light curves can also be used to detect binary asteroids, based on observing abrupt changes in reflected light when the bodies eclipse each other. Finally, observations of asteroids at precise wavelengths from optical into the infra-red allows their spectra and temperature to be determined. These measurements provide additional insight into the compositional properties of these bodies and their size. A complete introduction to asteroids and historical measurements that have been made of them can be found in the Asteroids, Asteroids II and Asteroids III texts [48, 102, 14].

Despite the importance of remote optical observations of asteroids, they can only provide limited detail on the specific bodies at hand. The best determinations possible usually occur for binary asteroids, where it is possible to measure the total system mass and infer mass distributions between the two bodies [135]. These observations provide tantalizing clues to the natures of these bodies, but cannot

answer fundamental questions about them. For that one needs to turn to other observation methodologies.

1.2.2 Radar Observations

Range-Doppler radar observations of asteroids is a powerful observation technique that has been used to open new windows on our understanding of asteroids [120]. These observation techniques allow for precise determination of an asteroid's orbit and provide measurement data from which the shape and spin state of the body can often be derived. Typical shape model reconstruction accuracies on the order of 10 meters are possible, enabling a relatively precise level of insight into the shape of these bodies. Such models were not available previously to the first shape model determined for the asteroid Castalia [75], which is notable for being a clearly defined contact binary asteroid (Fig. 1.1). Since that time many additional asteroid shape models have been determined, and some of them, such as the binary asteroid 1999 KW4 [121, 165], have been used to greatly advance our understanding of asteroids (Fig. 1.2).

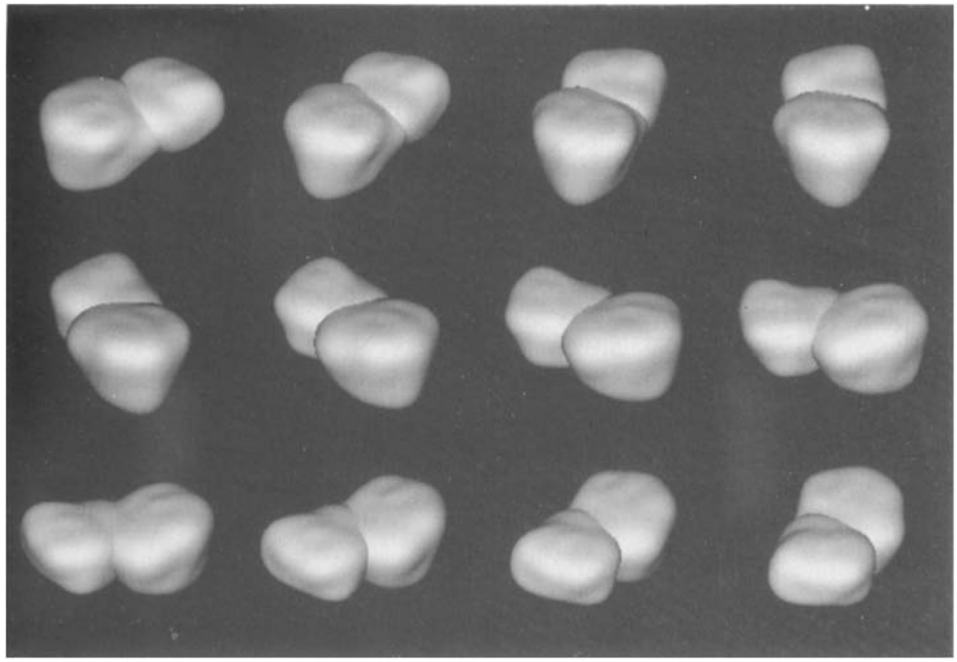


Fig. 1.1 Shape model of Castalia based on radar observations. Credit: JPL/CALTECH.

There are only two radio telescopes that can currently generate observations that can lead to the shape and spin models. These are the Arecibo Observatory in Puerto Rico and the Deep Space Network's 70-meter dish at Goldstone, California. These

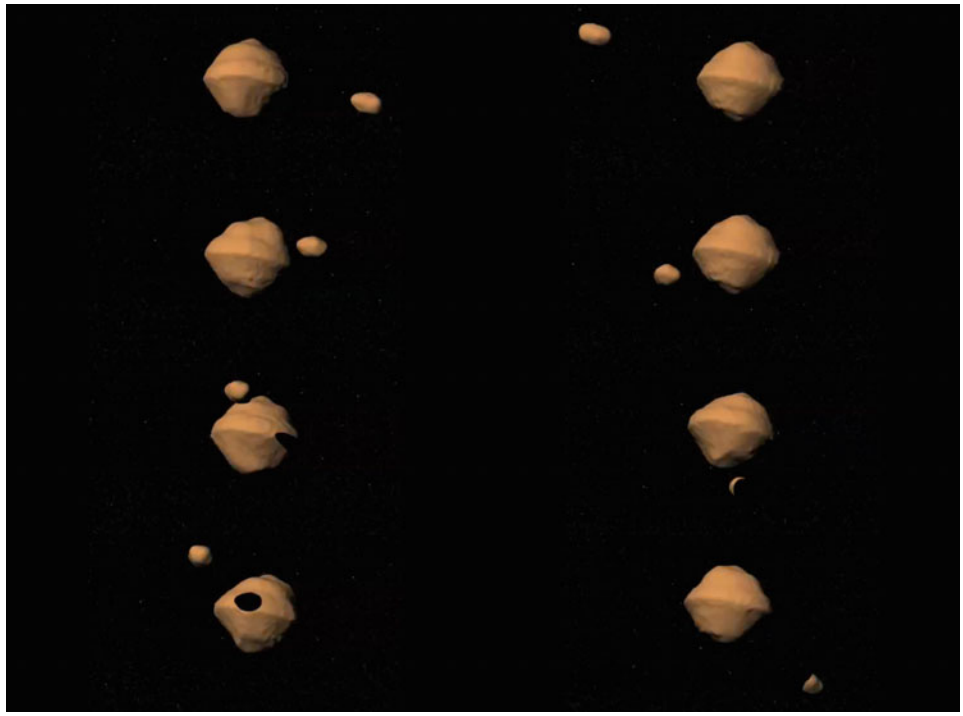


Fig. 1.2 Shape models and relative orbit of binary asteroid 1999 KW4 based on radar observations. Credit: DIAL/JPL/CALTECH.

observatories do not carry out discovery observations of asteroids, but are used to follow up discovered asteroids if they happen to pass close enough to the Earth to enable detection and observation. Perhaps the main drawback of radar observations is that they require the target bodies to have relatively close flybys to Earth for the observations to have sufficient signal-to-noise ratios for these detailed shapes to be determined. Despite this, radar observations of asteroids have revolutionized our understanding of these bodies.

1.2.3 Distant Flybys

Moving beyond ground-based observations can yield significant scientific returns through the use of instrumented spacecraft. The Galileo mission to Jupiter took advantage of its multi-year trajectory to that planet to carry out two flybys of Main Belt asteroids, obtaining the first high-resolution optical images of asteroids. The first flyby was of the asteroid Gaspra on October 29, 1991. The second flyby was of the asteroid Ida on August 28, 1993. It is significant that the second asteroid ever imaged at close range, Ida, was found to have a moon captured in orbit about it, Dactyl, shown in Fig. 1.3. Thus, the second asteroid observed at high-resolution was found to be a binary, presaging the many discoveries since that time. The Ida

Fig. 1.3 Galileo spacecraft image of the Ida–Dactyl asteroid system. Credit: NASA.

flyby enabled the density of that asteroid to be constrained to lie between 1.4 to 4 grams per cubic centimeter [11]. The large uncertainties arose as the orbit of Dactyl could not be pinned down very precisely.

Since that time there have been additional flybys of asteroids by spacecraft on their way to other solar system destinations. Figure 1.4 shows a compilation of all spacecraft-imaged small bodies. Despite such flybys being relatively inexpensive in terms of propellant, there are operational costs associated with such flybys which often make missions reluctant to incorporate them into their plans. Thus, there have been only 4 asteroid flybys of note since the initial Galileo flybys: the New Millennium DeepSpace 1 Spacecraft flyby of asteroid Braille, the NEAR-Shoemaker Spacecraft flyby of asteroid Mathilde, and the recent flybys of the Rosetta Spacecraft of asteroids Steins and Lutetia. Of these flybys, the most notable are the NEAR flyby of Mathilde and the Rosetta flyby of Lutetia, as the total mass of these asteroids were determined [199, 164, 122], although only a partial shape/volume of these asteroids were found. The total density of Mathilde was bounded and found to be on the order of 1.3 g/cm^3 [199], a surprisingly low value, indicating for the first time that asteroids can have a high level of porosity – a fact that has been repeatedly reconfirmed.

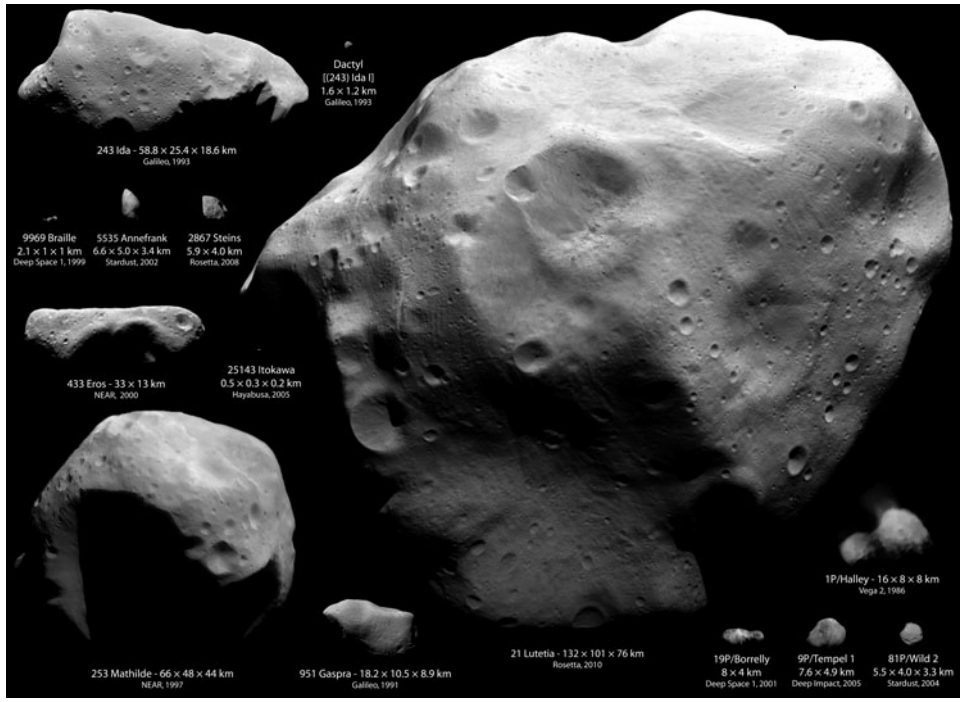


Fig. 1.4 Compilation of asteroids and comets imaged by spacecraft. Credits: Montage by Emily Lakdawalla. Ida, Dactyl, Braille, Annefrank, Gaspra, Borrelly: NASA/JPL/Ted Stryk. Steins: ESA/OSIRIS team. Eros: NASA/JHUAPL. Itokawa: ISAS/JAXA/Emily Lakdawalla. Mathilde: NASA/JHUAPL/Ted Stryk. Lutetia: ESA/OSIRIS team/Emily Lakdawalla. Halley:: Russian Academy of Sciences/Ted Stryk. Tempel 1: NASA/JPL/UMD. Wild 2: NASA/JPL.

1.2.4 Rendezvous

To date there have only been three spacecraft rendezvous missions to asteroids, the National Aeronautics and Astronautics Administration's (NASA's) Near Earth Asteroid Rendezvous (NEAR) mission to asteroid Eros, the Japanese Exploration Agency's (JAXA) Hayabusa mission to asteroid Itokawa, and most recently NASA's DAWN mission to asteroids Vesta and Ceres. Both the NEAR and Hayabusa missions successfully characterized the external morphology of their asteroids and measured some element of their mass and gravity field. The DAWN mission is just achieving its rendezvous with Vesta as of this writing, and will provide significant scientific returns on that asteroid's mass, gravity and morphology. Following the Vesta phase of the mission, the spacecraft will go on to explore the asteroid Ceres.

The NEAR Mission

The NEAR mission was launched in 1996 and achieved rendezvous with Eros on February 14, 2000. Its subsequent mission lasted for over one year, during which

time it remained in a bound orbit about the asteroid, seen in Fig. 1.5. During this period high-resolution images of the surface were obtained, a shape model created, and its gravity field and spin estimated, among other measurements. The total density for this body was determined to be 2.4 g/cm^3, and its internal mass distribution was found to be nearly homogeneous. At the end of its mission, the spacecraft was landed on the surface of the body, where it remained in contact with the Earth for a few weeks before it was no longer observable and contact was lost. At its next apparition an unsuccessful attempt was made to contact the spacecraft. The entire mission design for the NEAR spacecraft at Eros involved novel elements of mission design not needed for any previous orbital mission, due to the asteroid's highly irregular shape, rapid rotation period of 5.27 hours and obliquity near 90°. Although Eros is a Near-Earth asteroid, it is uncharacteristic of this population due to its large size, with its mean diameter being greater than 15 km.

Fig. 1.5 NEAR spacecraft image of asteroid Eros. Credit: NASA/JHU-APL.

The Hayabusa Mission

The Hayabusa mission was launched in 2003 and achieved rendezvous with Itokawa in mid-September 2005, shown in Fig. 1.6. Its subsequent mission at that asteroid lasted until November 25, at which time the spacecraft was subject to an anomaly and left the vicinity of Itokawa. During its short stay at the asteroid, Hayabusa was able to characterize the asteroid surface in detail, leading to a precise shape model being developed. The total mass of the asteroid was estimated, leading to a density

estimate of approximately 2 g/cm^3. Higher order gravity field estimations have not been carried out for this body as of yet, although relevant data for such determinations is in existence. The spacecraft achieved touchdown on the surface of the asteroid twice, and collected samples of regolith in its sample chamber. Following the anomaly of November 25, 2005, the spacecraft was brought back under control and navigated back to the Earth where it had a successful re-entry into the Earth's atmosphere over Australia in June 2010, at which time the sample container was retrieved. During its close proximity phase at the asteroid, the spacecraft achieved a bound orbit about the asteroid only a few times for only brief periods of time. The majority of the time the spacecraft "hovered" over the asteroid surface, meaning that it used its thrusters to actively maintain its location a distance of several kilometers from the surface of the asteroid. Only for the trial and actual descents to the surface did the spacecraft come close enough to the asteroid to actually achieve orbit for brief periods.

Fig. 1.6 Hayabusa spacecraft images of asteroid Itokawa. Credit: JAXA/ISAS.

1.2.5 Future Missions

Beyond the current DAWN mission there are future asteroid exploration missions still being developed and planned. Most recently NASA has selected the OSIRIS-REx asteroid sample return mission as the next New Frontiers mission. This mission will visit the asteroid 1999 RQ36 in 2020, characterize and sample its surface, and return its sample to Earth in 2022. This target asteroid is relatively small and appears to be similar to the class of binary asteroid primaries, with a relatively rapid rotation rate and a strongly oblate shape. The body has a total extent of ∼500 meters, and its exploration will represent an important next-step in the understanding of asteroids in general.

Other space exploration agencies also have plans for future missions to asteroids. Perhaps the most highly developed proposal is JAXA's Hayabusa 2 mission to asteroid 1999 JU3. This mission has been approved and is scheduled for launch in July 2014. It is a sample return mission but will spend an extended period of time at the asteroid (approximately 1.5 years). It will return to Earth in December 2020, and thus would bring its sample home prior to the return of the OSIRIS-REx sample. Also being proposed to the European Space Agency (ESA) is the Marco Polo-R mission, which is a sample return mission to a binary asteroid. This proposal is currently in assessment, with a decision for funding to be announced in the future. In addition to these specific plans, there are always several other compelling missions to asteroids being proposed to various space agencies. This is driven, in part, by the accessibility of the NEO population to spacecraft.

1.3 A Brief History of Comet Exploration

The history of comet observations predates that of asteroids for naked-eye and telescopic observations, yet has lagged behind asteroid exploration in terms of rendezvous. The lack of rendezvous explorations will soon be remedied by the Rosetta mission to comet Churyumov-Gerasimenko. Rosetta is scheduled to enter orbit about the comet in May 2014, and will have a 2-year mission of exploration. This will include observations from the spacecraft and the deployment of a lander to the comet surface.

1.3.1 Ground-Based Observations

Due to the observability of comets with the naked eye from the ground, the record of their observations goes back to the earliest recorded histories. The societal impacts of comets have been extremely significant, and have often been correlated with social upheavals and natural catastrophes by active imaginations due to their spectacular presentation in the sky. Thus, to provide a review of observations goes well beyond the scope of the current book, but has been considered from a scientific perspective in earlier texts such as the book by D.K. Yeomans [198]. Ground-based observations of comets continue to play a key role in their scientific understanding, yet such observations usually provide few clues to even the most simple physical aspects of comets such as their shape, size and rotation state. Exceptions occur for particularly well-studied comets such as Tempel 1 [183]. Further, radar observations of comets have provided some important information on their environment and other physical properties. Two specific examples are the radar observations of comet Hyakutake, which was the first to observe clouds of centimeter-sized particles in orbit about that body (recently seen optically by the Deep Impact spacecraft's flyby of comet Hartley 2). Observations of comet Tuttle [62] provided the first detailed shape model of a comet, showing a strongly bifurcated shape.

1.3.2 Spacecraft Observations

On the other hand, the history to date of spacecraft exploration of comets is much shorter and, until the Rosetta spacecraft has its rendezvous, will consist only of flyby missions. The first series of spacecraft observations of a comet were the Halley flotilla in 1986. This international group of spacecraft included craft from the European Space Agency, the Japanese space agency, and the then-Soviet Union's space agency. These flybys provided significant details about that comet's rotation state, composition, and shape, but could not provide mass estimates or even a fully three-dimensional shape and uncontroversial spin state estimate.

Following these observations the next significant comet observations from a spacecraft was the comet Borrelly flyby by the DeepSpace 1 spacecraft in 2001, which provided the first evidence that comet shapes could be strongly bifurcated. Following this was the Stardust flyby of comet Wild 2 in 2004, which showed a comet with a much different surface morphology than comet Borrelly and eventually provided the first direct sample return from a comet's coma.

In 2005 the Deep Impact spacecraft had its interaction with comet Tempel 1, impacting it with a 370-kg impactor designed to form a crater on that body to determine the overall strength of a cometary surface. Due to the unforeseen optically dense ejecta field raised by that impact, that mission was not able to directly image the crater formed on the surface. The mission also provided the highest-resolution imaging, to date, of a comet surface, raising a host of interesting scientific questions about the nature of comets. The recent flyby of the Stardust spacecraft of Tempel 1 in February 2011 was able to image the impact site of the original Deep Impact mission, although at a lower resolution than the initial mission, and identified the crater made by the Deep Impact impactor.

Finally, the Deep Impact spacecraft carried out a flyby of comet Hartley 2 in November 2010, obtaining images of that body and confirming the presence of significant amounts of "snow" in orbit about that nucleus. Figure 1.7 shows proportionate images of comets Tempel 1 and Hartley 2. The next significant advance in our understanding of comets will occur when the Rosetta spacecraft has its rendezvous with comet Churyumov-Gerasimenko in 2014. Although low-resolution shape models of this comet are available and preliminary orbit analyses have been performed, these analyses will be supplanted by the actual data after rendezvous.

It is relevant to ask what the difference between comets and asteroids are in terms of spacecraft and particle motion in their vicinity. There are three differences expected, although two of them are more minor. The first of the minor differences is that comets are expected to be less dense than asteroids by a factor of 2 or more, due to their high content of volatile ices and presumed porous structure. Estimates of comet densities have been estimated to be as low as 0.3 grams per cubic centimeter [138]. The second minor detail is that comets are expected to more commonly lie in excited rotation states, although there are known asteroids that are also in such states. The major difference between asteroids and comets is the outgassing present at comets, creating a gaseous environment that will potentially provide significant non-gravitational forces acting on the spacecraft. There has been some disagreement over the nature of these outgassing fields, and thus, while they

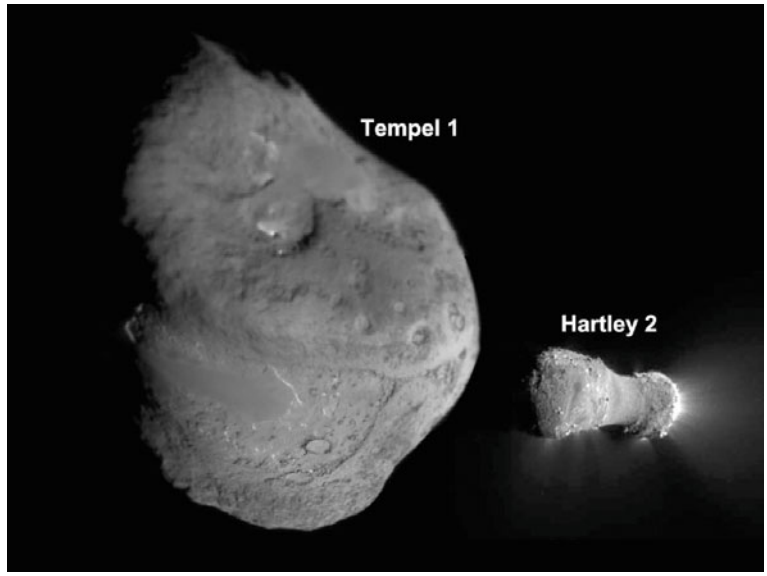


Fig. 1.7 Comparison image of the nuclei of comets Tempel 1 and Hartley 2. Credit: NASA.

are discussed in this book, the models that are used are likely to be significantly updated by results from the Rosetta mission to comet Churyumov-Gerasimenko.

1.4 A Brief History of Planetary Satellite Exploration

1.4.1 Early Ground-Based Observations

The history of planetary satellite observations extend to primeval times, given the close proximity and stark observability of the Earth's planetary satellite, the Moon. However, as will become clear when the dynamical environment of planetary satellite orbiters are discussed later, the Earth's moon is unique among planetary satellites due to its large mass relative to the Earth, and is rightly treated as a problem in its own right (see [179]).

Besides the Moon, the start of planetary satellite exploration occurs with Galileo's discovery of the four largest Jovian satellites in 1610: Io, Europa, Ganymede and Callisto. The discovery of Saturn's largest moon, Titan, followed in 1655 by Huygens. In the subsequent few centuries all of the major moons of the large outer planets were observed and cataloged. The Martian moons, Phobos and Deimos, were discovered by Hall in 1877, and represented a new class of objects, small planetary satellites. Whereas the major planetary satellites of the gas giant planets are all spheroidal in shape, the Martian moons have quite strongly non-spherical shapes reminiscent of asteroids. Indeed, current speculation is that these moons are a captured asteroid in orbit about Mars, although the detailed mechanics of how such an event could have occurred is not understood.

1.4.2 Flyby Spacecraft Observations

The largest advance in our knowledge of planetary satellites in the solar system occurred with the Pioneer and Voyager spacecraft flybys of the outer planets. These missions allowed, for the first time, the masses of the larger satellites to be determined and provided the first detailed, high-resolution observations of the surfaces of these bodies. They also led to the discovery of additional, small planetary moons. Perhaps the most significant realizations from these missions were that these planetary satellites exhibited a wide range of surface morphologies and levels of activity. From the active volcanism on Io to the significant differences between the smooth surface of Europa, battered surface of Callisto and the dense atmosphere of Titan.

1.4.3 Spacecraft Rendezvous Observations

On the heels of the successful flybys of the Jovian and Saturnian systems, NASA enacted plans to visit Jupiter and Saturn with dedicated rendezvous spacecraft. The Galileo spacecraft visited Jupiter and was in orbit in this system from 1995 until 2003 when it was sent on an impact trajectory with Jupiter. Recall that it was the Galileo spacecraft that also carried out the first high-resolution imaging observations of asteroids during its flight to Jupiter. Throughout its mission in the Jovian system the spacecraft provided significant insights into the diverse population of Jovian moons, and has fundamentally altered and shaped current scientific thinking on the formation and evolution of this system.

The counterpart of the Galileo spacecraft at Saturn is the Cassini spacecraft, part of the Cassini-Huygens mission. This craft entered orbit about the Saturnian system in 2004, and is still actively exploring this system. Current plans have the mission extending through 2017. One of the main focuses of this mission was ESA's Huygens probe sent to the surface of the largest Saturnian satellite, Titan. From its observations, combined with Cassini radar observations of the moon, it was determined that the surface and atmosphere of Titan is active with lakes, precipitation and varying winds. The Cassini spacecraft has also made detailed studies of many of the other moons of Saturn, with some of the most fascinating being the small satellites embedded in or shepherding the rings, the satellite Hyperion which is known to be in a chaotic rotation state, and the larger satellite Enceladus which is now speculated to harbor an ocean under its surface.

In addition to the satellites of the outer planets, many spacecraft have also observed the Martian moons of Phobos and Deimos. Most recently, NASA's Mars Reconnaissance Observer mission and ESA's Mars Express missions have both imaged these satellites and further constrained their mass. The history of such observations traces back to NASA's Viking mission, which produced the first high-resolution images of these moons.

Shown in Figs. 1.8 and 1.9 are images of several planetary satellites taken from spacecraft rendezvous missions. Figure 1.8 shows a number of smaller satellites, specifically bodies that have non-spheroidal shapes. Figure 1.9 shows a number of larger satellites, specifically those that have a more spheroidal shape.

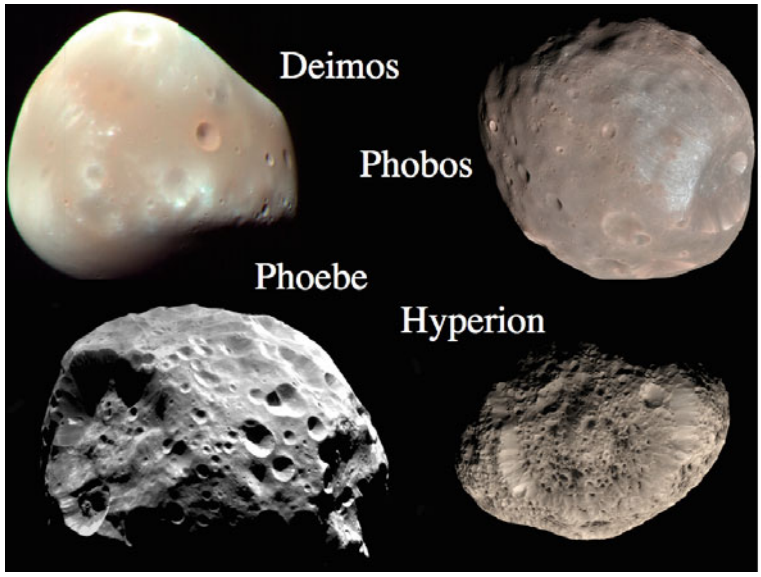


Fig. 1.8 Small and irregularly shaped planetary satellites (not shown to scale). Phobos and Deimos are moons of Mars. Phoebe and Hyperion are both satellites of Saturn. Phoebe is the largest irregular satellite of Saturn and Hyperion is in a non-uniform rotation state. Credit: NASA/JPL-Caltech.

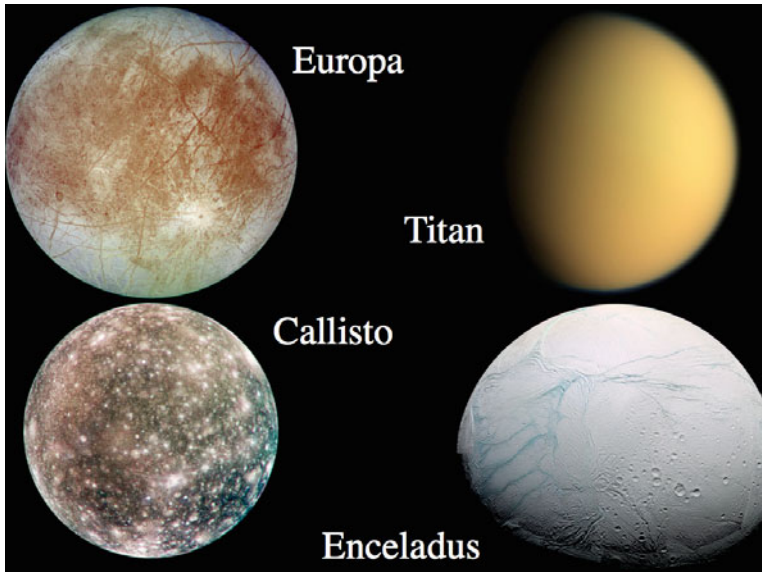


Fig. 1.9 Large and spheroidally shaped planetary satellites (not shown to scale). Europa and Callisto are Jovian satellites. Europa has an ice-covered surface which resurfaces itself frequently while Callisto's surface is extremely old and covered with craters. Titan and Enceladus are Saturnian satellites. Titan has a complex atmosphere that shrouds its surface while Enceladus has complex surface features. Credit: NASA/JPL-Caltech.

1.4.4 Modern Ground-Based Observations

In tandem with the spacecraft exploration of the satellite systems of the gas giant planets, there has been continued ground-based observation activities. The most significant observations have been of the so-called irregular satellites of the outer planets. These are small bodies that have been captured into relatively loosely bound orbits about the planets. These objects often have high inclinations (including retrograde) and eccentricities. They are currently thought to be primitive bodies captured during an early stage of planetary formation, with a presumed diffuse and gaseous disk around the planets aiding in their capture. Nearly one hundred of these satellites have been discovered since the late 1990s, with all four of the outer planets having at least some irregulars in orbit about them.

In addition to optical observations, radar observations of the Galilean planetary satellites have also been carried out in recent decades. These observations primarily use the NASA Deep Space Network's Goldstone Planetary Science Radar instrument, a 70-meter antenna at the Goldstone complex, and the Arecibo radio antenna. Unlike the asteroid observations made with these same instruments, which can produce global shape models and spin states, the radar observations of planetary satellites are used primarily to determine surface reflectance properties to understand the structure of the surface.

1.4.5 Future Spacecraft Missions

Finally, it is significant to discuss the possible future exploration of planetary satellites. There has been perennial interest in sending spacecraft to select moons of Jupiter and Saturn to study and determine the interior geophysical structure of these satellites. The Jovian moon Europa has been considered a high priority target since the late 1990s, and has seen several different mission proposals go through different stages of advancement. Unfortunately, all of these proposals have, to date, been stopped. The scientific priority of a Europa mission is still high, although it competes with many other worthy scientific endeavors. Also, the technical complexity and cost of a Europa orbiter mission has been long recognized as a significant issue. The orbital dynamics complexity of such missions, discussed later in this book in Chapters 16 and 17, are but one of the challenges such a mission must address.

More recently, with the discovery of cryovolcanism on Enceladus, that moon has also been proposed as a target for space missions. As will be shown later in the text, the orbital dynamics challenges at Enceladus are even greater than those faced by a Europa orbiter. However, the ultimate decision on any specific mission target should also be based on a mix of the science and engineering. These debates have yet to take place.

In addition to these two mission ideas, there are perennially many proposals that focus on different aspects of planetary satellites. Of most significance is the Russian Phobos-Grunt mission which was launched in November 2011, although it has had subsequent issues that remain to be resolved. This mission will visit and

perform close proximity operations in the vicinity of the Martian moon Phobos, also placing a lander on the surface of that moon. The discussion in Chapter 18 is specifically relevant for such a mission. In addition to this currently planned mission, missions to the Martian moons have been frequently proposed and will likely continue to be proposed in the future.

1.5 Notation and Definitions

This is a technically oriented book, and thus it is necessary to introduce the mathematical terminology and notation necessary for the following the discussions. As the theory is focused on relatively specific types of dynamical systems, our basic mathematical terminology and notation can be introduced here and used throughout the book. Due to the variety of systems discussed, we do not try to define a completely uniform notation and reserve the right to introduce specific details and updates as necessary. The following discussion is more focused on the overriding generalities.

1.5.1 Vectors, Dyadics and Higher-Order Tensors

In this text a tensor is always a Cartesian tensor, meaning that it is defined and specified with respect to an orthogonal set of coordinates. Tensors are denoted as upper- or lower-case latin letters with subscripts to denote the rank and different entries. A tensor's dimension is the number of dimensions each entry contains while its rank is the number of entries that it has. The traditional notation of subscripts will be used to denote a tensor. Thus, an n-dimensional tensor of rank m will be denoted as $A_{i_1 i_2 \ldots i_m} \in \mathbf{R}^{n \times m}$ where $i_j = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, m$. Note that a tensor can be a function of additional items, such as location relative to some object. Tensors of like dimension and rank can be added and subtracted, term by term, but the rules for multiplication are more involved. The main operation needed in this text is a generalization of the dot product (also called contraction) which allows two tensors of the same dimension but potentially different rank to be multiplied together.

Consider two tensors, A_{ijk} and B_{lm}. A new tensor can be defined with the operation

$$C_{ijm} = \sum_{k=1,n} A_{ijk} B_{km} \tag{1.1}$$

$$= A_{ijk} B_{km} \tag{1.2}$$

Note that two of the entries in the above equations have the same index, meaning that these terms are multiplied by each other and summed. In the second line this tensor product is restated using the Einstein summation convention, which automatically sums the terms with a similar index. Which index multiplies which is defined by the physics and mathematics of the operation where this occurs. The most common occurrence of this notation is when a number of rank-1 tensors are

multiplied with a tensor of higher rank. Now consider A_{ijk} again and a rank-1 tensor a_l. Then the following tensors have the denoted ranks:

$$B_{ij} = A_{ijk}a_k \text{ Rank 2} \tag{1.3}$$

$$B_i = A_{ijk}a_j a_k \text{ Rank 1} \tag{1.4}$$

$$B = A_{ijk}a_i a_j a_k \text{ Rank 0} \tag{1.5}$$

where a rank-0 tensor is a real number. In this text tensors of rank higher than 2 are only used for higher-order Taylor series expansions. Finally, note that tensors can be defined independent of any specific coordinate frame, although when they are to be realized in terms of specific numbers a frame must be specified.

Tensors of rank 1 and 2 play a special role in mechanics and are called vectors and dyadics, respectively. Using our tensor notation, a vector is a_i while a dyadic is A_{ij}. Special notation is used to denote these quantities along with some special operations. Vectors will generally be represented as lower-case, bold-faced letters and dyadics as upper-case, bold-faced letters and symbols. Thus, equivalent notations for vectors are a_i and $\boldsymbol{a}$ and for dyadics A_{ij} and $\boldsymbol{A}$. Exceptions to this notation as written here do occur.

Vector Operations

The tensor contraction operation plays an important role for vectors. When applied to a vector itself it defines the length-squared of a vector, which is a scalar quantity. Denote this as $a^2 = a_i a_i = \boldsymbol{a} \cdot \boldsymbol{a}$, and thus the tensor contraction applied to vectors is equivalent to the dot product. This operation is used to define the vector magnitude, or vector length, operation

$$|\boldsymbol{a}| = \sqrt{\boldsymbol{a} \cdot \boldsymbol{a}} \tag{1.6}$$

where a non-bolded letter will be used to denote the magnitude of a vector, $a = |\boldsymbol{a}|$.

Vectors that have a length equal to unity are called unit vectors and usually denoted as a lower-case bold-faced letters with a caret on top, or $\hat{\boldsymbol{u}}$, defined such that $|\hat{\boldsymbol{u}}| = 1$. Any vector with non-zero length can be transformed into a unit vector as $\hat{\boldsymbol{a}} = \boldsymbol{a}/|\boldsymbol{a}|$. The dot product is also defined for different pairs of vectors, with equivalent expressions $a_i b_i$ and $\boldsymbol{a} \cdot \boldsymbol{b}$. If the dot product of two vectors is equal to zero, this means that these vectors are orthogonal to each other, i.e., that they have zero projections onto each other.

If a set of unit vectors are chosen such that they are all mutually orthogonal and their number equals the dimension of the vector space, then any vector can be represented as a linear combination of these unit vectors. A common set are the orthogonal unit vectors in a 3-D space denoted as $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$ and $\hat{\boldsymbol{z}}$, where each of these unit vectors lives in $\mathbf{R}^3$. Then, given an arbitrary vector $\boldsymbol{a} \in \mathbf{R}^3$, this vector can always be expressed relative to an orthogonal set of unit vectors as $\boldsymbol{a} = a_x\hat{\boldsymbol{x}} + a_y\hat{\boldsymbol{y}} + a_z\hat{\boldsymbol{z}}$, where $a_x = (\boldsymbol{a} \cdot \hat{\boldsymbol{x}})$, etc. If an implied set of coordinates is used the vector can be written as a column or row vector of the general form $[a_x, a_y, a_z]$, where the unit vectors are implied.

Dyadic Operations

Dyadics can also be described using a set of unit vectors. Assume a general dyadic $\boldsymbol{A} \in \mathbf{R}^{3\times3}$, then its general form as expressed with respect to our orthogonal basis is $\boldsymbol{A} = A_{xx}\hat{\boldsymbol{x}}\hat{\boldsymbol{x}}+A_{xy}\hat{\boldsymbol{x}}\hat{\boldsymbol{y}}+A_{xz}\hat{\boldsymbol{x}}\hat{\boldsymbol{z}}+A_{yx}\hat{\boldsymbol{y}}\hat{\boldsymbol{x}}+A_{yy}\hat{\boldsymbol{y}}\hat{\boldsymbol{y}}+A_{yz}\hat{\boldsymbol{y}}\hat{\boldsymbol{z}}+A_{zx}\hat{\boldsymbol{z}}\hat{\boldsymbol{x}}+A_{zy}\hat{\boldsymbol{z}}\hat{\boldsymbol{y}}+A_{zz}\hat{\boldsymbol{z}}\hat{\boldsymbol{z}}$. The "product" of two unit vectors, $\hat{\boldsymbol{x}}\hat{\boldsymbol{y}}$, is called a dyad and, in column and row vector notation is equivalent to the outer product of two vectors (i.e., a row times a column). A dyadic can be written with an implied basis as a matrix, for the above example as a 3×3 matrix

$$[A] = \begin{bmatrix} A_{xx} & A_{xy} & A_{xz} \\ A_{yx} & A_{yy} & A_{yz} \\ A_{zx} & A_{zy} & A_{zz} \end{bmatrix} \tag{1.7}$$

Similar to matrices, dyadics can have their transpose taken, which consists of swapping the order of all the dyads while keeping the coefficients fixed. Thus, if $\boldsymbol{A} = A_{xy}\hat{\boldsymbol{x}}\hat{\boldsymbol{y}}$, then $\boldsymbol{A}^T = A_{xy}\hat{\boldsymbol{y}}\hat{\boldsymbol{x}}$, etc. There is also a unity dyadic corresponding to the identity matrix, expressed as $\boldsymbol{U} = \hat{\boldsymbol{x}}\hat{\boldsymbol{x}}+\hat{\boldsymbol{y}}\hat{\boldsymbol{y}}+\hat{\boldsymbol{z}}\hat{\boldsymbol{z}}$. This dyadic plays the same role as the identity matrix with the general property $\boldsymbol{U} \cdot \boldsymbol{a} = \boldsymbol{a} \cdot \boldsymbol{U} = \boldsymbol{a}$. Finally, the determinant operator can also be applied to dyadics, where the notation $\|\boldsymbol{A}\|$ denotes the determinant of a dyadic or its equivalent matrix.

The dot product, or contraction, operator can be applied to dyadics as well. If the product is between two dyadics a new dyadic is formed, while if the operation is between a dyadic and a vector a new vector is formed. In tensor notation, given A_{ij} and B_{kl} these can be combined to form $C_{ik} = A_{ij}B_{jk}$. The equivalent dyadic operation is $\boldsymbol{C} = \boldsymbol{A} \cdot \boldsymbol{B}$, and if using an implied basis this can also be expressed as a matrix multiplication, $[C] = [A][B]$. For products between dyadics and vectors, the same set of examples yields the equivalent statements $c_i = A_{ij}b_j$, $\boldsymbol{c} = \boldsymbol{A} \cdot \boldsymbol{b}$, and $[c] = [A][b]$.

Dyadics can also be used to define the other important vector operation, the cross product. It can be shown that the cross product between two vectors, $\boldsymbol{a}$ and $\boldsymbol{b}$ in $\mathbf{R}^3$, can be expressed using the dot product between a certain dyadic and one of the vectors. Notationally this is denoted as $\boldsymbol{a} \times \boldsymbol{b} = \tilde{\boldsymbol{a}} \cdot \boldsymbol{b} = \boldsymbol{a} \cdot \tilde{\boldsymbol{b}}$. Assuming that the vector has been expressed relative to a basis $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, $\hat{\boldsymbol{z}}$, the cross product operation is

$$\tilde{\boldsymbol{a}} = a_x\left(\hat{\boldsymbol{z}}\hat{\boldsymbol{y}} - \hat{\boldsymbol{y}}\hat{\boldsymbol{z}}\right) + a_y\left(\hat{\boldsymbol{x}}\hat{\boldsymbol{z}} - \hat{\boldsymbol{z}}\hat{\boldsymbol{x}}\right) + a_z\left(\hat{\boldsymbol{y}}\hat{\boldsymbol{x}} - \hat{\boldsymbol{x}}\hat{\boldsymbol{y}}\right) \tag{1.8}$$

Note that the cross product dyadic is skew-symmetric, as is expected for the cross product. Thus $\tilde{\boldsymbol{a}}^T = -\tilde{\boldsymbol{a}}$, and the additional equivalent statements of the cross product follow: $\boldsymbol{a} \times \boldsymbol{b} = -\tilde{\boldsymbol{b}} \cdot \boldsymbol{a} = -\boldsymbol{b} \cdot \tilde{\boldsymbol{a}}$. The cross product dyadic has a few additional, special properties stated below

$$\tilde{\boldsymbol{a}} \cdot \tilde{\boldsymbol{b}} = \boldsymbol{b}\boldsymbol{a} - (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{U} \tag{1.9}$$

$$\widetilde{\tilde{\boldsymbol{a}} \times \boldsymbol{b}} = \boldsymbol{b}\boldsymbol{a} - \boldsymbol{a}\boldsymbol{b} \tag{1.10}$$

At a deeper level, the cross product can actually be shown to arise from a rank-3 skew-symmetric tensor, but the above definition suffices for our purposes.

Scalars

Finally, a scalar is a rank 0 tensor, meaning that it has a dimensionality of 1, although it can still be a function of other quantities, such as position, but does not require a coordinate frame to specify its value. Examples of scalars include the energy integral, the Lagrangian function, the Hamiltonian function, and the potential function to name a few. Scalars are usually denoted as unbolded capital letters and indicate the terms of which they are functions. For example, the gravitational potential function of a spherical body is expressed as $U(\boldsymbol{r}) = \mu/|\boldsymbol{r}|$, where U is a scalar quantity, even though its value is a function of a position vector.

Gradient Operators

The gradient operator will raise the rank of a scalar, vector or tensor in general. Assume a tensor of some rank that is a function of another tensor (in particular, a vector). Then the term-by-term partial derivative of that tensor with respect to the vector will result in a tensor of a higher rank. Specifically, assume $A_{ij}(\boldsymbol{r})$ is a rank-2 tensor that is a function of a rank-1 vector. Then the rank-3 tensor gradient is defined as $A_{ij,k} = \partial A_{ij}/\partial r_k$. For a second example consider the gradient of a scalar with respect to a vector using our other notation. Let $U(\boldsymbol{r}) = \mu/|\boldsymbol{r}|$. Then the gradient of U with respect to $\boldsymbol{r}$ will define a vector. To carry this out note that $|\boldsymbol{r}| = \sqrt{\boldsymbol{r}\cdot\boldsymbol{r}}$, that the chain rule should be used, and that the partial of a vector with respect to itself equals the unity dyadic.

$$\frac{\partial U}{\partial \boldsymbol{r}} = -\frac{1}{2}\frac{\mu}{|\boldsymbol{r}|^3}\frac{\partial \boldsymbol{r}\cdot\boldsymbol{r}}{\partial \boldsymbol{r}} \tag{1.11}$$

$$= -\frac{1}{2}\frac{\mu}{|\boldsymbol{r}|^3}\left(\boldsymbol{U}\cdot\boldsymbol{r} + \boldsymbol{r}\cdot\boldsymbol{U}\right) \tag{1.12}$$

$$= -\frac{\mu}{|\boldsymbol{r}|^3}\boldsymbol{r} \tag{1.13}$$

Time Derivatives of Vectors

Let $\boldsymbol{r}$ denote a three-dimensional position vector with general entries r_j defined in an orthogonal reference frame. The expression $\dot{\boldsymbol{r}}$ denotes the time derivative of the vector $\boldsymbol{r}$, or $d\boldsymbol{r}/dt = \dot{\boldsymbol{r}}$. When taking the time derivative of a vector, it is important to note whether the vector is expressed relative to an inertially fixed frame or relative to a frame which rotates relative to inertial space. The general assumption is that all time derivatives take place relative to an inertially oriented frame unless otherwise stated. If the frame is rotating with respect to an inertial frame, say with an angular velocity $\boldsymbol{\Omega} \in \mathbf{R}^3$, then the time derivative of the quantity with

respect to an inertial (non-rotating) frame can be related to the time derivative in the rotating frame via the transport equation

$$\dot{\boldsymbol{r}}_I = \dot{\boldsymbol{r}}_R + \widetilde{\boldsymbol{\Omega}} \cdot \boldsymbol{r} \tag{1.14}$$

where the subscripts I and R denote time derivatives with respect to inertial and rotating frames, respectively. The angular velocity and the position can be expressed with respect to either frame, which frame the time derivative is being taken with respect to will generally be clear from the context. Applying the formula a second time yields a formula relating the acceleration vector in an inertial and a rotating frame

$$\ddot{\boldsymbol{r}}_I = \ddot{\boldsymbol{r}}_R + \dot{\widetilde{\boldsymbol{\Omega}}} \cdot \boldsymbol{r} + 2\widetilde{\boldsymbol{\Omega}} \cdot \dot{\boldsymbol{r}} + \widetilde{\boldsymbol{\Omega}} \cdot \widetilde{\boldsymbol{\Omega}} \cdot \boldsymbol{r} \tag{1.15}$$

1.5.2 Functions

Finally, some of our functional notation and assumptions must be defined. Unless otherwise stated, the uniform assumption is that all of the functions encountered in our analysis are analytic in all of their arguments, meaning that their higher-order partial and time derivatives are well-defined. Exceptions to these assumptions arise when spacecraft controls are considered, but these are only dealt with in isolated portions of the book. Thus, besides these few exceptions, the functions considered in this book are C_∞ in their arguments.

Functions can be scalars, vectors, or higher-order tensors, depending on the application. A general format for our most common example, a vector function, is

$$\boldsymbol{f}(\boldsymbol{x}, t) = \{\boldsymbol{x} \times t \in \mathbf{R}^n \times \mathbf{R} \rightarrow \boldsymbol{f} \in \mathbf{R}^m\} \tag{1.16}$$

where $\boldsymbol{x}$ is n-dimensional, t is a scalar quantity, and $\boldsymbol{f}$ is m-dimensional. In general the dimension of the arguments and the function will be evident from the context.

1.5.3 Differential Equations and Solutions

Of particular interest to us are ordinary differential equations that are derived from physical principles. These can generally be placed into the form

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, t) \tag{1.17}$$

where the dot signifies a total time derivative and $\boldsymbol{f}$ has the same dimension as $\boldsymbol{x}$ and is C_∞ in its arguments. The general assumption is that the ordinary differential equations encountered are well-defined and have unique solutions (i.e., are Lipschitz).

A solution to a set of ordinary differential equations is a function $\boldsymbol{x}(t) = \boldsymbol{\phi}(t; \boldsymbol{x}_o, t_o)$ defined such that

$$\frac{\partial \boldsymbol{\phi}}{\partial t} = \boldsymbol{f}(\boldsymbol{\phi}(t), t) \tag{1.18}$$

$$\boldsymbol{x}_o = \boldsymbol{\phi}(t_o; \boldsymbol{x}_o, t_o) \tag{1.19}$$

Note that the dependency of the solution on the so-called "epoch state" $\boldsymbol{x}_o$ is explicitly called out, as this state can be used to define the solution. This is done as one of the most important quantities associated with a solution is its gradient with respect to this epoch state, $\partial \boldsymbol{\phi} / \partial \boldsymbol{x}_o$, which forms the "state transition" matrix or dyadic.

2. Modeling Small Body Environments

Before motion in the small body environment is studied it is first necessary to define this environment. There exist many different scientific taxonomies for asteroids and comets, and we refer the interested reader to the "Asteroids and Comets" series of texts for a larger view of these bodies [48, 102, 14, 197, 41]. As our focus is on the dynamics of vehicles at these bodies, only a sub-set of all the properties of these bodies needs to be specified. Thus in the following we describe the orbits of small bodies in the solar system, the rotations of these bodies, their gravitational fields and shapes, and the forces that act on a spacecraft or particle in their vicinity.

2.1 Heliocentric and Planetocentric Orbits

Small bodies orbit the sun on trajectories that are dominated by the solar gravitational attraction. The implicit assumption we make in this book is that the orbital motion of a small body can be described accurately by the two-body problem solution. This is an assumption, and is technically not correct as small bodies are subject to additional perturbations that can have significant effects on their motion. These additional perturbations arise from the gravitational attraction of the planets, Jupiter in particular, and from non-gravitational forces acting on the bodies, usually driven by solar irradiation. However, except for some exceptional cases, the assumption that a small body follows an essentially two-body orbit over a timespan of a few orbit periods about the Sun is a good one.

In other situations, when a small body has a close approach to a planet, the interaction between the small body and the planet during the time when they are close to each other can generally be modeled as a hyperbolic orbit. These situations are also important for understanding the effect of a close planetary passage on orbiting particles about a small body. The "interpolation" between these extremes would have the small body subject to two gravitational attractions simultaneously, such as the Trojan asteroids which are attracted by the Sun and Jupiter.

Finally, we make similar assumptions when considering the motion of planetary satellites. For detailed models it is necessary to incorporate perturbations from the

D.J. Scheeres, *Orbital Motion in Strongly Perturbed Environments*, Springer Praxis Books,
DOI 10.1007/978-3-642-03256-1_2, © Springer-Verlag Berlin Heidelberg 2012

Sun and, in some cases, other satellites about the planet. However, for the general description of their environments over the relatively short time scale considered here the use of a two-body orbital motion suffices.

The initial classifications we focus on are based on the orbits of these bodies. Figures 2.1 and 2.2 graphically display the range of asteroid and comet orbits, and a snapshot of their location in the solar system. The main message from these representations is that the small bodies of the solar system are distributed throughout the solar system, with specific areas of concentration but almost no regions completely devoid of small bodies. Those few that exist arise due to specific perturbations of Jupiter on their orbits, with the best-known regions being the Kirkwood gaps. Figure 2.3 shows a population distribution of asteroids in the Main Belt, clearly indicating the lack of asteroids in regions that are in resonance with Jupiter.

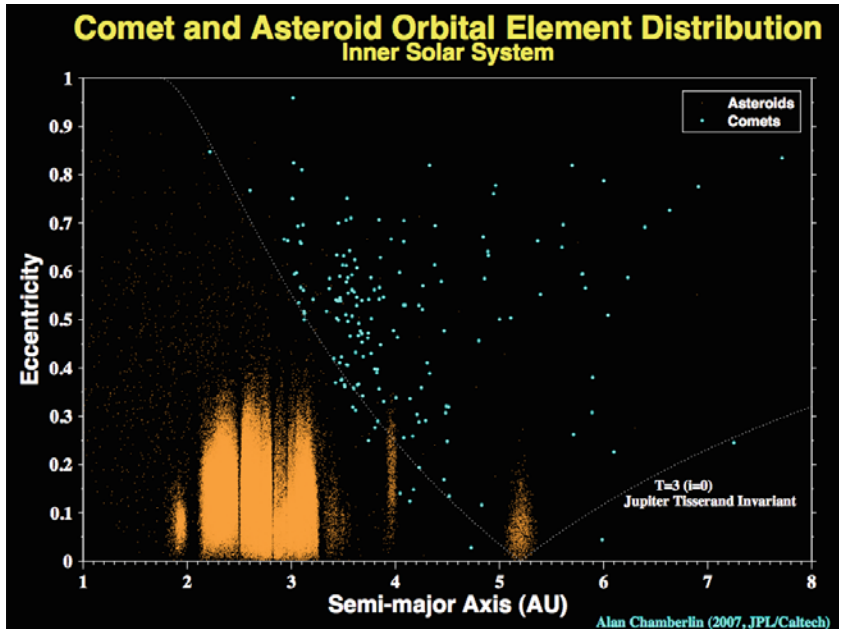


Fig. 2.1 Distribution of semi-major axis and eccentricity for asteroids and comets in the inner Solar System. Credit: Alan Chamberlin, NASA/JPL-Caltech.

2.1.1 Asteroids

Asteroids are found throughout the Solar System and are often discriminated according to their heliocentric orbits. The technical classification of an asteroid is that it is a minor planet and does not exhibit a visible coma or outgassing. This classification can be somewhat ambiguous, with recent controversies over the proper

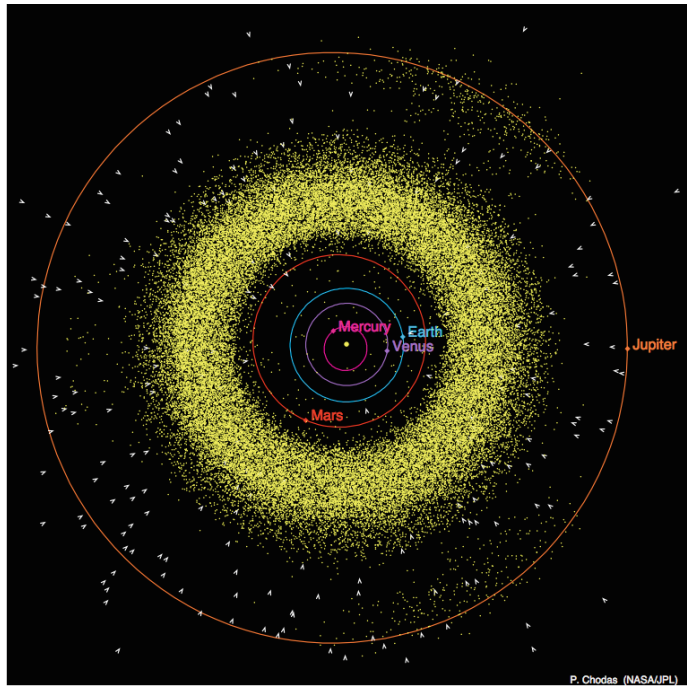


Fig. 2.2 A snapshot of asteroid and comet locations in the inner Solar System. Credit: Alan Chamberlin, NASA/JPL-Caltech.

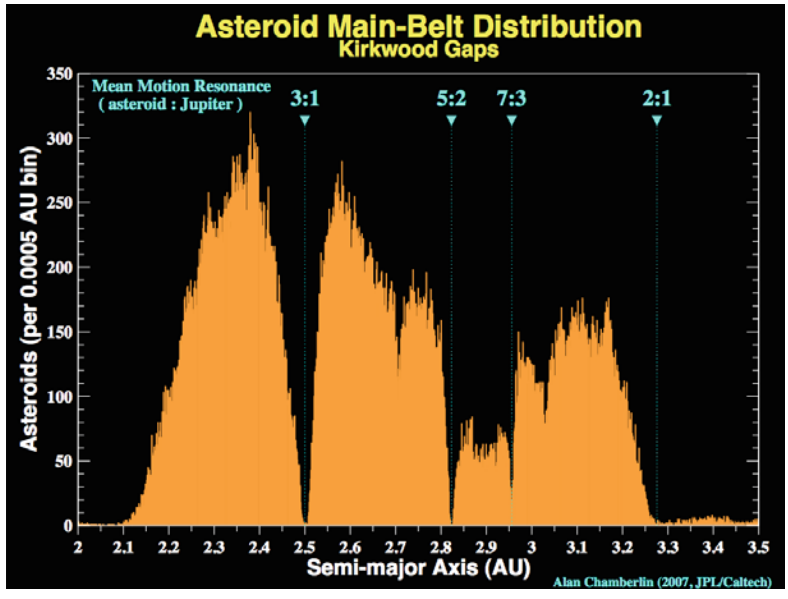


Fig. 2.3 Semi-major axis vs. asteroid population in the Main Belt, clearly showing the Kirkwood gaps. Credit: Alan Chamberlin, NASA/JPL-Caltech.

definition of a planet or dwarf planet, the discovery of some asteroids that seem to undergo periods of outgassing, and a growing realization that there may be a continuum of body types instead of distinct classifications of asteroids and comets. These issues are not of major concern to us here, however.

Near-Earth Objects

The Near-Earth asteroids (NEA) are often classified more generally as Near-Earth Objects (NEO) due to the presence of comet-like bodies that fit into the following orbital definitions. These bodies are clearly distinguished by their heliocentric orbits, as summarized in Table 2.1. NEOs are the most inexpensive bodies to rendezvous with as a class, due to their proximity to the Earth. Certain sub-sets of these bodies have been flagged as potential targets for future human exploration. NEOs are also the bodies with the highest probability of impact with the Earth, and are considered to be objects of interest due to this fact alone.

The lifetime of an NEO has been estimated to be on the order of 10 million years [49], and is limited by their chaotic interactions with the inner planets. These bodies can impact with the terrestrial planets, be ejected from the Solar System, or, most commonly, impact with the Sun. Due to their limited lifetime, there must be a source for these bodies. This source is now known to be the Main Belt, and is described in the following subsection.

Table 2.1 Definitions of Near-Earth Objects from the Near Earth Object Program Office at JPL's website [178]. q represents perihelion radius, Q represents aphelion radius, P represents orbit period.

Group	Description	Definition
NECs	Near-Earth Comets	$q < 1.3$ AU $P < 200$ years
NEAs	Near-Earth Asteroids	$q < 1.3$ AU
Atiras	NEAs whose orbits are contained entirely with the orbit of the Earth (named after asteroid 163693 Atira).	$a < 1.0$ AU $Q < 0.983$ AU
Atens	Earth-crossing NEAs with semi-major axes smaller than Earth's (named after asteroid 2062 Aten).	$a < 1.0$ AU $Q > 0.983$ AU
Apollos	Earth-crossing NEAs with semi-major axes larger than Earth's (named after asteroid 1862 Apollo).	$a > 1.0$ AU $q < 1.017$ AU
Amors	Earth-approaching NEAs with orbits exterior to Earth's but interior to Mars' (named after asteroid 1221 Amor).	$a > 1.0$ AU $1.017 < q < 1.3$ AU
PHAs	Potentially Hazardous Asteriods: NEAs whose Minimum Orbit Intersection Distance (MOID) with the Earth is 0.05 AU or less and whose absolute magnitude (H) is 22.0 or brighter.	MOID ≤ 0.05 AU H ≤ 22.0

Main Belt Asteroids

Main Belt asteroids are minor planets primarily found in orbit between Jupiter and Mars. Those observed range in diameter from 948 km (1 Ceres) down to dust detected by interplanetary spacecraft. They are characterized by a strong structure in their semi-major axis distribution, with clear gaps in the population at low-order resonances with Jupiter. The Main Belt is also dominated by families, which are collections of asteroids formed by the impact and break-up of larger bodies. In this sense, the main belt is continually eroding with larger bodies being broken up into cascades of smaller bodies, which themselves impact and continue this fractionation. Due to this the sizes of bodies in the Main Belt are in a collisional distribution, with the number of smaller bodies increasing with a power law with a size dependence on the order of $1/d^{2.3}$. This size distribution is relatively stable, with the largest few asteroids that have survived this process now isolated from other asteroids that could disrupt them, but with impacts continuing among the smaller members. The frequency of family formation is not that rapid, although dynamical astronomers have been able to identify several families that were formed relatively recently in the Main Belt (less than 10 million years ago in one case). Figure 2.4 shows the main groupings within the Main Belt, which to some extent are arranged along major family membership.

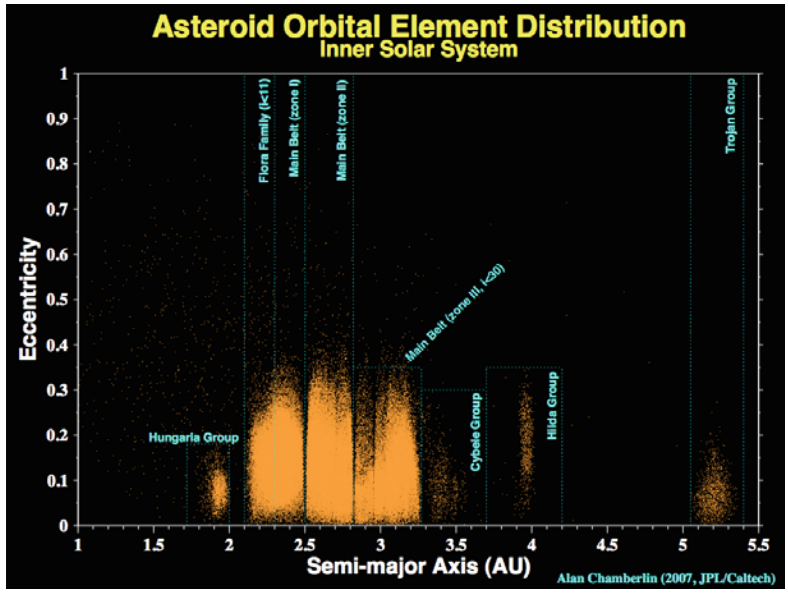


Fig. 2.4 Main asteroid groupings in the inner Solar System. Credit: Alan Chamberlin, NASA/JPL-Caltech.

Collisions between Main Belt asteroids is the main process that modifies and changes these bodies over time, although when one goes to smaller sizes, below 10 km in general, other processes become important and actually serve to migrate

these bodies through the solar system, serving as a source for the NEA. As has been established relatively recently, perturbations from Jupiter can cause asteroids to undergo unstable migrations within the Main Belt and be ejected from the Main Belt. Some of these bodies will scatter into the outer Solar System where interactions with the larger planets can cause these bodies to escape. Others will scatter down into the inner Solar System where they become NEA. A non-gravitational force known as the Yarkovsky effect [144, 25] has been identified as the dominant perturbation on small asteroids in the Main Belt. This effect causes a secular drift in their semi-major axes, allowing them to drift into destabilizing interactions with Jupiter. The Yarkovsky effect arises due to a combination of solar heating of asteroids, thermal inertia of these bodies (which delays the re-emission of absorbed solar photons), and their rotation which ensures that the re-radiated photons have a component that acts normal to the Sun-line. This net component causes a slow drift either towards or away from the Sun depending on the asteroid's obliquity. As it is a non-gravitational effect it preferentially acts on smaller bodies, which is consistent with the small sizes of NEA in general.

Trojans, Centaurs and Beyond

Traditional terminology defines asteroids to include Trojans (bodies captured in Jupiter's 4th and 5th Lagrange points), Centaurs (bodies in orbit between Jupiter and Neptune), and Trans-Neptunian objects (orbiting beyond Neptune). Many of these bodies are expected to be volatile-rich, however, and would likely outgas if they were in closer proximity to the sun. Among these bodies are objects significantly larger than the largest Main Belt asteroid, including the Pluto–Charon system, and thus are more correctly called "minor planets." The largest Trans-Neptunian objects are over 2,000 km in diameter and have been found to have multiple companions. These minor planets tend to be less rocky and more icy in composition since they formed farther from the Sun. In fact, a group of these objects known as the Kuiper Belt is thought to be the reservoir for short-period comets. The distinction between asteroid and comet has become increasingly difficult to characterize as we learn more about these primordial building blocks of our Solar System. Due to the lack of active outgassing of these bodies, however, a spacecraft in their vicinity would be subject to similar forces as from an asteroid. The major difference arises due to their distance from the Sun and their lower densities.

2.1.2 Comets

Comets are small bodies from the outer Solar System that migrate into orbits which bring them close enough to the Sun so that the volatiles on their surfaces and in their interiors begin to sublimate, causing them to outgas. This causes comets to take on a distinctly different appearance as seen from the Earth, forming large gaseous clouds in their vicinity which are affected by solar photon pressure and form their characteristic tails. The presence of volatiles on comets indicates that they are a pristine reservoir of chemicals and compounds from the formational

epoch of the Solar System. Due to this comets are currently a high scientific priority for both rendezvous missions and eventual sample return missions. The sizes of comets mirror that of asteroids and range from quite small – on the order of kilometers or less – to relatively large bodies up to 100 kilometers. Although comets primarily come from the outer Solar System, the designation of comet is only used for those bodies whose orbits are subsequently perturbed to the extent that their perihelion come close enough to the Sun for visible outgassing. Once such close interactions start, comets have substantial migration of their orbits over their lifetimes. The main force behind their migration is Jupiter, through its gravitational perturbations, and the Sun, through heating and generation of outgassing across their surface. The combination of these gravitational and non-gravitational forces makes it difficult to precisely predict comet orbits and can cause substantial migration over their lifetime. The time scales of these migrations, however, are measured in multiple years whereas the time scales of interest to us are on the order of days and years at most. Thus, it is reasonable for us to neglect changes in the orbits of the bodies studied here.

The primary differences between comets and asteroids are a much lower density for comets and the presence of a gaseous envelope that surrounds these bodies. This envelope is characteristically diffuse when far from the body and is only of dynamical interest when a spacecraft comes into close proximity to the nucleus. A simplistic model for a spacecraft interaction with a cometary jet is presented later in this chapter. The presence of these outgassing jets also perturb comet rotation states, with many of these bodies expected to be in an excited rotation mode. Again, as these are still relatively weak torques the rotational dynamics of these bodies can be modeled as following torque-free excited rotation states, discussed later. Finally, due to their source regions being in the outer Solar System, comet orbits tend to be highly eccentric, as can be seen in Fig. 2.1. This means that solar perturbations acting on orbiters in their vicinity can vary strongly over one orbit period, a consideration explicitly taken into account in our analysis.

For purposes of mission design and spacecraft dynamics, comets can be classified into two main types, short- and long-period comets.

Short-Period Comets

Short-period comets, or more generally periodic comets, are comets that have orbits which repeat in timespans on the order of 200 years or less, or which have been viewed multiple times. The comet Halley is the most famous of this set, and was actually the first comet to be imaged at close range. This class of comets are the usual targets for rendezvous or flyby missions, as their motion is relatively predictable – a crucial aspect for planning space missions.

What is particularly striking about these bodies is that they are transient and change their nature over time. Many comets are expected to eventually go dormant or become extinct, meaning that they no longer show outgassing effects. This can be caused by the depletion of volatiles on or near their surface – the fraction of asteroids that fall into this class is currently unknown. At the other end of the spectrum, they are also sometimes seen to undergo “bursting” following a perihelion

passage, indicating that the outgassing activities on these bodies are substantial enough to erode them and cause them to break apart. Another class of comets have been found that have perihelion extremely close to the Sun, called Kreutz or Sun-grazing comets. This class of comets have been discovered by spacecraft that image the Sun. A common end-state of a body undergoing frequent planetary flybys in the Solar System is that they impact with the Sun. Thus, it is not surprising to find comets with orbits in this state. Due to these different processes all comets currently visible are expected to have finite lifetimes in terms of their visible outgassing activity. This implies that they are replenished from source regions in the outer Solar System.

Within the short-period comets are several sub-classifications. Those specifically mentioned so far include extinct comets, dormant comets, and Kreutz comets. Another classification used is Main Belt comets, which are defined as comets with perihelion within the main asteroid belt. Similary, Near-Earth comets are comets that satisfy the Near-Earth definitions. Along with these diversity of names are included a diversity of properties, activity levels and even surface morphology.

Long-Period Comets

The other main class of comets are the long-period comets. These are comets that have only been viewed at one apparition. They are on orbits with extremely long periods (technically greater than 200 years but with effective periods up to millions of years), implying that they are on weakly bound, near-parabolic orbits about the Sun and emanate from the Oort Cloud. They can also be classified as inter-stellar comets, meaning that they are not gravitationally bound to the Sun and only have a single apparition before they leave the solar vicinity. These comets are typically found to be quite active and can be very large. Once observed, however, there is generally only a year or two at most until they have their perihelion passage and subsequently go far from the Sun again. Due to this, it is very difficult to plan for a rendezvous mission to such a body. At best, one could implement a flyby mission to such a comet, although the engineering development and approach to such a mission would have to be highly non-standard due to the tight time constraints.

2.1.3 Planetary Satellites

The final class of body considered in this text are planetary satellites. Planetary satellites are found around the majority of planets, only Mercury and Venus have none. Their orbits are generally near-circular, with some notable exceptions, and their orbit planes generally have small obliquities with respect to the heliocentric orbits of their primaries. These bodies are of great scientific interest for a variety of reasons. At one end of the spectrum, the larger planetary satellites – especially those about the gas giants – are thought to have formed via mechanisms similar to the formation of the Solar System's planets. Thus, their exploration would open new understandings of how these formation processes work. At the other end of the spectrum, some smaller planetary satellites, such as the irregular satellites about

Jupiter and Saturn, are thought to be captured bodies from early in the Solar System's formational process. As such, they may also retain early records of material properties and compositions from the formational epoch of the solar system. Other planetary satellites present challenges for understanding, such as the history of the Martian moons Phobos and Deimos, and the formation of the Moon, which now seems likely to have been formed in a large collision between a Mars-sized protoplanet and the early Earth. Perhaps most intriguing among the planetary satellites are the Jovian moon Europa and the Saturnian moons Titan and Enceladus. Europa and Enceladus are suspected of having liquid oceans trapped beneath their ice-covered shells, and have been identified as potential locations in the solar system where life could have emerged. Titan is now known to have a complex weather system that include methane precipitation and hydrocarbon lakes across its surface.

As a population, the size of planetary satellites range from a few kilometers across to objects larger than the planet Mercury. With only a few exceptions (notably the Martian moons of Phobos and Deimos and some of the smaller satellites of the gas giants) planetary satellites have shapes that are spheroidal, and thus satellite dynamics about these bodies are not affected by their gravity field distributions to the same extent as asteroid orbiters. What makes these strongly perturbed systems are their close proximity to their planets, resulting in significant perturbations from the combined gravitational attraction and centripetal accelerations that they experience due to their orbital motion. Due to these perturbations, whole classes of scientifically attractive orbits about these bodies such as polar orbits can be strongly unstable and present challenges for any space science mission. For a particular example, low-altitude circular polar orbits about Europa can impact with the satellite's surface within a few weeks if not properly designed or controlled. For these reasons a discussion of planetary satellites is included within this book, and the last three chapters are devoted to describing methods for characterizing orbital motion about these bodies.

2.2 Mass and Density of Small Bodies

A fundamental property of a small body is its total mass, as this ultimately controls the gravitational attraction it exerts on a spacecraft. Despite this, once the details of motion around these bodies are explored the total mass of a body is not as crucial of a parameter as its density. The mean density of a body can be computed if the volume and mass is known from $\rho = M/V$, where M is the body mass, ρ is its mean density, and V is its total volume. Another important concept for any body is its mean radius, which is defined as the radius of the sphere of equal volume, or $R = (3V/4\pi)^{1/3}$. By definition, the mean radius is essentially equivalent to the geometric mean of the body's size with the mean radius lying between the maximum and minimum radii of the body. As such, the mean radius is a reasonable measure of how close a spacecraft can come to a given body.

The importance of the density is exemplified if we consider the simple orbit period equation defined by Kepler's 3rd Law

$$T = \frac{2\pi a^{3/2}}{\sqrt{\mathcal{G}M}} \tag{2.1}$$

where T is the orbit period, $\mathcal{G}$ is the gravitational constant, M the body mass, and a the semi-major axis (or orbit radius for a circular orbit). Replacing the total mass using the mean density and mean radius the expression reduces to

$$T = \sqrt{\frac{3\pi}{\mathcal{G}\rho}} \left(\frac{a}{R}\right)^{3/2} \tag{2.2}$$

This shows the well-known result that an orbit period about a body, when specified in terms of body radii, is independent of the body size and only a function of its density. Thus, noting that the Earth's mean density is 5.5 grams per cubic centimeter, an orbit period at the surface of the Earth is 1.4 hours and at one radius altitude is 4 hours. Thus the orbit period about an asteroid of similar density will precisely mirror these periods. This is an important point as it means that the time scale of motion about small asteroids is similar to the same orbit time scales encountered at the Earth, albeit most asteroids have densities on the order of one-half that of the Earth with a commensurate increase of $\sqrt{2}$ in the orbit periods.

It is also instructive to consider the mass scales encountered at asteroids, as these can be directly compared with the spacecraft that visit them. A boulder with mean radius of 1 meter and mean density of 3.5 g/cm^3 will have a total mass of approximately 15 metric tons. For comparison, the mass of the NEAR-Shoemaker spacecraft when it was at the asteroid Eros was approximately 0.5 metric tons, or equivalent to a boulder of mean radius 30 centimeters (assuming a mean density of 3.5 g/cm^3). Thus, although it is true that small body masses are inconsequential as compared to planetary masses, spacecraft masses are similarly inconsequential as compared to asteroids even a few meters across.

Now consider the measured values of small body density. These are generally difficult measurements to make and require that the body either be a binary system – allowing the total mass to be estimated from Kepler's 3rd Law – or that perturbations to a spacecraft trajectory be measured when in close proximity to the body. As regards to planetary satellites or small binary asteroid secondaries, these can only be estimated if their mass is large enough to cause a measurable reflex motion in the primary. Even the large satellites of Jupiter and Saturn have only been precisely weighed by measuring the deflection of a spacecraft trajectory during close approaches, or by computing their mutual attractions. Comet densities have not been directly measured to date, although they have been estimated to be as low as 0.3 g/cm^3. A consistent theme for the estimated densities of asteroids and comets are that measurements have generally found these bodies to have a density significantly less than the density of material they are comprised of. Thus, although the expected grain density of asteroids Eros and Itokawa was on the order of 3.5 g/cm^3, the measured bulk densities of these bodies were 2.7 and 2.0 g/cm^3,

respectively. Similar results have been found for almost all precisely measured bulk densities, indicating that the internal structure of asteroids has high porosity in general. Based on these measured values, reviewed in [15], it is reasonable to assume bulk densities of between 0.3 and 1 g/cm^3 for comets and ranging from 1 to 5 g/cm^3 for asteroids. The densities of planetary satellites are not as important, as the perturbations in their vicinity are controlled by several additional factors beyond their mass density, including the mass of their planet and their orbit about the planet.

2.3 Spin States of Small Bodies

Spins are a crucial property of our target bodies as they define how the mass distributions of these bodies vary in inertial space, which controls how they will interact with an orbiting particle. Spin dynamics are easily divided into three classes: uniformly rotating, complex rotation, and synchronized rotation. Asteroids and comets inhabit the first two classes, and planetary satellites almost uniformly inhabit the last. Systems such as binary asteroids can exhibit several types of rotational motion, usually the uniform rotation of the primary and the synchronous rotation of the secondary. Spin states can be determined for asteroids and planetary satellites by observing time variations in the intensity of light scattered from their surface. However, the spin states of comets are not as well known or understood, arising from the difficulty of discriminating between internal light reflection within their comas and variations in reflected light from their surfaces. Still, some comets have had their spin states accurately determined, and we discuss them briefly where appropriate.

Each spin state has unique quantities that will define the body's rotational motion. An important specification shared by all of these spin types is the orientation of their total rotational angular momentum. For the analysis presented in this text a reasonable approximation will be that the bodies conserve their rotational angular momentum. Some exceptions to this occur for planetary satellites and binary asteroid systems, but these will be discussed as appropriate. The usual designation of a body's rotational angular momentum is relative to an inertially fixed frame. Such designations are important for providing a unique classification of a body's rotation; however, for our focus on motion close to these bodies, these inertial coordinates do not provide useful information by themselves. What is more dynamically relevant for our study is the orientation of the body's rotational angular momentum relative to the body's heliocentric (or planetocentric) orbit.

Of prime interest is the orientation of the rotational angular momentum vector, described as a unit vector with two angles relative to the body's orbit. First, define the obliquity of the body as the angle between the heliocentric or planetocentric orbit angular momentum and the rotation angular momentum vectors, denoted as β. Second, denote the right ascension of the body as the orientation of the rotation pole projected into the orbit plane, plus 90°, or geometrically as the planar orientation of the cross product between the orbit normal and the body's angular

momentum vector, denoted as α. As a reference direction in the orbit plane choose the periapsis, denoted as $\hat{\boldsymbol{E}}$, and the orbit normal, denoted as $\hat{\boldsymbol{H}}$. The cross product of these defines the triad, $\hat{\boldsymbol{E}}_\perp = \hat{\boldsymbol{H}} \times \hat{\boldsymbol{E}}$. With these definitions, the direction of the body's rotational angular momentum is

$$\hat{\boldsymbol{p}} = \sin\beta \sin\alpha \hat{\boldsymbol{E}} - \sin\beta \cos\alpha \hat{\boldsymbol{E}}_\perp + \cos\beta \hat{\boldsymbol{H}} \tag{2.3}$$

2.3.1 Uniform Rotators

The vast majority of asteroids are found to be in a uniform rotation state about their maximum moment of inertia. This is the expected spin state as, under energy dissipation caused by tides raised by complex rotation, it is the minimum energy rotation state of a body [20]. Figure 2.5 shows the size/spin distribution of all asteroids with estimated spin rates and states. Spin rates among uniform rotators range from extremely fast to extremely slow. There is a clear maximum spin rate for larger bodies, equivalent to a spin period on the order of 2.4 hours, evident in Fig. 2.5. This observation has led to the overall realization that larger asteroids are most probably rubble pile bodies, as this limiting spin period equals the spin period when material on the surface of a body with mean density of $\sim$2.1 g/cm^3 enter orbit. This maximum spin rate clearly increases for smaller-sized bodies. Current hypotheses that attempt to explain this are either that smaller bodies are monolithic rocks or that these smaller bodies may still be rubble piles but are bound together with cohesive bonds [133, 69, 169].

Uniform rotators can be specified with a given rotation period or rotation rate and their rotation pole, which is coincident with their rotational angular momentum. Given this, the attitude of a uniformly rotating body can be found by a simple integration of the constant rotation rate. The spin rates and spin poles of asteroids and comets are known to change over time. For asteroids the culprit is the YORP effect [145] while for comets it is torques due to outgassing [150]. The study and analysis of these longer-term dynamics are of interest, but are not relevant over the shorter timespans of interest for spacecraft missions design.

2.3.2 Complex Rotators

There is also a sizable population of asteroids and comets in non-uniform rotation, often referred to as "tumbling" or "complex rotation". Despite the frequent perception that such bodies are chaotically spinning, they actually follow a well-defined quasi-periodic rotational motion which is described in the next chapter. Larger bodies that are in a complex rotation state generally have a low overall spin rate. This has been clearly related to their relaxation time [63], with estimates for these bodies to relax to uniform rotation about their maximum moment of inertia shown in Fig. 2.5. At the small size scale there is a more puzzling set of tumbling bodies: fast spinners which are in complex spin states. A clear theoretical explanation for the existence or persistence of these bodies has not been fully worked out as of yet, but would present a unique challenge to a rendezvous mission to one of these

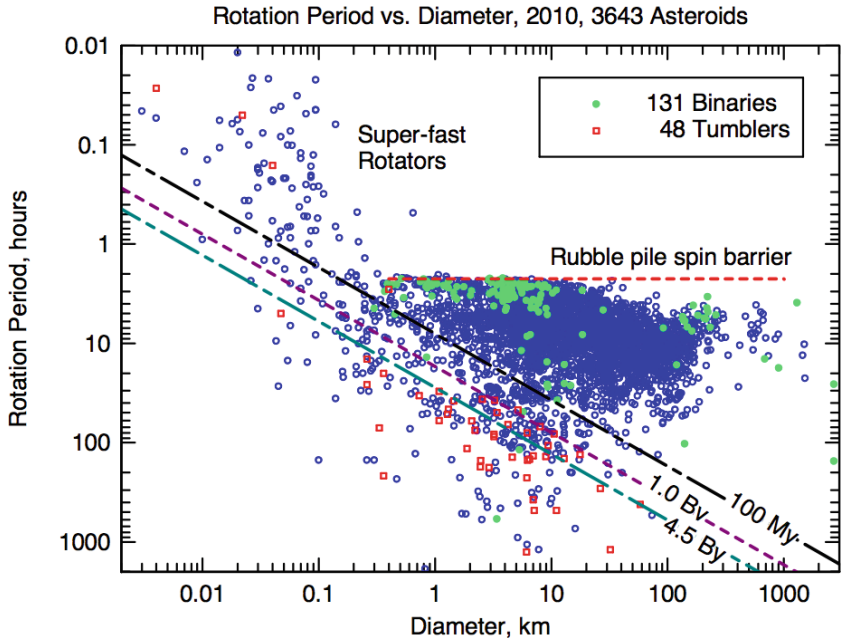


Fig. 2.5 Size–spin period distribution for asteroids. The smaller bodies are almost all in the NEO population with the larger bodies almost all in the Main Belt population. The dashed lines are the expected relaxation times for asteroids in non-principal-axis rotation. Credit: A.W. Harris, personal communication.

bodies. While the number of verified asteroids in non-principal-axis rotation are somewhat modest, the actual population of bodies in complex rotation states may be much larger due to the difficulty in detecting and confirming a non-principal-axis spin state based on light curve observations.

The spin state of a complex rotator is more difficult to specify. There are two fundamental frequencies that appear in their rotational motion, related to their nutation and to the spin rate of the body. Specifying these two quantities is not enough to provide a detailed description of the spin state in time, however. For that the moments of inertia of the body need to be specified as well (or at least the ratios of the moments of inertia). The detailed formula relating spin rates, attitudes and moments of inertia are provided in the next chapter.

2.3.3 Synchronous Rotators

The third distinct class of rotation states are those that are synchronous relative to their mutual orbit with their parent body. For asteroids this class is often found among binary asteroids, where the smaller member is frequently found in a synchronous rotation. In a fraction of binary asteroids synchronicity is also found between both bodies. Planetary satellites are almost all uniformly found in synchronous spin states. The most notable exception to this is the Saturnian moon

Hyperion, which is known to be in chaotic rotation. In a few larger systems, such as Pluto–Charon, the entire system is also doubly synchronous.

Such synchronicity is an expected dynamical state for bodies in orbit about each other. When they are non-synchronous, tides are raised on either body and act to de-spin the smaller body (or both bodies if they have relatively equal size). Eventually this leads them to be trapped into a 1:1 spin–orbit resonance, possibly with some libration. The continued dissipation of energy due to relative motion should eventually damp out this libration, or reduce it to a small value. The one binary asteroid imaged at high resolution, 1999 KW4, showed evidence that the secondary was librating, and thus these bodies are not necessarily completely relaxed to a synchronous state. On the other hand, capture into a synchronous rotation state is presumed to be the fastest tidal dissipation process in binary systems [50], and thus is the expected rotation state of asteroid binary secondaries.

The spin state of a synchronous rotator is specified by the orbit period of the system. If the body is in libration, the maximum libration angle of the satellite relative to the line connecting the two bodies should also be given. Alternatively, the orbit of the system may also be elliptic, which then creates a libration between the satellite and the body centers, even if the satellite spins at a constant rate. The detailed dynamics of interaction of a librating synchronous satellite can be relatively complex, involving interactions between the orbit eccentricity and the forced and free libration of the satellite.

2.4 Size, Shape and Morphology

Asteroids and comets are frequently defined by their shapes, whereas larger planetary satellites generally have shapes that only deviate slightly from spheroids. Thus, our focus is mainly on asteroids and comets in this section. As will be seen later, the detailed surface structure of these bodies is complex with unresolved variations down to the highest resolution at which they have been imaged. This level of fine detail is only needed if the surface motion of a vehicle is modeled. For gravitational and orbital studies, it is sufficient to have lower resolution models that capture the global morphology of a body's shape without providing the fine details of its surface. Such models are available from NASA's Planetary Data Systems-Small Bodies Node (PDS-SBN) [117].

2.4.1 Shape Model Formats

There are three fundamental formats that can be found at the PDS-SBN for describing specific shape models. Historically, use was made of a simple tabulation of radius, latitude and longitude, usually with equally space longitudes leading to oversampling of the polar regions. Such model formats are still available for some shapes, but this description is not amenable to our work and is not discussed any further. The most versatile shape models are specified as general polyhedra with triangular facets at the surface defined by vertices specified as vectors from a cen-

tral origin. Each facet of the surface can be uniquely modeled as a flat plate for a triangular facet. These shapes can be specified with two lists, one an ordered list of all vertices of the body expressed as vectors in a body-fixed frame, the other being vertex number triples that define the vertices involved with each surface facet. The defining vertices for a facet are generally given in counter-clockwise direction so that the resulting normal defined by the right-hand rule points away from the surface. From Euler's formula there is a simple relation between the number of faces for a triangular facet polygon, f, and the number of vertices, v, where $f = 2(v-2)$. Thus, to specify a shape model with f triangular facets fully requires $3v = 3(f/2+2)$ real numbers and $3f$ integers, or $9f/2+6$ total numbers. Due to this, storage requirements for high-resolution shape models can become onerous. Figure 2.6 shows an example model with surface faceting clearly shown.

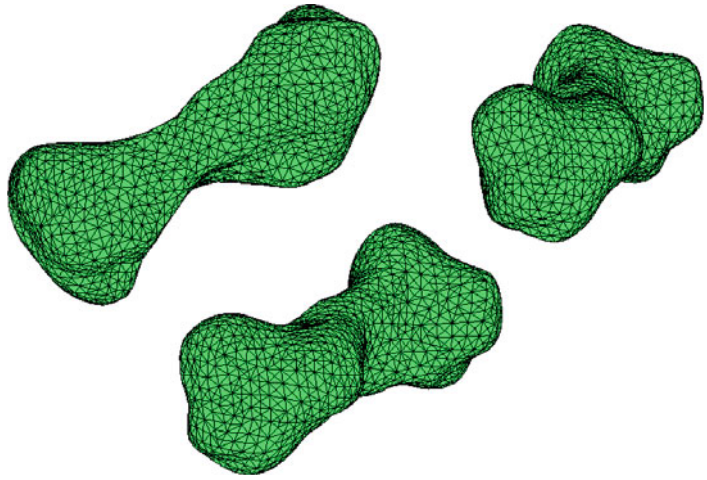


Fig. 2.6 Three views of asteroid Kleopatra in the polygon shape model format, with triangular surface facets.

The third format is the quadrilateral-implied format developed by R.W. Gaskell and also stored in the PDS-SBN for some models. This format describes each surface element of a shape by four vertices. Due to this, the surface cannot be modeled as a flat plate and, indeed, there is some ambiguity in the surface orientation. As the finest surface element is always at or beyond the limit of modeling resolution, it is appropriate to choose some reasonable rule for the surface orientations. Common choices are to define a diagonal between two vertices, splitting the quadrilateral into two triangular segments, or implementing an interpolation routine for the orientation of the surface normal across the body. The model format can be best envisioned as a parallelepiped drawn to enclose the body with the surface of each discretized into n squares on a side, creating $f = 6n^2$ facets. The shape can then be defined by conforming this structure onto the true surface, distorting the grid as necessary in 3-D space to make it cover the surface and account for non-convex surface features (see Fig. 2.7). Since the topology of the surface is fixed, to define the shape one only needs to know the vertices listed in the appropriate order, and

the total number of faces on an edge. Thus, to fully specify a Gaskell shape model one needs a total of $6n^2 + 2$ vertices. Since the shape format is defined implicitly, it is necessary to only have one file listing the vertices in the appropriate order with a header given the number of sides n. Thus, a shape model with a total of f surface facets requires storage of $18f + 12$ real numbers, a factor of four times more than the triangle plate format. However, there are still significant advantages to this form as it is not necessary to search through lists to find facets due to their implied storage. Due to this, it is possible to practically develop shape models of much higher resolution using this approach.

2.4.2 Shapes Based on High-Resolution Imagery

These are the most accurately determined shapes, as they are constructed from combining visual imagery taken at different viewing geometries and phase angles. The best shape models are those constructed from an orbiting spacecraft, as full coverage of the body at a uniform resolution is possible. Only asteroids Eros and Itokawa have shapes with uniformly high resolution as only these two asteroids have been visited by a rendezvous spacecraft to date. The two Mars satellites, Phobos and Deimos, have global shape models but have non-uniform surface resolution, due to flyby constraints for Mars orbiters. Flyby targets can sometimes have global

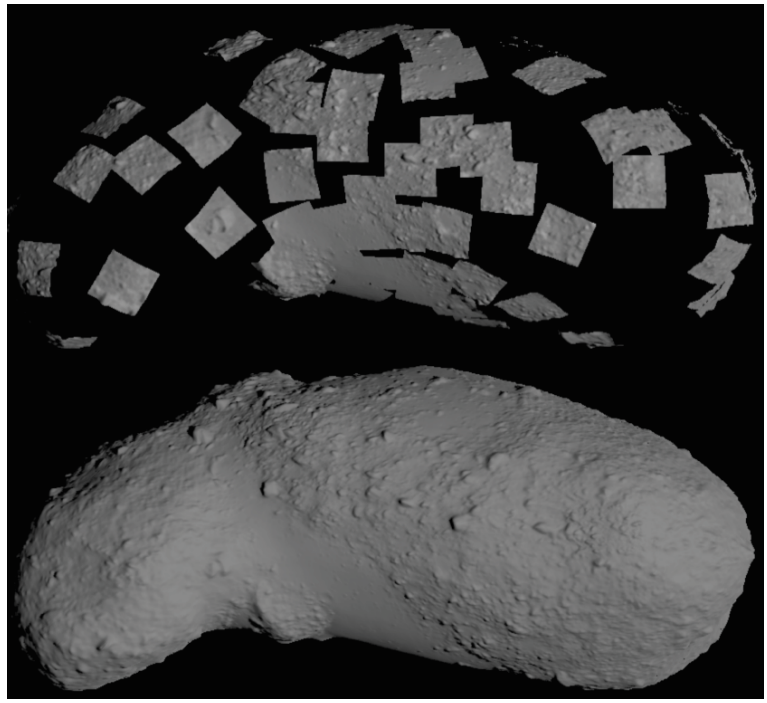


Fig. 2.7 Transitional and final shape model construction for the asteroid Itokawa based on Hayabusa imaging data [44].

shapes, but they are always non-uniform in resolution due to the rapid speeds at which flybys occur. Figure 2.7 graphically shows an intermediate step in the process of shape reconstruction based on asteroid imaging, in this case for asteroid Itokawa.

Shape models for specific bodies imaged by spacecraft are available at the PDS-SBN. These currently include full and partial shapes of asteroids and planetary satellites Phobos, Deimos, Gaspra, Ida, Eros, Mathilde, Itokawa, Lutetia, and Steins. Shape models of Vesta and Ceres will become available as the DAWN mission progresses. Partial comet shape models are also available, or will be soon, for comets Borrelly, Wild 2, Tempel 1 and Hartley 2.

2.4.3 Shapes Based on Radar Range-Doppler Imaging

Based on the pioneering work of Steve Ostro and Scott Hudson, small bodies that pass close enough to the Earth (generally much less than 0.1 AU) and with the proper trajectory can be imaged by the Arecibo or Goldstone radio antenna. During flyby range-Doppler imaging of asteroids can be obtained, and in a complex process the shape and spin state of that body can be estimated [120]. Shape models with surface resolution of less than ten meters have been obtained by this approach, although more typical resolutions are on the order of several tens of meters or larger. The best results also rely on the presence of light curve data for precise determination of the spin period. Shape models for these bodies have usually been produced in the triangular facet model format. Again, shape models for these asteroids are available at the PDS-SBN and currently include Castalia, Toutatis, Kleopatra, Betulia, 1999 KW4 Alpha and Beta, plus several additional ones. Figure 2.8 shows

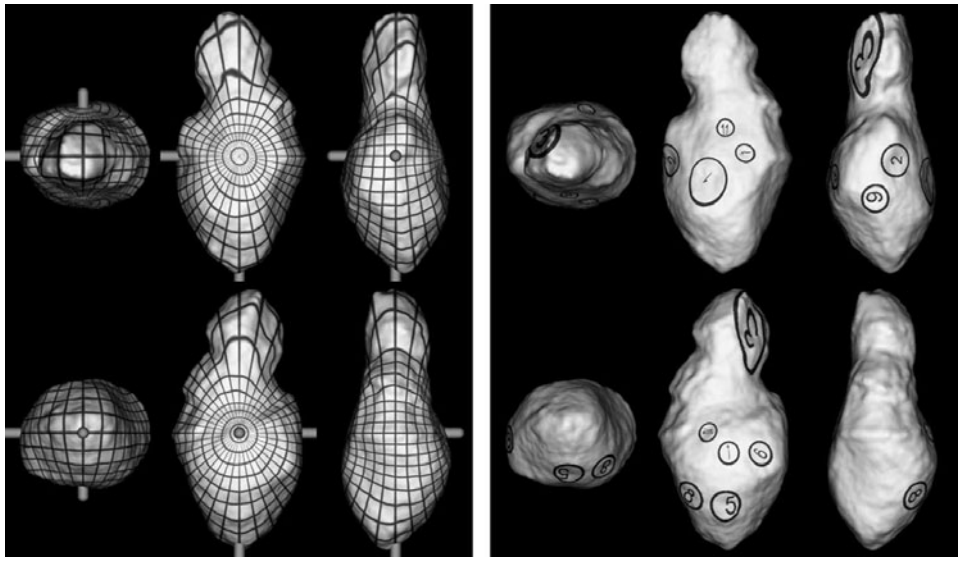


Fig. 2.8 High-resolution shape model of asteroid Toutatis based on its 1996 apparition [77].

the highest global resolution shape model estimated for an asteroid to date using this methodology, with surface resolution at the order of six meters.

2.4.4 Shapes Based on Light Curve Analysis

The simplest manner in which to estimate an asteroid or comet shape is by observing their light curves. These are photometric observations of a body sampled over extended observing periods. This allows the observer to measure the variation in the reflected light. Given an assumed albedo (i.e., reflectance) these measurements also allow the total projected area of the body to be tracked as a function of time. The simplest approach is to map the total variation in reflected light to an upper bound on the ratio between maximum and minimum ellipsoidal axes of the body. Newly developed techniques also enable a convex hull shape to be estimated [78, 79]. Although the resulting information does not provide great detail, it is sufficient to construct an estimate of the asteroid's second degree and order gravitational field, which plays the most significant role in controlling the dynamical evolution of a particle in orbit about the body.

2.4.5 Surface Characteristics

Precision observations of asteroid and comet surfaces are limited. Rendezvous missions have provided local observations of surfaces down to the centimeter scale over specific regions of the bodies, and globally to decimeter scale at Itokawa and meter scale at Eros. These are the highest-quality observations, as the surfaces are generally observed from multiple phase angles, which allow for stereo photoclinometry techniques [46] to determine albedo and surface topography simultaneously. Flyby observations have much lower resolution and generally only one view of the surface. Radar observations provide global surface features at the decameter scale. Resulting from these observations are a few important facts. First, most asteroids and comets have relatively relaxed surfaces, meaning that the vast majority of their surfaces are below the typical angle of repose for granular material, approximately 40 degrees, with many of these bodies having most of their surfaces significantly below slopes of 30 degrees. This relative smoothness does not extend down to small scales, however, as surfaces have strong segregation of surface morphology, with boulders dominating in some regions and others covered with apparent smooth pebbles or dust (in the case of Eros). The boulder terrains are rugged with steep small-scale slopes, while the covered regions generally have low slopes. On Itokawa there is evidence that the smooth regions migrated to their current locations, indicating that the surface of that body may be geologically active [108]. The distribution of material across asteroid surfaces seems to be controlled by the body's geopotential including gravity and rotation effects. Spin rates on some bodies can be fast enough to reverse the usual geopotential lows (which lie at the short axes of a non-rotating ellipsoid) and cause the long axes or equators of the bodies to be the potential low [57].

Comet surfaces have been viewed from a number of flyby missions, and will be investigated in situ once the Rosetta spacecraft has its rendezvous with comet Churyumov-Gerasimenko. The hallmark of comet surfaces are that each seems to be different from the last. Some surfaces, such as for comet Wild 2, have extremely rugged terrain while others, such as comet Tempel 1, have large regions of their surface apparently covered with smooth material. Our understanding of these bodies is still advancing and is likely to change in the future.

2.4.6 Mass Distribution Morphology

A significant and frequent morphology found among asteroids is that they form binary, or in some cases tertiary or higher, systems. Based on measurement statistics it is believed that at least 15% of NEOs are binary objects, with similar percentages being found in the small Main Belt population [135]. Binary asteroids are also very common among larger asteroids and are ubiquitous in the Kuiper Belt. The formation circumstances of binaries are believed to be quite different across the entire range of these bodies and is an active area of research. The modeling of these bodies can also be rather complex depending on the body's detailed circumstances [38], although this text adopts a relatively simple approach to their modeling that captures the main aspects of these bodies of relevance to spacecraft visits.

A different aspect of mass distribution morphology is in regard to the internal density distribution of an asteroid. The one asteroid whose shape and gravity field have been precisely measured to date, Eros, was found to have a large degree of homogeneity in its internal density distribution, with no detectable mass concentrations [107, 83]. This was a somewhat surprising result, and may not be true of all small bodies. In particular, the binary asteroid 1999 KW4 had a statistically significant difference in the density of the smaller and larger components, with the smaller component approximately 40% more dense. A key question for future missions to small asteroids will be to precisely determine the internal density distribution. Unfortunately the gravity field measurements of Itokawa were not of sufficient accuracy to provide strong constraints for that body.

2.5 Gravitational Potentials

The key distinguishing features of small bodies are their irregular shapes, as noted above, and hence their strongly non-spherical mass distributions. This feature drives many of the interesting dynamics when close to these bodies.

The following derivation and discussion assumes that the attraction is between a massless particle and an arbitrary mass distribution. To set up the potential for this situation, we integrate the potential of a differential mass element over the entire body

$$U(\boldsymbol{r}) = \mathcal{G} \int_{\mathcal{B}} \frac{dm(\boldsymbol{\rho})}{|\boldsymbol{r} - \boldsymbol{\rho}|} \tag{2.4}$$

where $\boldsymbol{\rho}$ is the position vector of the differential mass element dm, $\mathcal{B}$ is the collection of all mass elements, and $\mathcal{G}$ is the gravitational constant, taken to equal 6.673×10^{-8} cm^3 g^{-1} s^{-2}. Figure 2.9 shows the geometry of the mass distribution. It can be shown that the gravitational potential satisfies Laplace's equation outside of the body, $\nabla^2 U = 0$, and Poisson's equation inside the body, $\nabla^2 U = -4\pi\mathcal{G}\sigma$ where σ represents the local density. For a sphere with constant density, total mass M, and radius R, integration of this potential yields:

$$U = \begin{cases} \dfrac{\mathcal{G}M}{|\boldsymbol{r}|} & r > R \\ \dfrac{\mathcal{G}M}{2R^3}\left(3R^2 - r^2\right) & r \leq R \end{cases} \tag{2.5}$$

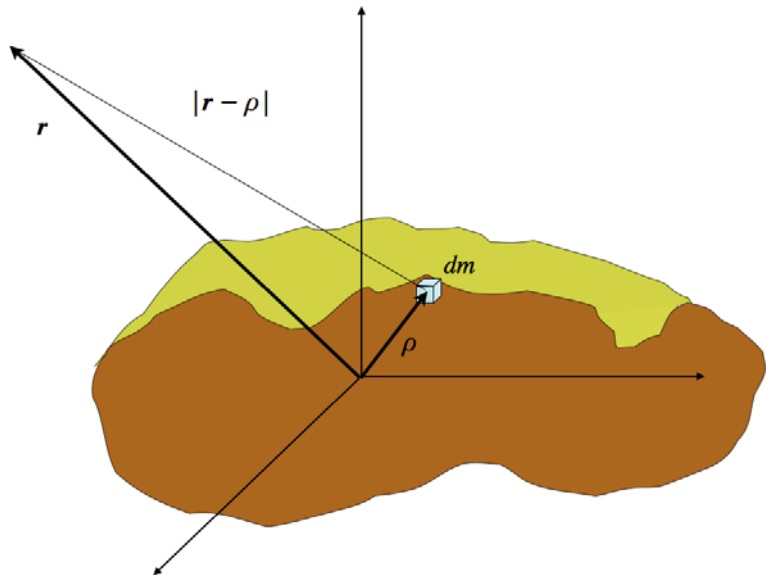


Fig. 2.9 Mass distribution geometry.

2.5.1 Spherical Harmonics Models

There are several approaches to specify the gravitational potential for a non-spherical body. Mathematically, any solution to Laplace's equation which corresponds to the physical mass distribution will work. This fact can be used to great benefit by constructing a set of orthogonal solutions to Laplace's equation and choosing the coefficients of these expansions to match with the actual potential function.

Specifically, Laplace's equation can be solved by separation of variables in terms of spherical coordinates [80]. Denote the spherical coordinates corresponding to a position vector $\boldsymbol{r} = x\hat{\boldsymbol{x}} + y\hat{\boldsymbol{y}} + z\hat{\boldsymbol{z}}$ as

$$r = \sqrt{x^2 + y^2 + z^2} \tag{2.6}$$

$$\sin\delta = \frac{z}{r} \tag{2.7}$$

$$\tan\lambda = \frac{y}{x} \tag{2.8}$$

where δ is the latitude and λ is the longitude. The general form for the spherical harmonic potential for a gravity field is then:

$$U(r,\delta,\lambda) = \frac{\mu}{r}\sum_{l=0}^{\infty}\sum_{m=0}^{l}\left(\frac{r_o}{r}\right)^l P_{lm}(\sin\delta)\left[C_{lm}\cos m\lambda + S_{lm}\sin m\lambda\right] \tag{2.9}$$

where $\mu = \mathcal{G}M$, the gravitational parameter of the body, r_o is the normalizing radius (often chosen as either the maximum radius or mean radius of the body), P_{lm} are the Associated Legendre Functions, and C_{lm} and S_{lm} are called the gravity field harmonic coefficients (or Stokes coefficients). Specification of these coefficients is analogous to defining the mass distribution and the potential of the body. The Associated Legendre Functions can be defined by the closed-form relationship [80]

$$P_{lm}(\sin\delta) = \cos^m\delta \sum_{i=0}^{\text{int}[(l-m/2]} T_{lmi}\sin^{l-m-2i}\delta \tag{2.10}$$

$$T_{lmi} = \frac{(-1)^i(2l-2i)!}{2^l i!(l-i)!(l-m-2i)!} \tag{2.11}$$

where the int$[x]$ function returns the integer part of x (see Appendix E for a partial list of these functions).

Functional orthogonality is an important property of the spherical harmonic expansion. Define the quantity S_{lmi} for $i = 0, 1$ as

$$S_{lm0} = C_{lm}P_{lm}(\sin\delta)\cos(m\lambda) \tag{2.12}$$

$$S_{lm1} = S_{lm}P_{lm}(\sin\delta)\sin(m\lambda) \tag{2.13}$$

Then the orthogonality property is most easily stated as

$$\int_{\delta=-\pi/2}^{\pi/2}\int_{\lambda=0}^{2\pi} S_{lmi}S_{hkj}\cos\delta\, d\delta\, d\lambda = \frac{4\pi(l+m)!\delta_i^j\delta_m^k\delta_h^l}{(l-m)!(2l+1)(2-\delta_m^0)} \tag{2.14}$$

where δ_i^j is the Kronecker delta function and equals 0 if $i \neq j$ and equals 1 if $i = j$. Thus, the integral is only non-zero if all three of the indices match with each other.

If a mass distribution is given, the gravity coefficients C_{lm} and S_{lm} can be directly computed using the orthogonality relations. Specifically, assume that the

density of the body is specified at each point as $\sigma(\boldsymbol{\rho})$. Then the gravity coefficients can be found from the integrals [110]:

$$(C,S)_{lm} = \frac{(2-\delta_m^0)}{M}\frac{(l-m)!}{(n+m)!}\int_{\mathcal{B}}\left(\frac{r}{r_o}\right)^l P_{lm}(\sin\delta)\,\mathrm{cs}(m\lambda)\sigma\,dV \tag{2.15}$$

$$\sigma\,dV = \sigma(r,\delta,\lambda)r^2\cos\delta\,dr\,d\delta\,d\lambda \tag{2.16}$$

where r_o is an arbitrary reference radius often taken to equal the maximum radius of the body or the mean radius of the body, and cs denotes the cosine or sine function for computation of the C_{lm} or S_{lm} gravity coefficient, respectively. The reference radius r_o ensures that the gravity coefficients are dimensionless numbers, otherwise they would have units of length to the lth power.

For a constant density body one integration can be performed over the radius and the formula can be reduced to:

$$(C,S)_{lm} = \frac{\sigma(2-\delta_m^0)}{M}\frac{(l-m)!}{(n+m)!} \int_S \frac{R^3(\delta,\lambda)}{l+3}\left(\frac{R(\delta,\lambda)}{r_o}\right)^l P_{lm}(\sin\delta)\,\mathrm{cs}(m\lambda)\cos\delta\,d\delta\,d\lambda \tag{2.17}$$

where the radius $R(\delta,\lambda)$ is now a function of the latitude and longitude and defines the shape of the body.

Gravity Coefficients and Mass Moments

There is a well-known relationship between the gravity coefficients of all degrees and orders and the high-order mass distribution moments of an arbitrary body [95]. Of most interest are these relationships up to degree and order 2. First note that $C_{00} = 1$, from the above relationships. Next, note that the first degree and order gravity coefficients are related to the center of mass of the body. Specifically, in the Cartesian coordinates defined by the spherical coordinate system the following relations hold

$$x_{CM} = C_{11}r_o \tag{2.18}$$

$$y_{CM} = S_{11}r_o \tag{2.19}$$

$$z_{CM} = C_{10}r_o \tag{2.20}$$

where x_{CM}, etc. denote the center of mass of the body in the chosen body-fixed coordinate system. It is always possible to redefine the origin of the body-fixed coordinate system to coincide with the center of mass of the body, and thus render the first degree and order terms identically zero, $C_{11} = S_{11} = C_{10} = 0$. For a specified gravity field this is usually done, meaning that the first term beyond the spherical attraction of the body arise at the second degree and order. It is important to note that the center of mass and center of figure of a body will coincide if the body has a constant density distribution. For space missions to small bodies this assumption

is usually made initially; however, subsequent tracking of a spacecraft close to the body can determine whether any first degree and order gravity coefficients exist. If they do, they are evidence of internal density heterogeneity [107, 83].

The second degree and order gravity coefficients are related to the moments of inertia of the body.

$$I_{xx} - I_{yy} = -4Mr_o^2 C_{22} \tag{2.21}$$
$$I_{yy} - I_{zz} = Mr_o^2 \left(C_{20} + 2C_{22}\right) \tag{2.22}$$
$$I_{zz} - I_{xx} = -Mr_o^2 \left(C_{20} - 2C_{22}\right) \tag{2.23}$$
$$I_{xy} = -2Mr_o^2 S_{22} \tag{2.24}$$
$$I_{yz} = -Mr_o^2 S_{21} \tag{2.25}$$
$$I_{zx} = -Mr_o^2 C_{21} \tag{2.26}$$

These results hold independent of internal density heterogeneity. There are a few important observations that can be drawn from these relations. First, it is not possible to uniquely specify the moments of inertia given the second degree and order gravity coefficients, at best one can only solve for the differences between them. This is an important limitation, as it is often possible to measure the gravity field of an object by tracking its motion, but it is not possible to determine its full moments of inertia from this information. In general, to determine the full moment of inertia the body must also have non-trivial rotational dynamics (i.e., non-principal-axis rotation). The second observation is that for any given mass distribution it is always possible to define a set of coordinates such that the products of inertia, i.e., I_{xy}, I_{yz}, and I_{zx}, are zero. Thus, it is always possible to define a coordinate frame where the second degree and order gravity coefficients $C_{21} = S_{21} = S_{22} = 0$, or that the second degree and order gravity field can be reduced to the two gravity coefficients C_{20} and C_{22}. Again, for a space mission where the gravity field is to be estimated, a usual initial assumption is to choose the coordinate frame fixed in the small body to be nominally aligned with the constant density principal axes. Then, estimation of any non-zero values for the gravity coefficients C_{21}, S_{21} or S_{22} is again evidence of density heterogeneity.

Thus the simplest form of the gravitational potential begins at the second degree and order with only the coefficients C_{20} and C_{22}. Note that in the literature the gravity field coefficient C_{20} is often equivalently specified as the "J_2" gravity coefficient, with the relationship $C_{20} = -J_2$. Then, the simplest, non-trivial gravity field for studying dynamics in the vicinity of a body is:

$$U = \frac{\mu}{r}\left[1 + \left(\frac{r_o}{r}\right)^2 \left\{C_{20}\left(1 - \frac{3}{2}\cos^2\delta\right) + 3C_{22}\cos^2\delta\cos(2\lambda)\right\}\right] \tag{2.27}$$

It has also been found in previous studies that the second degree and order gravity field accounts for the majority of the perturbations on the dynamical system, and thus this model can be taken as a simple stand-in for a more general system.

Limitations of the Spherical Harmonics Approach

The spherical harmonics approach to modeling an arbitrary gravitational field does not work in all situations. A fundamental assumption for the use of this modeling approach are that the series converge to the true gravity field. This assumption appears nowhere in the above statements, but can be identified when one derives the explicit expansion for the gravity field. Consider the general statement of the gravity field again, making a constant density assumption for convenience:

$$U = \frac{\mu}{V} \int_{\mathcal{B}} \frac{dV}{|\boldsymbol{r} - \boldsymbol{\rho}|} \tag{2.28}$$

To place this function into the form of the spherical harmonic expansion make the assumption that $|\boldsymbol{r}| > \max_{\mathcal{B}} |\boldsymbol{\rho}|$, meaning that the potential at the point in question lies completely outside the maximum radius of the mass distribution being considered. Under this assumption the integrand can be expanded into the classical Laplace series form

$$\frac{1}{|\boldsymbol{r} - \boldsymbol{\rho}|} = \frac{1}{r} \sum_{i=0}^{\infty} \left(\frac{\rho}{r}\right)^i P_{i0} \left(\frac{\boldsymbol{r} \cdot \boldsymbol{\rho}}{r\rho}\right) \tag{2.29}$$

where the P_{i0} are known as the Legendre polynomials (see Appendix E). As detailed in [95] there is a one-to-one correspondence between the integral

$$\frac{1}{V} \int_{\mathcal{B}} \left(\frac{\rho}{r}\right)^i P_{i0} \left(\frac{\boldsymbol{r} \cdot \boldsymbol{\rho}}{r\rho}\right) dV \tag{2.30}$$

and the ith degree and order spherical harmonic gravity field. If the point in question lies at a radius less than the maximum radius of the mass distribution the Laplace series is not defined when integrating at radii $|\boldsymbol{\rho}| = r$ and diverges when $|\boldsymbol{\rho}| > r$.

A direct remedy for this situation is to define two different expansions, one valid for the usual case when $|\boldsymbol{\rho}| < r$ and the following expansion when $|\boldsymbol{\rho}| > r$:

$$\frac{1}{|\boldsymbol{r} - \boldsymbol{\rho}|} = \frac{1}{\rho} \sum_{i=0}^{\infty} \left(\frac{r}{\rho}\right)^i P_{i0} \left(\frac{\boldsymbol{r} \cdot \boldsymbol{\rho}}{r\rho}\right) \tag{2.31}$$

If a finite density is assumed, then the contribution of the gravity field at precisely $r = |\boldsymbol{\rho}|$ vanishes. If this expansion is integrated over the mass distribution outside of radius r another solution of Laplace's equation can be found as a polynomial in r. Generally the coefficients of this external expansion are not used as when the test point radius moves within the maximum radius of the mass distribution, the gravity coefficients of each expansion become a function of the radius r, meaning that they must be recomputed as r changes as the mass distributions being integrated over change. This functional relationship means that the gravity coefficients must then be computed and tabulated at each radius. Recent work has been performed on such interior expansions, reported in [192].

These limitations mean that the usual spherical harmonic expansion cannot be used when considering the gravitational potential close to a small body with irregular shape. It is interesting to note that for an ellipsoid with eccentricities limited to less than $1/\sqrt{2}$, the external gravitational potential will converge [8]. This convergence is fragile and does not survive if the shape is distended sufficiently from a sphere or ellipsoid, and is discussed in more detail in [193]. For a general body the divergence is severe once one comes within the circumscribing sphere of the body (sometimes called the Brillouin sphere, see Fig. 2.10) and makes the gravitational potential essentially worthless for dynamical computations.

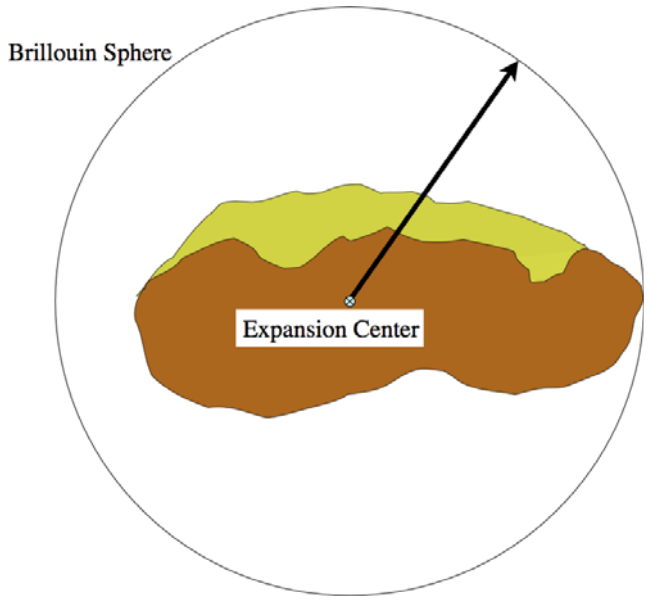


Fig. 2.10 Illustration of the circumscribing sphere about a body, otherwise known as the Brillouin sphere.

One innovative approach to remedying this situation while still using an expansion is the use of ellipsoid harmonic expansions to model the gravitational field, studied and applied to the small body gravitational potential problem in [43]. While ellipsoidal harmonic expansions suffer the same convergence problems, they are mitigated to some degree by having the circumscribing ellipsoid define the boundary between convergence and divergence of the external series. When on or close to the surface there are still divergence issues, but the potential can in general be used much closer to the surface. The cost of this approach is the complexity involved with the computation of the ellipsoidal harmonics orthogonal functions. Ultimately, if one wishes to describe the environment on the surface of a small body this approach also fails.

2.5.2 Closed-Form Gravitational Potentials

To circumvent this limitation there are a number of approaches that rely on the existence of a specified shape for the body in question and strong assumptions on density distribution. The ideal approach to this problem is to find a closed-form solution to Laplace's equation. There are several classical solutions to this problem that can be applied to the gravitational potential problem. At simplest, the sphere with a radially-varying density distribution has a known solution, similar in form to the point mass potential. Of more interest are the closed-form solutions for the constant density ellipsoid and the constant density polyhedron. These have been used in the majority of studies of motion close to and on the surface of small bodies. As exact solutions, they are valid up to and on the surface of the shape in question. They also satisfy Poisson's equation inside of the bodies, which is the analog to the gravitational potential inside of a mass distribution. The definition of these potentials is given in the following.

Constant Density Ellipsoid

Consider a constant density ellipsoid with semi-major axes $\gamma \leq \beta \leq \alpha$. Then the shape of this body is defined by the equation $(x/\alpha)^2 + (y/\beta)^2 + (z/\gamma)^2 \leq 1$ (Fig. 2.11). The total mass of a constant density ellipsoid is $M = 4\pi/3\sigma\alpha\beta\gamma$ and $\mu = \mathcal{G}M$.

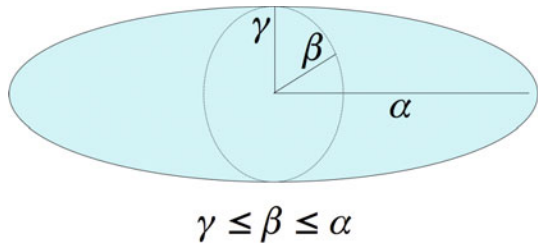


Fig. 2.11 Tri-axial ellipsoid.

The gravitational potential of the exterior region can be found by an application of Ivory's Theorem [112]. Without derivation the ellipsoidal gravitational potential is defined as

$$U(\boldsymbol{r}) = -\frac{3\mu}{4}\int_{\lambda(\boldsymbol{r})}^{\infty} \phi(\boldsymbol{r}, u)\frac{du}{\Delta(u)} \tag{2.32}$$

$$\phi(\boldsymbol{r}, u) = \frac{x^2}{\alpha^2 + u} + \frac{y^2}{\beta^2 + u} + \frac{z^2}{\gamma^2 + u} - 1 \tag{2.33}$$

$$\Delta(u) = \sqrt{(\alpha^2 + u)(\beta^2 + u)(\gamma^2 + u)} \tag{2.34}$$

where the parameter $\lambda(\boldsymbol{r})$ is defined by the equation $\phi(\boldsymbol{r}, \lambda) = 0$. This equation is equivalent to a cubic polynomial and λ is defined to be the maximum real root, which will always exist. The position vector $\boldsymbol{r}$ is specified in the principal axis frame with the x-axis along the long axis α, y along the intermediate axis β, and z along the short axis γ. The partial derivatives of this potential with respect to the coordinates can be computed by applying Leibniz's rule:

$$U_x = -\frac{3\mu x}{2} \int_{\lambda(\boldsymbol{r})}^{\infty} \frac{du}{(\alpha^2 + u)\Delta(u)} \tag{2.35}$$

$$U_y = -\frac{3\mu y}{2} \int_{\lambda(\boldsymbol{r})}^{\infty} \frac{du}{(\beta^2 + u)\Delta(u)} \tag{2.36}$$

$$U_z = -\frac{3\mu z}{2} \int_{\lambda(\boldsymbol{r})}^{\infty} \frac{du}{(\gamma^2 + u)\Delta(u)} \tag{2.37}$$

The term λ_x does not appear as the integrand is, by definition, zero when evaluated at λ. These terms can no longer be ignored when computing the second partials of the potential. Doing so yields

$$\begin{aligned} U_{xx} = &-\frac{3\mu}{2} \int_{\lambda(\boldsymbol{r})}^{\infty} \frac{du}{(\alpha^2 + u)\Delta(u)} \\ &+ \frac{3\mu x^2}{(\alpha^2 + \lambda)^2 \Delta(\lambda)} \frac{1}{\left[\frac{x^2}{(\alpha^2+\lambda)^2} + \frac{y^2}{(\beta^2+\lambda)^2} + \frac{z^2}{(\gamma^2+\lambda)^2}\right]} \end{aligned} \tag{2.38}$$

$$\begin{aligned} U_{yy} = &-\frac{3\mu}{2} \int_{\lambda(\boldsymbol{r})}^{\infty} \frac{du}{(\beta^2 + u)\Delta(u)} \\ &+ \frac{3\mu y^2}{(\beta^2 + \lambda)^2 \Delta(\lambda)} \frac{1}{\left[\frac{x^2}{(\alpha^2+\lambda)^2} + \frac{y^2}{(\beta^2+\lambda)^2} + \frac{z^2}{(\gamma^2+\lambda)^2}\right]} \end{aligned} \tag{2.39}$$

$$\begin{aligned} U_{zz} = &-\frac{3\mu}{2} \int_{\lambda(\boldsymbol{r})}^{\infty} \frac{du}{(\gamma^2 + u)\Delta(u)} \\ &+ \frac{3\mu z^2}{(\gamma^2 + \lambda)^2 \Delta(\lambda)} \frac{1}{\left[\frac{x^2}{(\alpha^2+\lambda)^2} + \frac{y^2}{(\beta^2+\lambda)^2} + \frac{z^2}{(\gamma^2+\lambda)^2}\right]} \end{aligned} \tag{2.40}$$

$$U_{xy} = \frac{3\mu xy}{(\alpha^2 + \lambda)(\beta^2 + \lambda)\Delta(\lambda)} \frac{1}{\left[\frac{x^2}{(\alpha^2+\lambda)^2} + \frac{y^2}{(\beta^2+\lambda)^2} + \frac{z^2}{(\gamma^2+\lambda)^2}\right]} \tag{2.41}$$

$$U_{xz} = \frac{3\mu xz}{(\alpha^2 + \lambda)(\gamma^2 + \lambda)\Delta(\lambda)} \frac{1}{\left[\frac{x^2}{(\alpha^2+\lambda)^2} + \frac{y^2}{(\beta^2+\lambda)^2} + \frac{z^2}{(\gamma^2+\lambda)^2}\right]} \tag{2.42}$$

$$U_{yz} = \frac{3\mu yz}{(\beta^2 + \lambda)(\gamma^2 + \lambda)\Delta(\lambda)} \frac{1}{\left[\frac{x^2}{(\alpha^2+\lambda)^2} + \frac{y^2}{(\beta^2+\lambda)^2} + \frac{z^2}{(\gamma^2+\lambda)^2}\right]} \tag{2.43}$$

The parameter $\lambda = \lambda(\boldsymbol{r})$ is always a function of the position vector, this dependency has been suppressed for conciseness in the above.

The computation of the ellipsoidal gravity field appears to be somewhat challenging, given its general formulation as an integral. However, the necessary terms to be computed uniformly fall into the general form of functions known as Carlson's Elliptic Integrals, as summarized (with associated computational algorithms) in [136]. Further, these can be re-expressed in terms of the Jacobi elliptic integrals of the first and second kind. In practice, computation as Carlson's Elliptic Integrals is numerically the most robust approach and can be rapidly evaluated to specified degrees of precision.

Due to the symmetry in an ellipsoid's shape, the gravitational coefficients of its spherical harmonic expansion taken about its center of mass have a relatively simple form. First, all S_{lm} coefficients are identically equal to zero. Second, all coefficients C_{lm} such that either l or m are odd are equal to zero as well. Thus, the only non-zero gravity coefficients are those of the form $C_{2l,2m}$, $l, m = 0, 1, 2, \ldots$. In [8] an explicit formula for the ellipsoid's gravity field coefficients are given, with the first few terms specified as:

$$C_{20} = \frac{1}{5r_o^2}\left(\gamma^2 - \frac{\alpha^2+\beta^2}{2}\right) \tag{2.44}$$

$$C_{22} = \frac{1}{20r_o^2}\left(\alpha^2 - \beta^2\right) \tag{2.45}$$

$$C_{40} = \frac{15}{7}\left(C_{20}^2 + 2C_{22}^2\right) \tag{2.46}$$

$$C_{42} = \frac{5}{7}C_{20}C_{22} \tag{2.47}$$

$$C_{44} = \frac{5}{28}C_{22}^2 \tag{2.48}$$

Constant Density Polyhedron

In a series of papers Werner detailed closed-form solutions for arbitrary constant density polyhedra [190, 193, 191]. Although the potential is given in closed form, the specific computations necessary to implement this potential are complex enough to require computer code to be written. This particular formulation of the gravity potential was used operationally for the NEAR mission, and has been used in a number of studies of the dynamical environment about asteroids. If local density variations are present or specified, it is possible to add additional potential terms to model such density heterogeneity [170].

The following summarizes the results in [193] for an arbitrary polyhedron with triangular faces. Assume such a body is specified, then from the Euler–Descartes formula for a polyhedron with triangular faces specified by v vectors there must be $f = 2v-4$ faces and $e = 3(v-2)$ edges. In Werner's approach the polyhedron is split into its edges and faces. Each face is associated with a set of three vertex vectors $\boldsymbol{r}_i^f$, $i = 1, 2, 3$, such that the three vertices taken in order are counter-clockwise about the normal to the face, $\hat{\boldsymbol{n}}_f$. Associated with each edge e are two vertices $\boldsymbol{r}_i^e$,

$i = 1, 2$, and two faces, f and f', which join at the edge. Denote the edge normal corresponding to face f as the unit vector $\hat{\boldsymbol{n}}_e^f$ perpendicular to the edge and to $\hat{\boldsymbol{n}}_f$ and pointing away from the face center. Thus the edge normals $\hat{\boldsymbol{n}}_e^f$ and $\hat{\boldsymbol{n}}_e^{f'}$ point in different directions and are not parallel to each other in general. Given these definitions the general formula for the potential of a polyhedron can be stated as [193]

$$U(\boldsymbol{r}) = \frac{\mathcal{G}\sigma}{2}\left[\sum_{e\in\text{edges}} \boldsymbol{r}_e \cdot \boldsymbol{E}_e \cdot \boldsymbol{r}_e L_e - \sum_{f\in\text{faces}} \boldsymbol{r}_f \cdot \boldsymbol{F}_f \cdot \boldsymbol{r}_f \omega_f\right] \tag{2.49}$$

$$\frac{\partial U}{\partial \boldsymbol{r}} = -\mathcal{G}\sigma\left[\sum_{e\in\text{edges}} \boldsymbol{E}_e \cdot \boldsymbol{r}_e L_e - \sum_{f\in\text{faces}} \boldsymbol{F}_f \cdot \boldsymbol{r}_f \omega_f\right] \tag{2.50}$$

$$\frac{\partial^2 U}{\partial \boldsymbol{r}^2} = \mathcal{G}\sigma\left[\sum_{e\in\text{edges}} \boldsymbol{E}_e L_e - \sum_{f\in\text{faces}} \boldsymbol{F}_f \omega_f\right] \tag{2.51}$$

where $\boldsymbol{r}_e$ denotes the vector from any point in the edge e to $\boldsymbol{r}$ and similarly $\boldsymbol{r}_f$ denotes the vector from any point in the face f to $\boldsymbol{r}$. The following terms are defined as

$$\boldsymbol{E}_e = \hat{\boldsymbol{n}}_f \hat{\boldsymbol{n}}_e^f + \hat{\boldsymbol{n}}_{f'} \hat{\boldsymbol{n}}_e^{f'} \tag{2.52}$$

$$\boldsymbol{F}_f = \hat{\boldsymbol{n}}_f \hat{\boldsymbol{n}}_f \tag{2.53}$$

$$L_e = \ln \frac{r_1^e + r_2^2 + e_e}{r_1^e + r_2^2 - e_e} \tag{2.54}$$

$$\omega_f = 2 \arctan \frac{\boldsymbol{r}_1^f \cdot \tilde{\boldsymbol{r}}_2^f \cdot \boldsymbol{r}_3^f}{r_1^f r_2^f r_3^f + r_1^f \boldsymbol{r}_2^f \cdot \boldsymbol{r}_3^f + r_2^f \boldsymbol{r}_3^f \cdot \boldsymbol{r}_1^f + r_3^f \boldsymbol{r}_1^f \cdot \boldsymbol{r}_2^f} \tag{2.55}$$

where e_e denotes the actual length of the edge, or $|\boldsymbol{r}_1^e - \boldsymbol{r}_2^e|$.

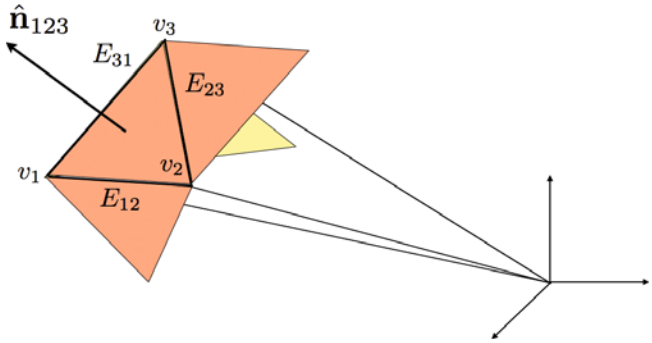


Fig. 2.12 Three vertices, three edges and one surface normal associated with a single facet.

The Laplacian of the potential can be shown to be

$$\nabla^2 U = -\mathcal{G}\sigma \sum_{f \in \text{faces}} \omega_f \tag{2.56}$$

The term ω_f denotes the signed area of the face f projected onto the unit sphere centered at the point $\boldsymbol{r}$. Thus, when the point is outside of the body the total signed projection equals 0, and when inside the body it equals 4π and thus the polyhedron potential satisfies Laplace's and Poisson's equations. Due to this, the potential and partials can be used to compute the gravitational field down to and on the surface of an arbitrary body. Note that, due to this property of the Laplacian of U it is also possible to determine whether a material point is outside or inside of a body by evaluating the Laplacian, thus providing a rigorous manner to determine when impact with a body has occurred.

It is significant to note that closed-form expressions for the spherical harmonic gravity coefficients can be computed given a polyhedron shape model. The procedure and code for performing these computations is detailed in [191]. As the computation of the polyhedron gravity field can be computationally expensive, the use of a spherical harmonic expansion of this field when outside of the circumscribing radius of the body can yield significant computational speed increases. As the harmonic series becomes divergent at the circumscribing sphere, usual practice is to make the transition between polyhedron and spherical harmonic field at a radius larger than the circumscribing sphere.

2.5.3 Other Gravity Field Representations

There are a number of other approximation approaches for the gravitational potential of an arbitrary body that have been used. Most of these use a collection of simple closed-form potential solutions combined in a manner that mimics the gravity field of an actual body. There is scientific heritage for this approach, dating to the use of "mascons" to estimate density inhomogeneities within the lunar gravity field [114]. The most common approach to this problem is to approximate the mass distribution by a collection of point masses. Some past research has taken an existing shape model and filled it with a collection of point masses to mimic the total gravity field. This approach is considered in [193] but shown to be computationally inefficient as compared to the polyhedron model and to provide relatively poor modeling accuracy at the surface of the asteroid. One problem with this approach is that the gravity field of a "cube" of material is modeled as a "sphere". These two objects have demonstrably different gravity fields when in close proximity leading to systematic errors. The approach was also used by the navigation team for the Hayabusa mission to asteroid Itokawa [101] to approximate the irregular gravity field of that object. The precision of such models are low, however, and it is difficult to uniquely estimate the location and masses of these particles while at the same time finding a robust fit for the true gravity field of an object.

A second approach developed recently is to use the logarithmic potential, which is the gravitational potential of a line element with linear density. There is a simple closed form for these potentials which only involves the use of natural logarithms. By constructing a collection of these it is possible to mimic other gravity fields, especially when point masses are also incorporated. Some relevant research on this topic is reported in [9]. While this can be used to approximately represent a gravity field, no systematic estimation procedure for their placement has been developed as of yet. Thus, their use in practical navigation scenarios is very limited. They are useful for constructing simple systems with which to simulate the dynamical evolution of bodies in non-point-mass potentials.

2.6 Other Forces and Perturbations

When dealing with the orbital motion of a particle in the small body environment the dominant force is usually from the small body gravitational field. Significant perturbations also arise from other gravitational and non-gravitational sources, however. The prime source of these perturbations are the Sun, both gravitational and non-gravitational. Additional sources of perturbation include planetary gravitational perturbations, which could become significant when the small body has a close passage by a planet.

2.6.1 Third-Body Gravitational Perturbations

Solar and planetary gravitational perturbations are simple to specify, if one assumes the small body, or small body system center of mass, lies in a two-body orbit relative to the sun or planet. This is a reasonable approximation, although it tacitly assumes that we will not deal with more complex situations where both planetary and solar gravitational perturbations become significant (which would be a 4+ body problem). The absolute gravitational attraction that a particle would experience is simply $-\left[(\mathcal{G}M_p)/(|\boldsymbol{r}-\boldsymbol{d}|^3\right](\boldsymbol{r}-\boldsymbol{d})$ where M_p is the mass of the perturbing body, $\boldsymbol{r}$ is the vector from the small body center of mass to the particle, and $\boldsymbol{d}$ is the position vector from the small body center of mass to the perturbing body. In this formulation the sun/perturber is viewed as orbiting the small body, which is only a change of perspective and just implies that the relative position vector goes from the small body to the perturber. The absolute acceleration is not the relevant acceleration to consider for finding the motion of the particle relative to the small body. For that situation the relative or perturbing acceleration is needed, which is just the difference between the absolute acceleration that the particle feels and the absolute acceleration which the small body feels, $+\left[(\mathcal{G}M_p)/|\boldsymbol{d}|^3\right]\boldsymbol{d}$, where $\boldsymbol{d}$ is the position vector of the perturbing body relative to the central body. The

difference of these forms the perturbation gravitational acceleration from a body with mass M_p:

$$\boldsymbol{a}_p = -\mathcal{G}M_p \left[\frac{(\boldsymbol{r}-\boldsymbol{d})}{|\boldsymbol{r}-\boldsymbol{d}|^3} + \frac{\boldsymbol{d}}{|\boldsymbol{d}|^3} \right] \tag{2.57}$$

For use in perturbation analysis it is convenient to recast this in terms of a perturbing acceleration potential

$$\mathcal{R}_p = \mathcal{G}M_p \left[\frac{1}{|\boldsymbol{r}-\boldsymbol{d}|} - \frac{\boldsymbol{d}\cdot\boldsymbol{r}}{|\boldsymbol{d}|^3} \right] \tag{2.58}$$

where

$$\boldsymbol{a}_p = \frac{\partial \mathcal{R}_p}{\partial \boldsymbol{r}} \tag{2.59}$$

Of particular note, if the distance between the central body and the perturbing body, d, is much larger than the distance of the particle from the central body, or $d \gg r$, the perturbing potential can be expanded using the Legendre expansion:

$$\mathcal{R}_p = \frac{\mu_p}{d} \left[\sum_{i=0}^{\infty} \left(\frac{r}{d}\right)^i P_{i,0}(\boldsymbol{r}\cdot\boldsymbol{d}/rd) - \frac{\boldsymbol{d}\cdot\boldsymbol{r}}{|\boldsymbol{d}|^2} \right] \tag{2.60}$$

From Appendix E the Legendre Polynomial $P_{1,0}(x) = x$, which leads to a cancellation between the $i = 1$ term and the second term in the summation. Also, $P_{0,0}(x) = 1$, and thus the first term of the summation has only a constant contribution. Thus, the perturbation terms that matter do not start until $i = 2$, resulting in the perturbation acceleration:

$$\mathcal{R}_p = \frac{\mu_p}{d} \sum_{i=2}^{\infty} \left(\frac{r}{d}\right)^i P_{i,0}(\boldsymbol{r}\cdot\boldsymbol{d}/rd) \tag{2.61}$$

When this expansion is used, usually only the first term is kept. Evaluating this term explicitly, with the Legendre polynomial $P_{2,0}(x) = \frac{1}{2}\left(3x^2 - 1\right)$, yields

$$\mathcal{R}_p = \frac{1}{2}\frac{\mu_p}{d^3} \left[3(\boldsymbol{r}\cdot\hat{\boldsymbol{d}})^2 - r^2 \right] \tag{2.62}$$

This expansion is the fundamental approximation made in the Hill problem, and will be applied later in this book. It should be noted that the definition of whether $\boldsymbol{d}$ points from the central body to the perturber or from the perturber to the central body makes no difference in this form of the perturbation. This perturbation, when combined with a rotating reference frame, creates the tidal acceleration discussed in Chapters 12, 17, 18.

If multiple bodies are attracting the spacecraft, a different perturbation potential can be specified for each one, and they can be summed together to create the full potential. It is important to note that the perturbation potential is time-varying

in general, as the position $\boldsymbol{d}$ varies with time. In this formulation the motion of $\boldsymbol{d}$ with time can either be supplied from theory, i.e. the two-body solution, or can be provided by an ephemeris.

2.6.2 Solar Photon Perturbation

Also associated with the sun and the heliocentric orbit are non-gravitational perturbations due to momentum transfer from solar photons striking and recoiling off the orbiting body. These non-gravitational forces also act on the small body orbit through the Yarkovsky effect [144], although the relevant timespan involved is many orders of magnitude longer than what interests us here. These same non-gravitational forces also induce a torque on the small body and can modify its rotation state over time, called the YORP effect [145]. Again, the time scale for these are usually much longer than orbital time scales about the small body. If it is desired to incorporate these effects, it is in general sufficient to allow the orbit or rotation state of the small body to change slowly over time without accounting for the rates of change of these quantities in the equations of motion for the orbiting body.

With regard to the spacecraft, the most general and relevant description of these solar forces come from the solar sail literature [103, 140]. From a momentum budget point of view, there are four main contributors to the total force acting on a spacecraft or particle from the incident solar photons. First is the momentum of the impinging photons on the surface. This component acts in the direction of the traveling photons, i.e., along the line from the sun to the particle. Next is the recoil momentum of those photons specularly reflected from the surface. This fraction is found by multiplying the incident radiation by two terms. The first is ρ which is the total reflectance or albedo of the body in question. The next is a factor s which is the fraction of specularly reflected light (i.e., the portion that is like a mirror). Thus the total multiplier is ρs. The direction of the reflected light depends on the local orientation of the surface, with the surface normal bisecting the angle between the incident and specularly reflected light. Third is the recoil momentum of the photons diffusely scattered from the surface. There are two factors involved here, the first being the complement of the specularly reflected light, $\rho(1 - s)$. The second is a scattering coefficient B that describes the fraction of light scattered normal to the surface. For an ideal Lambertian surface this factor is 2/3, although for a real body this scattering factor will deviate from this ideal value. This reflected light leaves perpendicular to the surface on average, providing an additional momentum transfer against this direction. The final contribution arises from thermal considerations. Of the incident photons, a fraction $(1 - \rho)$ are absorbed by the body and raise its temperature. The body subsequently re-radiates these photons, on average a balance existing between the influx of photons, the temperature of the body, and the outflow of re-emitted photons. For a thin surface the relative emittance of the forward and back surfaces must also be accounted for. The direction of this net momentum transfer is also along the local normal to the surface with its magnitude again decreased by the factor 2/3 for a Lambertian surface. Finally, the total

radiation pressure of the incident sunlight must be modeled. A typical model is to assume the sun acts as a point source, leading to a simple total pressure:

$$P(d) = \frac{P_0}{d^2} \tag{2.63}$$

where P_0 is a solar constant approximately equal to 1×10^8 kg km^3/s^2/m^2. Combined together, the solar radiation force acting on a unit area A is then

$$\begin{aligned} \boldsymbol{F} = -P(d)\big[\big\{\rho s(2\hat{\boldsymbol{n}}\hat{\boldsymbol{n}} - \boldsymbol{U}) + \boldsymbol{U}\big\} \cdot \hat{\boldsymbol{u}}\hat{\boldsymbol{u}} \cdot \hat{\boldsymbol{n}} \\ + \big\{B(1-s)\rho + (1-\rho)B\big\}\hat{\boldsymbol{n}}\hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{u}}\big] H(\hat{\boldsymbol{u}})A \end{aligned} \tag{2.64}$$

where $\hat{\boldsymbol{n}}$ is the unit normal to the surface, $\hat{\boldsymbol{u}}$ is the unit vector from the surface to the sun, and $H(\hat{\boldsymbol{u}})$ is the visibility function for the surface and is equal to 1 when the sun is in view and 0 otherwise. When summed over all surface elements of the spacecraft and divided by the total mass this yields the net acceleration due to the solar photons.

As opposed to this more complex model, the simplest model for computing solar radiation accelerations is to assume that the spacecraft presents a constant area perpendicular to the sun-line, and that the total momentum transfer is modeled as insolation plus reflection. Then the net acceleration will act away from the sun-line and have the general form

$$\boldsymbol{a}_{SRP} = -\frac{(1+\rho)P_0 A_{SC}}{M_{SC}} \frac{(\boldsymbol{d} - \boldsymbol{r})}{|\boldsymbol{d} - \boldsymbol{r}|^3} \tag{2.65}$$

where again $\boldsymbol{d}$ is the vector from the small body to the sun. An important parameter is the mass to area ratio, denoted here as $B_{SC} = M_{SC}/A_{SC}$, as it controls the relative strength of this perturbation. A typical range of values of B_{SC} for spacecraft is tabulated in [171] and ranges between 20 and 40 kg/m^2. Spacecraft with large solar arrays will have smaller values of B, and will be more susceptible to these perturbations, which can often be the dominant perturbation for an orbiter.

This simple form of the solar radiation pressure model can be rewritten as a potential

$$\mathcal{R}_{SRP} = -\frac{(1+\rho)P_0}{B_{SC}} \frac{1}{|\boldsymbol{d} - \boldsymbol{r}|} \tag{2.66}$$

with $\boldsymbol{a}_{SRP} = \partial \mathcal{R}_{SRP}/\partial \boldsymbol{r}$. If the spacecraft is close to the small body, or $r \ll d$, the potential can be further simplified by expanding $1/|\boldsymbol{d} - \boldsymbol{r}|$ and keeping the first term that contains the position vector $\boldsymbol{r}$.

$$\mathcal{R}_{SRP} = -\frac{(1+\rho)P_0}{B_{SC}} \frac{\boldsymbol{r} \cdot \boldsymbol{d}}{d^3} \tag{2.67}$$

with the gradient giving a solar radiation pressure acceleration independent of the spacecraft's position relative to the small body

$$\boldsymbol{a}_{SRP} \sim -\frac{(1+\rho)P_0}{B_{SC}d^3}\boldsymbol{d} \tag{2.68}$$

which is often an acceptable approximation for our applications. The magnitude of the solar radiation pressure acceleration is then equal to

$$a_{SRP} = \frac{(1+\rho)P_0}{B_{SC}d^2} \tag{2.69}$$

This model is approximate, but is commonly used to model the non-gravitational accelerations acting on satellites. Higher-precision developments of spacecraft solar radiation pressure or non-gravitational models have been developed [202, 140, 104], and are appropriate for use at a small body. Even with these more general formulations, however, a spacecraft in the vicinity of a small body will frequently be oriented to keep a constant projected area towards the sun to minimize deviations in the non-gravitational perturbations. Thus, in this case it is appropriate to still model the solar radiation pressure as a constant force, perhaps generalized so that the acceleration vector does not necessarily lie along the sun-line.

2.6.3 Comet Outgassing Pressure Models

Finally, a simple model that accounts for the effect of outgassing pressure from a comet acting on an orbiting body is presented. It must be stressed that the detailed modeling and physics of the flow of gases sublimated from the surface of a comet is still an active area of research, and thus the models presented here are simplistic and not grounded in any *in situ* measured comet models. A general description for the construction of comet outgassing models is given in [23], and is summarized here. This model assumes that the outgassing jet emanates from the rotating comet surface and travels radially outwards and expands. Further, the strength of the outgassing jet is a function of the distance of the comet from the sun (which accounts for the total solar insolation on the comet surface) and the relative orientation of the jet site from the sun (which accounts for the relative solar insolation).

The total acceleration felt by an orbiting spacecraft is expressed as

$$\boldsymbol{a}_J = \frac{\boldsymbol{p}_0(\boldsymbol{r}_{SJ})}{B_{SC}}\left(\frac{r_o}{r}\right)^2 \tag{2.70}$$

where $\boldsymbol{p}_0$ is the pressure field and direction at the comet surface, $\boldsymbol{r}_{SJ}$ is the position vector of the spacecraft relative to the center of the jet on the comet surface and r_o is the reference radius of the comet. The parameter B_{SC} is again the spacecraft mass to area ratio, now computed for the spacecraft area exposed to the outgassing jet. This model assumes that the jet strength varies as $1/r^2$ from the comet surface, which may not be particularly valid yet provides a tractable and conservative model

as in reality the drop off may be significantly faster. In the detailed model the pressure field $\boldsymbol{p}_0$ spans a cone about the center of the jet; however, for practical computations it often suffices to treat the pressure field as concentrated along a line in space.

A practical implementation of the model specified in [23] is

$$\boldsymbol{p}_0 = Q_J V_{og} h(\hat{\boldsymbol{e}}_J \cdot \hat{\boldsymbol{r}}_J)\hat{\boldsymbol{r}}_J \tag{2.71}$$

$$h(\hat{\boldsymbol{e}}_J \cdot \hat{\boldsymbol{r}}_{SJ}) = \begin{cases} 1 & \hat{\boldsymbol{e}}_J \cdot \hat{\boldsymbol{r}}_J > \cos(\delta_J) \\ 0 & \hat{\boldsymbol{e}}_J \cdot \hat{\boldsymbol{r}}_J < \cos(\delta_J) \end{cases} \tag{2.72}$$

The angle δ_J is the half-angle of the outgassing jet, as measured from a point on the surface, $\boldsymbol{e}_J$ is the unit vector defining the direction of the jet relative to the comet surface, $\boldsymbol{r}_{SJ} = \boldsymbol{r} - \boldsymbol{r}_J$ is the relative vector between the spacecraft and the jet's location on the surface of the comet, V_{og} is the speed of the gas as it leaves the surface, and Q_J is a combination of parameters

$$Q_J = Sf(\theta_S)g(d)Q_* \tag{2.73}$$

Here Q_* is the mass ejection rate of the comet surface at 1 AU, S is the relative intensity of the comet surface at the jet site, $g(d)$ is a complex function that models the outgassing strength as a function of distance to the sun, and $f(\theta_S)$ computes the relative insolation of the jet site as a function of θ_S, the angle between the jet surface normal and the sun direction. The function form of $g(d_{AU})$ is an empirically derived function taken from [99]

$$g(d_{AU}) = g_0 \left(\frac{d_{AU}}{d_0}\right)^{-c_1} \left[1 + \left(\frac{d_{AU}}{d_0}\right)^{c_2}\right]^{-c_3} \tag{2.74}$$

where $c_1 = 2.15$, $c_2 = 5.093$, $c_3 = 4.6142$, $d_0 = 2.808$ and $g_0 = 0.111262$, and d_{AU} is measured in astronomical units. For the insolation function $f(\theta_S)$ a simple formula can be used

$$f(\theta_S) = \max \begin{cases} 0 \\ \\ 1 - \alpha_J \left(1 - \theta_S\right) \end{cases} \tag{2.75}$$

Here α_J is a parameter between 0 and 1 that controls the outgassing strength as a function of solar insolation, and $\cos(\theta_S) = \boldsymbol{e}_J \cdot \hat{\boldsymbol{d}}$ is the cosine of the angle between the jet surface normal and the sun. The values for the outgassing parameters vary widely from reference to reference. In [23] the following values are used, $V_{og} = 0.35 \rightarrow 0.95$ km/s and $Q_* = 3.4 \times 10^6$ kg/h, this value varying strongly with the comet in general.

The model presented here is wholly based on a heuristic understanding of comet outgassing jets. It also assumes that the gas velocity is large relative to the speed of the spacecraft relative to the comet, and that the spacecraft is relatively close to the comet. If these are not true, then it is necessary to model the time between

gas ejection and the spacecraft crossing into the jet stream, detailed in [23]. It also assumes that the comet rotation is being modeled, meaning that the inertial location of the vectors $\boldsymbol{r}_J$ and $\hat{\boldsymbol{e}}_J$ varies in time with the rotation of the comet.

A simplified form of this model, for use in analytical evaluation, assumes that the jet normal points radially outwards from the center of the comet (modeled as a sphere) and that the crossing of the jet occurs impulsively. Under these approximations the outgassing jet model simplifies to

$$\boldsymbol{a}_J = \frac{Q_J V_{og}}{B_{SC}} \delta(1 - \hat{\boldsymbol{r}}_J \cdot \hat{\boldsymbol{r}}) \frac{\boldsymbol{r}}{r^3} \tag{2.76}$$

where $\delta(x)$ is the Dirac delta function.

Improved models of comet outgassing are currently under development and will be tested and refined once the Rosetta spacecraft arrives at comet Churyumov-Gerasimenko. A current description of some of these modeling developments can be found in [115]. The main advance in these models is the inclusion of non-radial components of the gas drag, based on sophisticated models of the cometary outgassing process.

Part II
Dynamics

3. Non-Perturbed Solutions

At the heart of our dynamical analyses are known solutions to fundamental dynamical problems of orbital and rotational motion. In this chapter we present the two core solutions necessary for more in-depth study of the problem: the two-body problem under mutual gravitation and the rotational dynamics of a torque-free rigid body. These are both classical solutions, but we detail their derivation here for completeness and to introduce specific concepts and notations used in this book.

3.1 The Two-Body Problem

An understanding of the two-body problem is required to motivate our dynamical understanding of motion about small bodies, as this solution serves as the starting point for many theories of motion. Thus, instead of providing a description of the solution we provide a brief but thorough derivation of the two-body problem.

3.1.1 The Two-Body Problem Statement

Assume that two bodies P_1 and P_2, with mass M_1 and M_2 and position vectors $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, respectively, are attracting each other according to Newton's law of gravitation. If the mutual gravitational potential between these bodies is derived assuming both bodies have spherical mass distributions we find the following dynamical system:

$$M_i \ddot{\boldsymbol{r}}_i = \frac{\partial \mathcal{U}}{\partial \boldsymbol{r}_i} \tag{3.1}$$

$$i = 1, 2$$

$$\mathcal{U} = \frac{\mathcal{G} M_1 M_2}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|} \tag{3.2}$$

where $\mathcal{U}$ is the gravitational force potential (opposite in sign to the potential energy).

D.J. Scheeres, *Orbital Motion in Strongly Perturbed Environments*, Springer Praxis Books,
DOI 10.1007/978-3-642-03256-1_3, © Springer-Verlag Berlin Heidelberg 2012

3.1.2 Classical Integrals of Motion

The system conserves translational and angular momentum, as well as energy. Conservation of translational momentum allows us to reduce the problem by three degrees of freedom by recasting it in terms of its center of mass and the relative motion between the two bodies. Specifically, by choosing new coordinates:

$$\boldsymbol{r}_C = \frac{1}{M_1 + M_2}\left(M_1\boldsymbol{r}_1 + M_2\boldsymbol{r}_2\right) \tag{3.3}$$

$$\boldsymbol{r} = \boldsymbol{r}_2 - \boldsymbol{r}_1 \tag{3.4}$$

it is easy to show that the new equations of motion for these variables becomes

$$\ddot{\boldsymbol{r}}_C = 0 \tag{3.5}$$

$$\ddot{\boldsymbol{r}} = \frac{M_1 + M_2}{M_1 M_2}\frac{\partial \mathcal{U}}{\partial \boldsymbol{r}} \tag{3.6}$$

Define the relative potential for the system as:

$$U = \frac{M_1 + M_2}{M_1 M_2}\mathcal{U} \tag{3.7}$$

$$= \frac{\mathcal{G}(M_1 + M_2)}{|\boldsymbol{r}|} \tag{3.8}$$

which leads to the usual statement of the two-body problem in relative form:

$$\ddot{\boldsymbol{r}} = \frac{\partial U}{\partial \boldsymbol{r}} \tag{3.9}$$

The resulting system still conserves energy and angular momentum

$$\mathcal{E} = \frac{1}{2}\frac{M_1 M_2}{M_1 + M_2}\dot{\boldsymbol{r}}\cdot\dot{\boldsymbol{r}} - \mathcal{U}(\boldsymbol{r}) \tag{3.10}$$

$$\boldsymbol{\mathcal{H}} = \frac{M_1 M_2}{M_1 + M_2}\boldsymbol{r}\times\dot{\boldsymbol{r}} \tag{3.11}$$

More commonly used are the specific energy and angular momentum of the system, found by dividing through by the effective mass $M_1M_2/(M_1 + M_2)$,

$$E = \frac{1}{2}\boldsymbol{v}\cdot\boldsymbol{v} - U(\boldsymbol{r}) \tag{3.12}$$

$$\boldsymbol{H} = \boldsymbol{r}\times\boldsymbol{v} \tag{3.13}$$

where $\boldsymbol{v} = \dot{\boldsymbol{r}}$. Conservation of energy E relies on the potential U not being a function of time, while conservation of the angular momentum $\boldsymbol{H}$ relies on U corresponding to a central field, i.e., only a function of the distance between the two bodies. The simple two-body potential satisfies both of these conditions, as can be verified.

The geometrical interpretation of the angular momentum integral can be split into two pieces, and is most easily noted by rewriting it as $\boldsymbol{H} = H\hat{\boldsymbol{H}}$, the product of the magnitude and the unit vector in which it points, both of which must be constant. Constancy of $\hat{\boldsymbol{H}}$ implies that the solution to the two-body problem lies on a fixed plane perpendicular to $\hat{\boldsymbol{H}}$. This plane can be oriented relative to an inertial coordinate system by two angles, traditionally chosen as the inclination, i, which is the angle between $\hat{\boldsymbol{H}}$ and the $\hat{\boldsymbol{z}}$-axis of the coordinate frame, and the longitude of the ascending node Ω, which is the angle from the $\hat{\boldsymbol{x}}$-axis to the vector $\hat{\boldsymbol{z}} \times \hat{\boldsymbol{H}}$, which lies in the $\hat{\boldsymbol{x}} - \hat{\boldsymbol{y}}$ plane. The angular momentum unit vector can be specified using these two angles as:

$$\hat{\boldsymbol{H}} = \sin i \sin \Omega \hat{\boldsymbol{x}} - \sin i \cos \Omega \hat{\boldsymbol{y}} + \cos i \hat{\boldsymbol{z}} \tag{3.14}$$

The angular momentum being a constant of the motion, these two "orbital elements" are also constants.

The constancy of the orbit energy and angular momentum magnitude provide explicit relationships between the velocity and the position of the relative solution in the plane defined by $\hat{\boldsymbol{H}}$. If we define the angle between the position and velocity vector as $\pi/2 - \gamma$, where γ is commonly termed the flight path angle (see Fig. 3.1), then $H = rv\cos\gamma$. It is interesting to consider the projection of the velocity vector along the radius vector and perpendicular to it:

$$v_r = \boldsymbol{v} \cdot \hat{\boldsymbol{r}} \tag{3.15}$$

$$= v \sin\gamma \tag{3.16}$$

$$v_h = v\cos\gamma \tag{3.17}$$

$$= H/r \tag{3.18}$$

where v_r stands for radial and v_h stands for horizontal. If we introduce the angular rotation rate $\dot{\theta}$ of the position vector relative to inertial space, using an implied polar coordinate system in the orbit plane with center at body P_1, and note the familiar kinematic relationship $v_h = r\dot{\theta}$, we find

$$H = r^2\dot{\theta} \tag{3.19}$$

Next consider the orbit energy, E, and what constraints it places on the motion. For natural systems there is a strong distinction between positive and negative energy solutions, with fundamentally different behavior resulting between the two cases. Consider a system with total negative energy, $E < 0$. Then from the energy equation it is simple to derive an upper bound on the distance between the two bodies. For the two-body problem our specific argument rests on the fact that $\lim_{r\to\infty} U(\boldsymbol{r}) = 0$, $\partial U/\partial r < 0$ for all r, and $\lim_{r\to 0} U(\boldsymbol{r}) \to \infty$. Then, there exists a radius R_{max} such that $U(R) + E \leq 0$ for $R \geq R_{max}$. But $U(r) + E = \frac{1}{2}v^2 \geq 0$. Thus, $U(r) \geq U(R)$ which leads to $r \leq R_{max}$, an upper bound on the distance between the two bodies. When the system reaches this upper bound, the kinetic energy of the system is zero and the energy is completely contained in the potential energy of the system. For the case of positive energy there are no upper bounds

on the motion, and for the two-body problem all solutions will eventually travel to arbitrarily large separations between the bodies.

Combining the angular momentum magnitude and the energy integrals we can establish a lower bound on the energy for a given angular momentum. Specifically we find that for a given angular momentum H, the energy must be greater than or equal to $E_{min} = -\frac{\mu^2}{2H^2}$. To show this start with the supposition

$$E = \frac{1}{2}v^2 - \frac{\mu}{r} \geq -\frac{\mu^2}{2H^2} \tag{3.20}$$

Removing the speed v from the above inequality with the substitution $v = H/(r\cos\gamma)$, we can find the quadratic equation

$$r^2 - \frac{2H^2}{\mu}r + \frac{H^4}{\mu^2\cos^2\gamma} \geq 0 \tag{3.21}$$

and its factorization

$$r = \frac{H^2}{\mu}\left[1 \pm \sqrt{1 - \frac{1}{\cos^2\gamma}}\right] \tag{3.22}$$

Note that the radicand is less than or equal to 0, and thus that a physical value of radius r can never equal its root except when $\gamma = 0$ and $r = H^2/\mu$. At this value of γ and r, however, the energy is identically equal to E_{min}. For any other value of γ the energy is greater than E_{min}, establishing the result.

Continuing with the angular momentum and energy we can also find a lower bound on the radius. First, we note that $H > 0$ and that a minimum separation distance between the bodies exists. Then, at a minimum distance we must have $d(\boldsymbol{r}\cdot\boldsymbol{r})/dt = 0$, or $\boldsymbol{r}\cdot\boldsymbol{v} = 0$, and thus the flight path angle $\gamma = 0$ at a minimum of the separation radius leading to $H = r_p v_p$. Eliminating v_p from the energy and angular momentum then provides the result:

$$E = \frac{H^2}{2r_p^2} - \frac{\mu}{r_p} \tag{3.23}$$

which can be changed into the quadratic equation and its solution

$$r_p^2 + \frac{\mu}{E}r_p - \frac{H^2}{2E} = 0 \tag{3.24}$$

$$r_p = \frac{-\mu}{2E}\left[1 \pm \sqrt{1 + \frac{2EH^2}{\mu^2}}\right] \tag{3.25}$$

For a physical orbit the radicand must always be greater or equal to zero, given the lower limit on energy. Thus the solution always has at least one physical (i.e., real) root. If $E > 0$, then the radicand is always real and greater than 1, and the minus sign has a positive separation between the two bodies. If $E < 0$ we note that $E \geq E_{min}$ and the radicand is positive and less than 1. Thus in this case there are

two real solutions; inspection indicates that one is a minimum and one a maximum. Thus the minimum radius of the system can be denoted as:

$$r \geq r_p \tag{3.26}$$

$$r_p = \frac{\mu}{2E}\left[\sqrt{1+\frac{2EH^2}{\mu^2}}-1\right] \tag{3.27}$$

Conversely, whether there is a maximum radius depends on the energy of the system, leading to the result $r \leq r_a$ with

$$r_a = \begin{cases} \frac{-\mu}{2E}\left[1+\sqrt{1+\frac{2EH^2}{\mu^2}}\right] & E < 0 \\ \infty & E \geq 0 \end{cases} \tag{3.28}$$

3.1.3 Additional Integrals and the Orbit Trajectory

While strong constraints can be placed on the planar motion for the two-body motion, the solution is not completely described as of yet. The general n-body problem has the same integrals of motion discussed to this point, translational momentum, angular momentum, and energy, resulting in 10 separate integrals of motion that can be used to reduce the dimension of the system and place constraints on its motion. In the two-body problem we note that the system can be reduced to a one-degree-of-freedom system by use of the Clairaut solution substitution [6]. We will proceed in a different direction. First note the existence of an additional integral of motion which is only defined for the two-body problem, the Laplace vector or equivalently the eccentricity vector. Second note Jacobi's final integral theorem, which states that any system that has been reduced to a one-dimensional dynamical system can always be solved by quadratures.

The Laplace vector can be formed by the combination:

$$\boldsymbol{b} = \boldsymbol{v} \times \boldsymbol{H} - \mu\hat{\boldsymbol{r}} \tag{3.29}$$

while the eccentricity vector is just:

$$\boldsymbol{e} = \frac{\boldsymbol{b}}{\mu} \tag{3.30}$$

$$= \frac{1}{\mu}\boldsymbol{v} \times \boldsymbol{H} - \hat{\boldsymbol{r}} \tag{3.31}$$

It can be shown that $\dot{\boldsymbol{e}} = 0$ by direct substitution of the previous integrals of motion and the equations of motion. Specifically,

$$\dot{\boldsymbol{e}} = -\frac{1}{r^3}\boldsymbol{r} \times \boldsymbol{H} - \frac{1}{r}\left[\boldsymbol{U} - \hat{\boldsymbol{r}}\hat{\boldsymbol{r}}\right] \cdot \boldsymbol{v} \tag{3.32}$$

Note that $\boldsymbol{r} \times \boldsymbol{H} = \tilde{\boldsymbol{r}} \cdot \tilde{\boldsymbol{r}} \cdot \boldsymbol{v} = r^2 \left[\hat{\boldsymbol{r}}\hat{\boldsymbol{r}} - \boldsymbol{U}\right] \cdot \boldsymbol{v}$, showing that these terms are equal and opposite, leading to $\dot{\boldsymbol{e}} = 0$.

The eccentricity vector only has one additional item of information apart from the classical integrals. As a vector, it lies in the orbit plane of motion and its magnitude equals $e^2 = 1 + 2H^2E/\mu^2$, neither of which provides new information. The additional information that $\boldsymbol{e}$ provides is related to its orientation in the plane of motion, which provides an inertially fixed reference point for the relative motion between the bodies. Denote the angle between $\boldsymbol{e}$ and the node vector defined by the unitized vector $\hat{\boldsymbol{z}} \times \boldsymbol{H}$ as ω, this is called the argument of periapsis and is another orbital element.

To solve for the trajectory of the bodies, take the dot product of the constant vector $\boldsymbol{e}$ with the radius vector $\boldsymbol{r}$ to find:

$$\boldsymbol{r} \cdot \boldsymbol{e} = \frac{H^2}{\mu} - r \tag{3.33}$$

$$re\cos f = \frac{H^2}{\mu} - r \tag{3.34}$$

and where the angle between $\boldsymbol{r}$ and $\boldsymbol{e}$ is denoted as f and called the true anomaly. This equation can be explicitly solved for the distance between the two bodies to find:

$$r = \frac{H^2/\mu}{1 + e\cos f} \tag{3.35}$$

which gives us the relative distance between the two bodies as a function of the angle f measured from the $\boldsymbol{e}$ vector, which in turn is specified by the angle ω measured in the orbital plane from the node crossing (see Fig. 3.1 for a description of the geometry). This simple relationship captures Kepler's first law of planetary motion.

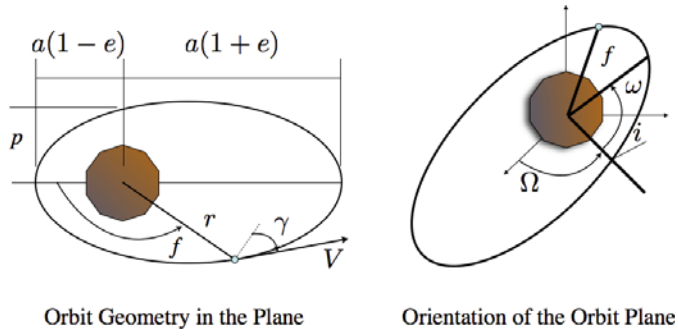


Fig. 3.1 Geometric definition of the orbit elements.

The eccentricity, e, equals $\sqrt{1 + 2H^2E/\mu^2}$. Due to the minimum energy value constraint $e \geq 0$, with $e = 0$ occurring at the minimum energy for a given angular momentum. Thus, the eccentricity is a measure of the excess energy that an orbit

has. If $E \geq 0$ then $e \geq 1$, which yields a fundamental change in the radius formula as now the denominator can disappear and, apparently $r \to \infty$. Denote the value H^2/μ as the orbit parameter, p, a sole function of the angular momentum of the orbit with units of length. With this new notation the trajectory solution for the two-body problem takes on its standard form:

$$r = \frac{p}{1 + e\cos f} \tag{3.36}$$

Using this relation it is now possible to relate the speed and flight path angle as functions of the true anomaly f. From the energy equation the speed can be evaluated, and from the angular momentum the flight-path angle can be evaluated

$$v = \sqrt{\frac{\mu}{p}\left(1 + 2e\cos f + e^2\right)} \tag{3.37}$$

$$\tan\gamma = \frac{e\sin f}{1 + e\cos f} \tag{3.38}$$

Again, these formula are valid for all energy and angular momentum. The only constraint which occurs is when $e \geq 1$, as then $1 + e\cos f = 0$ can occur at two particular values of true anomaly, $f = \pm f_\infty$ where $f_\infty = \arccos(-1/e)$, at which point the radius formula is undefined. Thus $f \in (-f_\infty, f_\infty)$. The additional quantities p, e and ω are all considered to be "classical" orbital elements in addition to the previously defined inclination i and longitude of the ascending node Ω. Note that another, alternate, orbit element is usually used known as the semi-major axis, a. This is a direct function of the energy only and can be computed as $a = -\frac{\mu}{2E}$, or alternately $a = p/(1 - e^2)$. Thus when $E < 0$, $a > 0$ and vice versa. When $E = 0$ the quantity a is undefined.

For an orbit with zero energy, a parabolic orbit, $f_\infty = \pi$. The radius, velocity and flight path angle all take on special forms, arising from trigonometric reductions of the formula:

$$r = \frac{p}{2\cos^2(f/2)} \tag{3.39}$$

$$v = 2\sqrt{\frac{\mu}{p}}\cos(f/2) \tag{3.40}$$

$$\tan\gamma = \tan(f/2) \tag{3.41}$$

In this case as the true anomaly approaches $\pm f_\infty$ the speed approaches zero and the flight path angle approaches $\pi/2$.

For an orbit with positive energy, a hyperbolic orbit, f_∞ lies in the interval $(\pi, \pi/2)$, with its value approaching π as $e \to 1$ and $\pi/2$ as $e \to \infty$. The latter situation occurs when $\mu \to 0$, which in fact mimics the situation when a spacecraft has a flyby of a small asteroid. While the form of the radius, speed and flight path angle formulae do not change for this case, their limiting values do. Now as

true anomaly approaches $\pm f_\infty$ the following limits occur, $v \to v_\infty$ and $\gamma \to \gamma_\infty$, where

$$v_\infty = \sqrt{\frac{\mu(e^2-1)}{p}} \tag{3.42}$$

$$= \sqrt{\frac{\mu}{-a}} \tag{3.43}$$

$$\gamma_\infty = \frac{\pi}{2} \tag{3.44}$$

For $E > 0$ recall that $e > 1$ and $a < 0$, thus yielding positive quantities within all of the relevant square roots.

With the results found to this point it is possible to define the three-dimensional motion of the two-body solution. Using the derived directions of $\hat{\boldsymbol{H}}$ and $\hat{\boldsymbol{e}}$ the vector $\hat{\boldsymbol{e}}_\perp = \hat{\boldsymbol{H}} \times \hat{\boldsymbol{e}}$ can be defined, and using $\hat{\boldsymbol{r}}$ can also define $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{H}} \times \hat{\boldsymbol{r}}$. Then, by applying these definitions the full position and velocity vector describing the orbit is:

$$\boldsymbol{r} = r\left[\cos f\hat{\boldsymbol{e}} + \sin f\hat{\boldsymbol{e}}_\perp\right] \tag{3.45}$$

$$\boldsymbol{v} = \frac{H}{r}\left[\hat{\boldsymbol{\theta}} + \tan\gamma\hat{\boldsymbol{r}}\right] \tag{3.46}$$

In these expressions r, v, γ, $\hat{\boldsymbol{r}}$, and $\hat{\boldsymbol{\theta}}$ all vary with true anomaly, with the appropriate limits for the cases of $E \geq 0$.

3.1.4 Motion in Time

What has been defined is the path, or trajectory, that a solution to the two-body problem takes in position and velocity space, also called phase space. The link between the true anomaly and time has not been developed yet. Also, the reader can note that we have only defined 5 independent integrals to this point. These are the angular momentum $\boldsymbol{H}$, energy E, and eccentricity vector $\boldsymbol{e}$, which have 5 independent quantities that can be expressed in terms of the classical orbital elements. These are the orientation of the orbit plane, inclination i and longitude of ascending node Ω, orientation of the trajectory in the orbit plane, the argument of periapsis ω, and the shape and size of the orbit, the eccentricity e and either the semi-major axis a or the orbit parameter p.

What is lacking is the integral that locates where the solution is located within the trajectory at a given time. This final integral always exists due to a theorem by Jacobi. This theorem states that a one-dimensional ordinary differential equation $\dot{x} = f(x)$ can always be solved by quadratures as

$$t - \tau_o = \int_{x_o}^{x} \frac{dx'}{f(x')} \tag{3.47}$$

with the epoch time τ_o serving as the final integral of motion. This result can either hold globally or locally between zeros of $f(x)$. In the two-body problem it is easiest to find it by noting a simple relationship between the time rate of change of the true anomaly and the angular momentum magnitude.

Recall the angular momentum magnitude formula, $H = rv\cos\gamma$, and its alternate statement $H = r^2\dot{\theta}$, where $\dot{\theta}$ is the time rate of change of the angular location of the solution. For the two-body problem this will equal the time rate of change of the true anomaly, $\dot{f}$. The quantity $r^2\dot{\theta}$ equals twice the rate at which the radius vector sweeps out area, which is the classical statement of Kepler's Second Law. Solving for $\dot{f}$ and evaluating r yields

$$\dot{f} = \sqrt{\frac{\mu}{p^3}}(1 + e\cos f)^2 \tag{3.48}$$

This differential equation can be solved by separation of variables, putting it into the form

$$\sqrt{\frac{\mu}{p^3}}(t - \tau_o) = \int_{f_o}^{f} \frac{df'}{(1 + e\cos f')^2} \tag{3.49}$$

We have the freedom to choose one of the initial conditions τ_o or f_o, and following convention choose $f_o = 0$, at periapsis, with τ_o then signifying the time at which the orbit was at periapsis, which then becomes the final integral of motion.

First consider the case $E < 0$ and $0 < e < 1$. The quadrature can be carried out for this case, yielding

$$\begin{aligned}\sqrt{\frac{\mu}{p^3}}\,(t - \tau_o) = {} & \frac{-e}{1 - e^2}\frac{\sin f'}{1 + e\cos f'}\bigg|_{f_o}^{f} \\ & + \frac{2}{(1 - e^2)^{3/2}}\arctan\left\{\sqrt{\frac{1 - e}{1 + e}}\tan\frac{1}{2}f'\right\}\bigg|_{f_o}^{f}\end{aligned} \tag{3.50}$$

Evaluating this integral from $f = 0$ to 2π defines the period of the orbit, T, which takes on a particularly simple form when stated in terms of the semi-major axis:

$$T = \frac{2\pi a^{3/2}}{\sqrt{\mu}} \tag{3.51}$$

and is Kepler's Third Law. Thus the period is only a function of the orbit energy. From this the mean motion can be defined as $n = \sqrt{\mu/a^3} = 2\pi/T$, as this is by definition the average rate at which the orbit is traversed.

The expression in Eq. 3.50 is not convenient to work with, leading to the definition of the eccentric anomaly, E. The arc-tangent function will define an angle, leading to the natural definition

$$\tan E/2 = \sqrt{\frac{1 - e}{1 + e}}\tan f/2 \tag{3.52}$$

Since this is a half-angle formula it yields angles resolved over the full interval $[0, 2\pi)$. Corresponding formulae, and their inverses, for $\cos E$ and $\sin E$ in terms of $\cos f$ and $\sin f$ are stated in Appendix A. Using these definitions gives us the standard form of Eq. 3.50, which is usually called Kepler's equation:

$$n(t - \tau_o) = E - e \sin E \tag{3.53}$$

The mean anomaly is defined as $M = nt$, an additional angle. Then the final integral of motion can also be restated as $M_o = n\tau_o$. This gives the usual definition of Kepler's equation

$$M - M_o = E - e \sin E \tag{3.54}$$

Given a true anomaly the mean anomaly can be directly computed, first transforming into the eccentric anomaly and then using Kepler's equation. Evaluation of the true anomaly given a mean anomaly is not as straightforward, as Kepler's equation must be solved to find the eccentric anomaly that corresponds to a given mean anomaly. Despite the simplicity of this equation, it cannot be solved in closed form given its transcendental nature. Significant research and analysis of this equation has occurred historically, and indeed still occurs today. See [26] for a detailed history of this equation and methods for its inversion.

Equivalent quadratures and definitions can be carried out for the case $e = 1$ and $e > 1$, for parabolic and hyperbolic orbits. These relations are needed when modeling the flyby of a small body relative to a planet or of a spacecraft relative to a small body. As mentioned above, the other trajectory relations still hold, the only modification needed is to Kepler's equation. Going back to the fundamental quadrature,

$$\sqrt{\frac{\mu}{p^3}}(t - \tau_o) = \int_0^f \frac{df}{(1 + e \cos f)^2}$$

when $e > 1$ the result is

$$\begin{aligned}\sqrt{\frac{\mu}{p^3}}(t - \tau_o) = & \left. \frac{e}{e^2 - 1} \frac{\sin f'}{1 + e \cos f'} \right|_{f_o}^{f} \\ & - \left. \frac{2}{(e^2 - 1)^{3/2}} \operatorname{arctanh} \left\{ \sqrt{\frac{e-1}{e+1}} \tan \frac{1}{2} f' \right\} \right|_{f_o}^{f}\end{aligned} \tag{3.55}$$

Following the same approach for the elliptic case, define the Hyperbolic Anomaly, F, as

$$\tanh F/2 = \sqrt{\frac{e-1}{e+1}} \tan f/2 \tag{3.56}$$

with similar relationships between $\cos f$ and $\sin f$ and $\cosh F$ and $\sinh F$. Applying these substitutions yields the hyperbolic version of Kepler's equation:

$$\sqrt{\frac{\mu}{|a|^3}}(t - \tau_o) = e \sinh F - F \tag{3.57}$$

For parabolic orbits, the quadrature takes on a simpler form and yields a direct relationship between time and true anomaly:

$$\sqrt{\frac{\mu}{p^3}}\,(t-\tau_o) = \frac{1}{2}\tan\left(\frac{f}{2}\right)\left[1+\frac{1}{3}\tan^2\left(\frac{f}{2}\right)\right] \tag{3.58}$$

where p is generally well defined (as it is computed from the angular momentum) even if $e = 1$.

Finally, note that for orbits with zero angular momentum, $p = 0$, the entire analysis must be redone given the singular definition of the true anomaly in this case. Such a derivation is not performed here, however, as it is not needed in the following. We only note that the eccentric anomaly remains well-defined and can be used to describe the geometry of motion in this case.

3.2 Rotational Dynamics of Small Bodies

For many applications it is of interest to have a compact, analytical description of rotational motion which closely matches the expected motion of a body rotating in space. For precision modeling and estimation the rotational dynamics will usually be modeled using numerical integration incorporating all relevant torques acting on the body [107]. However, the torques acting on a rotating body will often be small in magnitude, meaning that over reasonable lengths of time it may be feasible to use the Euler solution for the torque-free rotation of a rigid body to describe rigid body rotation. Such an analytical solution is of interest as it allows the body's orientation to be specified at arbitrary epochs without requiring numerical integration of Euler's equations from an initial epoch. In this section the classical solution for all possible cases of rotational motion is restated – generalizing the analysis given by MacMillan [94]. The solution is not derived in as much detail as for the two-body problem, as the remainder of the text only deals with orbital motion, and introducing a detailed discussion of the solution of the rotational equations of motion would take us a bit far afield. Closed-form solutions for the orientation of the rigid body are also provided in the form of Euler angles as a function of time. The solutions are stated in terms of elliptic integrals and functions, which are usually available from standard mathematical libraries. The following derivation avoids the use of theta functions to express the Euler angles of the body (which the classical solution relies on, [195]). The qualitative specification of a small body's rigid body rotation state is shown to be given by only two numbers once the moments of inertia are defined for the body.

3.2.1 The Inertia Dyadic

First define the mass moments and products of inertia of a general body. These arise from the fundamental definition of the rotational angular momentum of a body. Consider a collection of mass elements $dm(\boldsymbol{\rho})$ located by position vectors $\boldsymbol{\rho}$

and rigidly connected to each other. Assume these elements are all rotating about the center of mass of this rigid body collection with an instantaneous rotation rate $\boldsymbol{\omega}$. Then the velocity of each mass element equals $\boldsymbol{\omega} \times \boldsymbol{\rho} = \tilde{\boldsymbol{\omega}} \cdot \boldsymbol{\rho}$ and enables us to define the rotational angular momentum of the body as an integral over the mass distribution denoted by β

$$\boldsymbol{H} = \int_{\beta} \boldsymbol{\rho} \times \boldsymbol{v} \, dm \tag{3.59}$$

$$= -\int_{\beta} \tilde{\boldsymbol{\rho}} \cdot \tilde{\boldsymbol{\rho}} \, dm \cdot \boldsymbol{\omega} \tag{3.60}$$

where the angular velocity vector can be pulled out of the integral as it is constant throughout the body. The integral in front of the angular velocity is then defined as the inertia dyadic, or if resolved into a specific basis the inertia matrix where the diagonal elements are the moments of inertia and the off-diagonal elements are the products of inertia. The inertia dyadic $\boldsymbol{\mathcal{I}}$ can then be defined as

$$\boldsymbol{\mathcal{I}} = \int_{\beta} \left[\rho^2 \boldsymbol{U} - \boldsymbol{\rho}\boldsymbol{\rho}\right] dm(\boldsymbol{\rho}) \tag{3.61}$$

where $\boldsymbol{U}$ is the unity dyadic and the second term is a dyad. Using this notation the angular momentum vector of the body is then $\boldsymbol{H} = \boldsymbol{\mathcal{I}} \cdot \boldsymbol{\omega}$. A frame for the inertia dyadic has not been defined as of yet, but it is convenient to fix it in the body-frame, as then the integral quantities are all constant.

3.2.2 Orientation of the Rigid Body

A second crucial definition is the orientation of the rigid body relative to inertial space. The rotation dyadic (or equivalently matrix), denoted as $\boldsymbol{C}$, represents a linear operator that takes a vector from a body-fixed frame and expresses it in inertial space. The specification of the rotation dyadic as a function of time is the ultimate goal for describing the attitude motion of a body. To carry this out the relevant differential equations must be posed and integrated.

Consider the operation of $\boldsymbol{C}$ taking a vector specified in a body-fixed frame, $\boldsymbol{\rho}_B$, into a vector specified in an inertial frame, $\boldsymbol{\rho}_I = \boldsymbol{C} \cdot \boldsymbol{\rho}_B$. Taking the time derivative of this expression and recalling the transport theorem on time derivatives in a rotating frame, the equivalence $\boldsymbol{C} \cdot [\boldsymbol{\omega} \times \boldsymbol{\rho}_B + \dot{\boldsymbol{\rho}}_B] = \dot{\boldsymbol{C}} \cdot \boldsymbol{\rho}_B + \boldsymbol{C} \cdot \dot{\boldsymbol{\rho}}_B$ can be established. As this relation must hold for arbitrary vector $\boldsymbol{\rho}_B$, the relevant differential equation can be extracted

$$\dot{\boldsymbol{C}} = \boldsymbol{C} \cdot \tilde{\boldsymbol{\omega}} \tag{3.62}$$

3.2.3 Euler's Equations

Now consider the derivation of Euler's equations for the rigid body. It can be assumed that there are no external torques operating on the body, and thus the total rotational angular momentum must be conserved in inertial space. This conserved vector is specified, using the above notations and definitions, as

$$\boldsymbol{H} = \boldsymbol{C} \cdot \boldsymbol{\mathcal{I}} \cdot \boldsymbol{\omega} \tag{3.63}$$

Noting that $\dot{\boldsymbol{H}} = 0$ yields

$$0 = \dot{\boldsymbol{C}} \cdot \boldsymbol{\mathcal{I}} \cdot \boldsymbol{\omega} + \boldsymbol{C} \cdot \dot{\boldsymbol{\mathcal{I}}} \cdot \boldsymbol{\omega} + \boldsymbol{C} \cdot \boldsymbol{\mathcal{I}} \cdot \dot{\boldsymbol{\omega}} \tag{3.64}$$

but $\dot{\boldsymbol{\mathcal{I}}} = 0$ in a body-fixed frame, and substituting the differential equation for $\boldsymbol{C}$ yields $\boldsymbol{C} \cdot [\boldsymbol{\mathcal{I}} \cdot \dot{\boldsymbol{\omega}} + \tilde{\boldsymbol{\omega}} \cdot \boldsymbol{\mathcal{I}} \cdot \boldsymbol{\omega}] = 0$. This must hold for any rotation dyadic and thus defines the general statement of Euler's equations in the body-fixed frame

$$\boldsymbol{\mathcal{I}} \cdot \dot{\boldsymbol{\omega}} + \tilde{\boldsymbol{\omega}} \cdot \boldsymbol{\mathcal{I}} \cdot \boldsymbol{\omega} = 0 \tag{3.65}$$

3.2.4 Conserved Quantities

The fundamental conserved quantities that are defined for the torque-free rotation of a body are its rotational angular momentum and kinetic energy. The rotational kinetic energy is defined as

$$T = \frac{1}{2} \boldsymbol{\omega} \cdot \boldsymbol{\mathcal{I}} \cdot \boldsymbol{\omega} \tag{3.66}$$

Taking its time derivative yields $\dot{T} = \boldsymbol{\omega} \cdot \boldsymbol{\mathcal{I}} \cdot \dot{\boldsymbol{\omega}}$, and substituting Euler's equation from above immediately verifies that $\dot{T} \equiv 0$.

It is more instructive, for the moment, to consider the magnitude of the angular momentum vector, or $H^2 = \boldsymbol{H} \cdot \boldsymbol{H}$, which is naturally conserved, even in the body-fixed frame. Assume that the inertia dyadic is diagonal and is aligned along the principal moments of inertia with axes $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$ following the convention $I_x \leq I_y \leq I_z$. Then when written in expanded form both the kinetic energy and the angular momentum each constrain the angular velocity vector (in the body-fixed frame) to lie on an ellipsoid

$$H^2 = I_x^2 \omega_x^2 + I_y^2 \omega_y^2 + I_z^2 \omega_z^2 \tag{3.67}$$

$$2T = I_x \omega_x^2 + I_y \omega_y^2 + I_z \omega_z^2 \tag{3.68}$$

For these two ellipsoids to be consistent with each other, the following condition must hold

$$\boldsymbol{\omega} \cdot \boldsymbol{\mathcal{I}} \cdot \left[\boldsymbol{U} - \frac{H^2}{2T} \boldsymbol{\mathcal{I}}^{-1} \right] \cdot \boldsymbol{\mathcal{I}} \cdot \boldsymbol{\omega} = 0 \tag{3.69}$$

where the inertia dyadic always has an inverse for finite density distributions. To simplify the discussion, define the ratio $H^2/(2T)$ to be the "dynamic inertia" I_D, which turns out to be a fundamental quantity. Then for the equality to hold the dyadic $\boldsymbol{U} - I_D\boldsymbol{\mathcal{I}}^{-1}$ must be negative semi-definite. Since the inertia dyadic is diagonal, this reduces to the condition that the determinant be ≤ 0, which can be modified to

$$(I_x - I_D)(I_y - I_D)(I_z - I_D) \leq 0 \tag{3.70}$$

The condition is identically equal to zero if I_D equals any of the principal moments of inertia. Also if $I_D < I_x$ or $I_D > I_z$ then the inequality is violated. Thus the dynamic inertia must satisfy (with the usual order of moments of inertia)

$$I_x \leq I_D \leq I_z \tag{3.71}$$

From this fundamental inequality a restriction on the possible values of kinetic energy for a given angular momentum can be inferred.

$$\frac{1}{2}\frac{H^2}{I_x} \geq T \geq \frac{1}{2}\frac{H^2}{I_z} \tag{3.72}$$

The lower bound on T corresponds to the body rotating about its maximum moment of inertia, which is the minimum energy state and is the stable final rotational state in the presence of energy dissipation. Conversely, the upper bound corresponds to rotation about the minimum moment of inertia, which is the maximum energy state for a given angular momentum and which is unstable in the presence of energy dissipation.

To finish, define the effective spin rate of the body

$$\omega_l = \frac{2T}{\sqrt{H^2}} \tag{3.73}$$

The effective spin rate can, in combination with the dynamic inertia, be used as an alternate statement of energy and angular momentum

$$H = \sqrt{\boldsymbol{\omega}\cdot\boldsymbol{\mathcal{I}}\cdot\boldsymbol{\mathcal{I}}\cdot\boldsymbol{\omega}} = I_D\omega_l \tag{3.74}$$

$$T = \frac{1}{2}\boldsymbol{\omega}\cdot\boldsymbol{\mathcal{I}}\cdot\boldsymbol{\omega} = \frac{1}{2}I_D\omega_l^2 \tag{3.75}$$

Both of these parameters figure prominently in our analytic solution and represent the rotation rate and moment of inertia of a "sphere" with the given angular momentum and kinetic energy.

3.2.5 Problem Statement and Parameter Definitions

There are five special solution cases to be considered depending on the value of I_D relative to the principal moments of inertia.

1. If $I_D = I_z$ then the small body is in principal axis rotation about the z-axis, the largest moment of inertia, and is in its minimum energy state for a given angular momentum.
2. If $I_y < I_D < I_z$ then the small body has a non-zero nutation and precesses about the z-axis, commonly termed a short-axis rotation mode (SAM), as the maximum moment of inertia axis $\boldsymbol{z}$ is generally the shortest axis of a constant density body.
3. If $I_y = I_D$ then the small body is in principal axis rotation about the y-axis, or its angular velocity moves along a heteroclinic connection between rotations about the y-axis in opposite directions.
4. If $I_x < I_D < I_y$ then the small body has a non-zero nutation and precesses about the x-axis, commonly termed a long-axis rotation mode (LAM), as the minimum moment of inertia axis $\boldsymbol{x}$ is generally the longest axis of a constant density body.
5. If $I_x = I_D$ then the small body in in principal axis rotation about the x-axis, which is the smallest moment of inertia, and is in its maximum energy state for a given angular momentum.

The solution for cases 1 and 5 are trivial and just give uniform rotation about a principal axis of the body, that axis being fixed in inertial space. For cases 2 and 4 the motion becomes more complex and explicit solutions are needed and given here. For case 3 the uniform rotation case is also trivial, although the case for motion on the heteroclinic orbit has added complication. This solution is a special case of 2 or 4 and is discussed for completeness.

3.2.6 Angular Velocity in the Body-Fixed Frame

Solutions for Short-Axis Mode (SAM) and Long-Axis Mode (LAM)

If the dynamic inertia lies in the interval $I_y < I_D < I_z$ the body is in a short-axis rotation mode and the angular velocity vector will circulate about the z-axis of the body. This is thought to be the usual case for Solar System objects, as any body dissipating energy eventually enters this rotation mode on its way to uniform rotation about its maximum moment of inertia. If instead the dynamic inertia lies in the interval $I_x < I_D < I_y$ the angular velocity vector will circulate about the x-axis. Under energy dissipation this mode will asymptotically transition to a SAM mode of rotation and finally to uniform rotation about the z-axis [20]. Despite this, it is important to note that at least some asteroids reside in a LAM rotation state ([76], [63]).

The angular velocity vector of the body can be expressed in the principal axis, body-fixed frame in terms of elliptic functions:

$$\omega_x = \omega_l \sqrt{\frac{I_D\,(I_z - I_D)}{I_x\,(I_z - I_x)}} \sqrt{1 - \sigma^2\,\mathrm{sn}(\tau)^2} \tag{3.76}$$

$$\omega_y = \omega_l \sigma \sqrt{\frac{I_D\,(I_z - I_D)}{I_y\,(I_z - I_y)}}\,\mathrm{sn}(\tau) \tag{3.77}$$

$$\omega_z = \omega_l \sqrt{\frac{I_D\,(I_D - I_x)}{I_z\,(I_z - I_x)}} \sqrt{1 - \gamma^2\,\mathrm{sn}(\tau)^2} \tag{3.78}$$

where

$$\sigma = \begin{cases} 1 & I_y < I_D < I_z \\ k & I_x < I_D < I_y \end{cases} \tag{3.79}$$

$$\gamma = \begin{cases} k & I_y < I_D < I_z \\ 1 & I_x < I_D < I_y \end{cases} \tag{3.80}$$

$$k^2 = \begin{cases} \frac{(I_y - I_x)(I_z - I_D)}{(I_D - I_x)(I_z - I_y)} & I_y < I_D < I_z \\ \\ \frac{(I_D - I_x)(I_z - I_y)}{(I_y - I_x)(I_z - I_D)} & I_x < I_D < I_y \end{cases} \tag{3.81}$$

The parameter k is the usual "parameter" used in the computation and definition of the elliptic sine function sn. See Appendix F for a brief discussion and definition of elliptic functions and integrals. Note that $\mathrm{sn}(\tau)$ is periodic in τ with a period of $4K(k)$, where $K(k)$ is the complete elliptic integral of the first kind and k is its parameter. It should be noted that $k_{LAM} = 1/k_{SAM}$ and that for the general case $k < 1$.

The relation of the time parameter τ to the "real time" t is:

$$\tau - \tau_o = \sqrt{\frac{I_D\,(I_z - I_y)\,(I_D - I_x)}{I_x I_y I_z}}\,\frac{\omega_l\,(t - t_o)}{\sigma} \tag{3.82}$$

where t_o is the initial epoch. The period of the vector $\boldsymbol{\omega}$ in the body-fixed frame is thus

$$P = \sqrt{\frac{I_x I_y I_z}{I_D\,(I_z - I_y)\,(I_D - I_x)}}\,\frac{4\sigma}{\omega_l} K(k) \tag{3.83}$$

Motion along the Heteroclinic Connection ($I_D = I_y$)

For the special case of $I_D = I_y$ either formulation can be used to arrive at a solution. First note that $k = 1$ for this special case and that $\mathrm{sn}(\tau, k = 1) = \tanh(\tau)$, where tanh is the hyperbolic tangent function. The explicit solution for the angular rates become:

$$\omega_x = \omega_l \sqrt{\frac{I_y\,(I_z - I_y)}{I_x\,(I_z - I_x)}} \frac{1}{\cosh(\tau)} \tag{3.84}$$

$$\omega_y = \omega_l \tanh(\tau) \tag{3.85}$$

$$\omega_z = \omega_l \sqrt{\frac{I_y\,(I_y - I_x)}{I_z\,(I_z - I_x)}} \frac{1}{\cosh(\tau)} \tag{3.86}$$

$$\tau - \tau_o = \sqrt{\frac{(I_z - I_y)\,(I_y - I_x)}{I_x I_z}} \omega_l (t - t_o) \tag{3.87}$$

where cosh is the hyperbolic cosine function. The period of motion becomes infinite, of course. As $\tau \to \pm\infty$ the functions $\tanh(\tau) \to \pm 1$ and $1/\cosh(\tau) \to 0$ so that the motion reduces to a uniform rotation about the y-axis with rotation rate ω_l in either direction.

Principal Axis Rotations

Finally, note that substitution of $I_D = I_x$ or $I_D = I_z$ makes the corresponding k parameters equal to zero. Further, $\mathrm{sn}(\tau, k = 0) = \sin(\tau)$ and that $K(0) = \pi/2$. Thus when $I_D = I_z$ then $\omega_x = \omega_y = 0$ and $\omega_z = \omega_l$. It is interesting to note that the rotation period in the body frame approaches $\sqrt{\{[I_x I_y]/[(I_z - I_y)(I_z - I_x)]\}}\, 2\pi/\omega_l$, which is different than the inertial period of $2\pi/\omega_l$.

Analogously, when $I_D = I_x$, $\omega_y = \omega_z = 0$ and $\omega_x = \omega_l$. The rotation period in the body frame is again different than the inertial period.

Symmetric Moment of Inertia Rotations

If any two of the moments of inertia of the rotating body are equal a significant simplification occurs in the rotational dynamics. These cases are widely discussed in textbooks, and hence we will only state the simplified form of their rotational dynamics for this case. Two different cases can be distinguished, where the minimum moments of inertia are equal to each other, which corresponds to an oblate body, and where the two maximum moments of inertia are equal to each other, which is a prolate body.

Oblate Bodies Setting $I_x = I_y = I_m$ in the solution has a number of consequences. First, since $I_m \leq I_D \leq I_z$ the body is always in the SAM rotation mode. Note that $k = 0$, $\gamma = 0$ and $\sigma = 1$. Thus, the elliptic sine function degenerates to a circular sine function. Substituting these changes gives us the new simplified equations for rotation in the body-fixed frame

$$\omega_x = \omega_l \sqrt{\frac{I_D\,(I_z - I_D)}{I_m\,(I_z - I_m)}}\cos(\tau) \tag{3.88}$$

$$\omega_y = \omega_l \sigma \sqrt{\frac{I_D\,(I_z - I_D)}{I_m\,(I_z - I_m)}}\sin(\tau) \tag{3.89}$$

$$\omega_z = \omega_l \sqrt{\frac{I_D\,(I_D - I_m)}{I_z\,(I_z - I_m)}} \tag{3.90}$$

The spin rate is constant about the maximum moment of inertia and varies sinusoidally between the two minimum axes. It is important to note that the angular velocity is still moving in the body-fixed frame and remains periodic. The relationship between τ and time becomes

$$\tau - \tau_o = \sqrt{\frac{I_D\,(I_z - I_m)\,(I_D - I_m)}{I_z}}\frac{\omega_l\,(t - t_o)}{I_m} \tag{3.91}$$

The period of the angular velocity vector $\boldsymbol{\omega}$ in the body-fixed frame is now

$$P = I_m \sqrt{\frac{I_z}{I_D\,(I_z - I_m)\,(I_D - I_m)}}\frac{2\pi}{\omega_l} \tag{3.92}$$

Prolate Bodies Setting $I_y = I_z = I_M$ in the solution yields different changes. Now $I_x \leq I_D \leq I_M$ and the body is always in a LAM rotation mode. The parameter $k = 0$ again, but now $\gamma = 1$ and $\sigma = 0$. For the prolate case, care must be made in setting these moments of inertia equal to each other, as there are cancellations of zeros that occur. Substituting these into the angular velocity equations now yields

$$\omega_x = \omega_l \sqrt{\frac{I_D\,(I_M - I_D)}{I_x\,(I_M - I_x)}} \tag{3.93}$$

$$\omega_y = \omega_l \sqrt{\frac{I_D\,(I_D - I_x)}{I_M\,(I_M - I_x)}}\sin(\tau) \tag{3.94}$$

$$\omega_z = \omega_l \sqrt{\frac{I_D\,(I_D - I_x)}{I_M\,(I_M - I_x)}}\cos(\tau) \tag{3.95}$$

Similar dynamics occur, except now the minimum moment of inertia has a constant spin rate and the circulation occurs across the two maximum moments. The relation of the time parameter τ to t is

$$\tau - \tau_o = \sqrt{\frac{I_D\,(I_M - I_x)\,(I_M - I_D)}{I_x}}\frac{\omega_l\,(t - t_o)}{I_M} \tag{3.96}$$

The period of the vector $\boldsymbol{\omega}$ in the body-fixed frame is thus

$$P = I_M \sqrt{\frac{I_x}{I_D\,(I_M - I_x)\,(I_M - I_D)}}\,\frac{2\pi}{\omega_l} \tag{3.97}$$

3.2.7 Analytic Solution for Type-II Euler Angles

To complete the analytical specification of torque-free rigid body solution an expression must also be found for the rotation matrix. We choose to express this using Euler angles of the body as a function of time, which can be solved for completely in terms of elliptic functions and integrals. The specific realization of these angles as a function of time will depend on the set of Euler angles specified, in the following the standard Type-II Euler angles are used [55]. Starting from an inertial frame X–Y–Z the first angle is a ccw rotation about the Z-axis by an angle ϕ (the right ascension), followed by a ccw rotation about the X'-axis by an angle θ (the declination), followed by a ccw rotation about the Z''-axis by an angle ψ (the hour angle). The range of the angles are $\phi \in [0, 2\pi]$, $\theta \in [0, \pi]$, and $\psi \in [0, 2\pi)$. The rotation matrix from the inertial frame to the body-fixed frame, $\boldsymbol{C}^T$ in our notation, is specified as:

$$\boldsymbol{C}^T = \begin{bmatrix} \cos\phi\cos\psi & \sin\phi\cos\psi & \sin\theta\sin\psi \\ -\sin\phi\cos\theta\sin\psi & +\cos\phi\cos\theta\sin\psi & \\ -\cos\phi\sin\psi & -\sin\phi\sin\psi & \sin\theta\cos\psi \\ -\sin\phi\cos\theta\cos\psi & +\cos\phi\cos\theta\cos\psi & \\ \sin\phi\sin\theta & -\cos\phi\sin\theta & \cos\theta \end{bmatrix} \tag{3.98}$$

The rotation matrix from the body-fixed frame to the inertial frame is the transpose.

Solution for Euler Angles θ and ψ

First note that the hour angle and declination (Euler angles ψ and θ) can be solved using only algebraic manipulations. To see this choose the inertial coordinate frame orientation such that the inertial z-axis is aligned with the angular momentum vector $\boldsymbol{H}$ and note that $|\boldsymbol{H}| = I_D\omega_l$. Then, the angular momentum vector in the body-fixed frame is $\boldsymbol{H} = I_D\omega_l\,[\sin\theta\sin\psi; \sin\theta\cos\psi; \cos\theta]^T$, which must be equal to the vector $\boldsymbol{I}\cdot\boldsymbol{\omega} = [I_x\omega_x; I_y\omega_y; I_z\omega_z]^T$. Equate the two vectors to find the solutions:

$$\tan\psi = \frac{I_x\omega_x}{I_y\omega_y} \tag{3.99}$$

$$\cos\theta = \frac{I_z\omega_z}{I_D\omega_l} \tag{3.100}$$

The specific form of these relations will depend on the rotational mode the asteroid is in.

Evaluating the ratios for the case of SAM and LAM rotation yields:

$$\tan\psi = \frac{1}{\sigma}\sqrt{\frac{I_x\left(I_z - I_y\right)}{I_y\left(I_z - I_x\right)}}\frac{\sqrt{1-\sigma^2\mathrm{sn}^2(\tau)}}{\mathrm{sn}(\tau)} \tag{3.101}$$

$$\cos\theta = \sqrt{\frac{I_z\left(I_D - I_x\right)}{I_D\left(I_z - I_x\right)}}\sqrt{1-\gamma^2\mathrm{sn}^2(\tau)} \tag{3.102}$$

For a SAM rotation the angle ψ will circulate through all possible values as τ increases monotonically. For a LAM rotation this angle will instead librate about $\pm\pi/2$, never crossing the angles 0 or π. For the angle θ note that the quantity $\sqrt{\left\{[I_z(I_D - I_x)]/[I_D(I_z - I_x)]\right\}} \leq 1$ for $I_D \leq I_z$. Thus the relationship is always defined in terms of real arguments. For the case of SAM rotation the angle θ is constrained to lie in the interval $(0, \pi/2)$ while for the case of a LAM rotation the angle will instead lie in the interval $(0, \pi)$ and librate about $\pi/2$.

Solution for Euler Angle ϕ

The expression of the right ascension (Euler angle ϕ) is not as simple. Classical discussions usually give the closed form for this final angle in terms of ratios of theta functions [195], which are generally not available for numerical evaluation in standard packages. This final angle can be found in terms of an elliptic integral of the third kind, which has a much simpler form than the classical solutions and can be easily implemented on a computer.

First note the equation for $\dot{\phi}$ [55]:

$$\dot{\phi} = \frac{1}{\sin\theta}\left[\sin\psi\omega_x + \cos\psi\omega_y\right] \tag{3.103}$$

From Eqs. 3.99 and 3.100

$$I_D\omega_l \sin\theta\sin\psi = I_x\omega_x \tag{3.104}$$

$$I_D\omega_l \sin\theta\cos\psi = I_y\omega_y \tag{3.105}$$

Multiply Eqs. 3.104 and 3.105 by ω_x and ω_y respectively and add, and square Eqs. 3.104 and 3.105 and add. Then substitute the resultant quantities into Eq. 3.103 to find:

$$\dot{\phi} = I_D\omega_l\frac{I_x\omega_x^2 + I_y\omega_y^2}{I_x^2\omega_x^2 + I_y^2\omega_y^2} \tag{3.106}$$

Next, use relations $I_D^2\omega_l^2 = \omega_x^2 I_x^2 + \omega_y^2 I_y^2 + \omega_z^2 I_z^2$ and $I_D\omega_l^2 = \omega_x^2 I_x + \omega_y^2 I_y + \omega_z^2 I_z$ to simplify Eq. 3.106 further to find:

$$\dot{\phi} = \frac{I_D\omega_l}{I_z}\left[1 + \frac{I_D\left(I_z - I_D\right)\omega_l^2}{I_D^2\omega_l^2 - I_z^2\omega_z^2}\right] \tag{3.107}$$

where ω_z is the only function of time in the expression. Making the substitution for ω_z from Eq. 3.78 and carrying out a series of algebraic manipulations puts Eq. 3.107 into the form:

$$\dot{\phi} = \frac{I_D\omega_l}{I_z}\left[1 + \frac{I_z - I_x}{I_x}\frac{1}{1 + n\,\mathrm{sn}^2}\right] \tag{3.108}$$

$$n = \sigma^2\left(\frac{I_z}{I_x}\right)\left(\frac{I_y - I_x}{I_z - I_y}\right) \tag{3.109}$$

Equation 3.108 consists of a linear increase in time plus a quadrature involving an elliptic function. The quadrature can be stated as:

$$\begin{aligned}\phi - \phi_o = {} & \frac{I_D\omega_l}{I_z}(t - t_o) \\ & + \sigma\sqrt{\frac{I_D I_y}{I_z I_x}}\frac{\left(I_z - I_x\right)}{\sqrt{\left(I_z - I_y\right)\left(I_D - I_x\right)}}\int_{\tau_o}^{\tau}\frac{d\tau}{1 + n\,\mathrm{sn}^2(\tau)}\end{aligned} \tag{3.110}$$

The integral $\Pi(\tau, n) = \int_0^{\tau}\frac{d\tau}{1+n\,\mathrm{sn}^2(\tau)}$ is the elliptic integral of the third kind, for which robust evaluation routines exist [136]. A few brief properties should be noted here:

$$\Pi(\tau + 2K(k), n) = \Pi(\tau, n) + 2\Pi(K(k), n) \tag{3.111}$$

$$\Pi(\tau + K(k), n) = 2\Pi(K(k), n) - \Pi(K(k) - \tau, n) \tag{3.112}$$

Thus, to have the function $\Pi(\tau, n)$ defined unambiguously and to accumulate the appropriate increase in time, we use the following definition:

$$\begin{aligned}\bar{\Pi}(\tau, n, k) = {} & 2\Pi\left(K(k), n\right)\mathrm{int}\left[\tau/2K(k)\right] \\ & + \Pi\left(\tau - 2K(k)\,\mathrm{int}\left[\tau/2K(k)\right], n\right)\end{aligned} \tag{3.113}$$

The function int$[x]$ denotes the truncation of decimal points of the number x, defined for positive or negative values.

Thus, the right ascension is described as:

$$\begin{aligned}\phi = {} & \phi_o + \frac{I_D\omega_l}{I_z}(t - t_o) + \\ & \sigma\sqrt{\frac{I_D I_y}{I_z I_x}}\frac{\left(I_z - I_x\right)}{\sqrt{\left(I_z - I_y\right)\left(I_D - I_x\right)}}\left[\bar{\Pi}(\tau, n, k) - \bar{\Pi}(\tau_o, n, k)\right]\end{aligned} \tag{3.114}$$

Note that ϕ consists of two main terms, a linear drift in time described by the angular momentum and the I_z moment of inertia, and a more complex increasing term involving an elliptic integral. The linear drift in ϕ is significant as it is the rotation rate of an arbitrary body whose kinetic energy has dissipated to a minimum value while keeping its angular momentum constant [20].

It is of interest to compute the average rate of change of the right ascension. This can be computed by differencing the value of ϕ at τ_o and $\tau = \tau_o + 4K(k)$ and dividing by the corresponding period P. This yields the result:

$$\overline{\left(\frac{d\phi}{dt}\right)} = \frac{\omega_l I_D}{I_z}\left[1 + \frac{I_z - I_x}{I_x}\frac{\Pi(K(k), n)}{K(k)}\right] \tag{3.115}$$

The Euler angles ψ and θ are periodic with period P. The angle ϕ, on the other hand, will not (in general) increase by a multiple of 2π over this time period. Thus the inertial attitude of the body will be quasi-periodic in time.

Finally, in the limit as $I_D \to I_z$ the Euler angles approach a singular value for $\theta = 0$ with $\phi + \psi$ equalling the total rotation angle. Taking the appropriate limits and adding, both ϕ and ψ will contribute non-trivial computations to this sum that converge to the total rotation angle $\phi + \psi = \phi_o + \pi/2 + \omega_l(t - t_o)$. Alternate angle definitions can, of course, be used to avoid some of the singular aspects of the Euler angles as specific configurations are approached. It is important to note that non-singular algorithms can be robustly coded to handle all possible issues with the Euler angles.

3.2.8 Specification of the Rotational Dynamics

Finally, it is important to specify the rotational dynamics of a given body relative to an inertial frame. There are two ways to do this, either by specifying the qualitative nature of the dynamics or by specifying the initial conditions of the rotational dynamics. Both procedures are reviewed below.

Starting from the Qualitative Nature of the Rotational Dynamics

The qualitative nature of its motion is controlled by the effective rotation rate ω_l and its dynamic inertia I_D. By specifying these two quantities the qualitative properties of the body's rotation motion are at once known, and by using the relationships given herein the quantitative values for its motion are also specified. By specifying the initial epoch τ_o the location of the angular velocity vector in its periodic orbit is defined. Next the initial epoch of the motion t_o must be specified as well as the initial right ascension (ϕ_o). Once these are chosen the attitude of the body at the initial epoch is fixed relative to the initial angular momentum vector.

The above four constants specify the rotational motion of the body relative to its initial angular momentum vector. To completely specify its motion one needs only specify the orientation of the angular momentum vector in inertial space, which can be done with the choice of two angles. This defines an additional, constant,

transformation matrix from the angular momentum relative system to inertial space that must be applied whenever the attitude of the body is required in inertial space. These six initial conditions then completely specify the ideal rotational dynamics of a body, such as an asteroid or comet.

Specifying the rotational dynamics in this way has benefits for simulation, as the rotation mode for the asteroid is directly chosen as well as the direction about which the body will precess and nutate (i.e., the angular momentum vector of the body).

Specifying the Initial Conditions

Now consider the reverse problem, the specification of the rotational motion of the body given an initial attitude and body-fixed angular velocity. Assume the initial attitude of the body is specified as a transformation matrix from the body-fixed frame to the inertial frame, and is specified as $\boldsymbol{\mathcal{T}}_o$. The initial body-fixed angular velocity of the body is specified as $\boldsymbol{\omega}_o$.

Given the initial body-fixed angular velocity $\boldsymbol{\omega}_o$ and inertia tensor $\boldsymbol{\mathcal{I}}$ the constants ω_l and I_D are computed using Eqs. 3.74 and 3.75. Once this computation is performed the qualitative aspects of the rotational motion are known and all the relevant parameters of the motion can be computed, such as k, σ, and γ. The specification of the initial parameter τ_o must be made more carefully, however. First find the values of $\mathrm{sn}(\tau_o)$ and $\mathrm{cn}(\tau_o)$ from Eqs. 3.76–3.78.

$$\mathrm{sn}(\tau_o) = \frac{\omega_y(t_o)}{\sigma\omega_l}\sqrt{\frac{I_y(I_z - I_y)}{I_D(I_z - I_D)}} \tag{3.116}$$

where the value of σ depends on which mode the rotation state is in. The evaluation of the $\mathrm{cn}(\tau_o)$ term also depends on which rotation mode the asteroid is in:

$$\mathrm{cn}(\tau_o) = \begin{cases} \dfrac{\omega_x(t_o)}{\omega_l}\sqrt{\dfrac{I_x(I_z - I_D)}{I_D(I_z - I_D)}} & I_y < I_D \leq I_z \\ \\ \dfrac{\omega_z(t_o)}{\omega_l}\sqrt{\dfrac{I_z(I_z - I_x)}{I_D(I_D - I_x)}} & I_x \leq I_D < I_y \end{cases} \tag{3.117}$$

These elliptic functions must be inverted to compute τ_o.

To perform the inversion note the following definitions $\mathrm{sn}(\tau_o) = \sin\varphi_o$ and $\mathrm{cn}(\tau_o) = \cos\varphi_o$, and derive the amplitude angle $\varphi_o = \arctan\left(\mathrm{sn}(\tau_o)/\mathrm{cn}(\tau_o)\right)$, where this angle will lie in the interval $[0, 2\pi)$ in general. The quantities τ_o and φ_o are related via the elliptic integral of the first kind, as defined in Eq. F.2:

$$\tau_o = F(\varphi_o, k) \tag{3.118}$$

To properly carry out this computation requires the proper accounting for which quadrant the angle φ_o lies in

$$\tau_o = \begin{cases} F(\varphi_o, k) & \text{if } \varphi \leq \pi/2 \\ 2F(\pi/2, k) - F(\pi - \varphi_o, k) & \text{if } \pi/2 < \varphi \leq \pi \\ 2F(\pi/2, k) + F(\varphi_o - \pi, k) & \text{if } \pi < \varphi \leq 3\pi/2 \\ 4F(\pi/2, k) - F(2\pi - \varphi_o, k) & \text{if } 3\pi/2 < \varphi < 2\pi \end{cases} \tag{3.119}$$

where $F(\pi/2, k) = K(k)$ is the complete elliptic integral of the first kind. If the case $I_D = I_y$ should occur, then τ_o can be solved uniquely from the initial condition $\omega_y(t_o) = \omega_l \tanh(\tau_o)$ without using elliptic functions or integrals.

Having computed ω_l, I_D, and τ_o the Euler angles are defined and hence that the transformation matrix from the inertial frame with angular momentum along the z-axis to the body-fixed frame at the initial epoch, $\boldsymbol{C}_o^T$, is defined. The matrix $\boldsymbol{C}_o$ takes us from the body-frame to the inertial frame at the initial epoch. Note that, in general, there is a disconnect between the inertial frame defined by the $\boldsymbol{C}_o$ matrix and the initial inertial frame in which the body orientation is defined. To close the loop an additional transformation matrix must be applied to $\boldsymbol{C}_o$ to rotate from the inertial frame with angular momentum aligned with the z-axis to the inertial frame the body's attitude is specified in, specifically the transformation $\boldsymbol{T}$ such that:

$$\boldsymbol{T}\boldsymbol{C}_o = \boldsymbol{\mathcal{T}}_o \tag{3.120}$$

$$\boldsymbol{T} = \boldsymbol{\mathcal{T}}_o \boldsymbol{C}_o^T \tag{3.121}$$

Then the inertial attitude of the body is specified by $\boldsymbol{T} \cdot \boldsymbol{C}(t)$.

4. Equations of Motion for a Small Body Orbiter

This chapter considers the basic equations of motion for a spacecraft or particle relative to a small body. Several forms of the equations are derived, each with their own special applications.

4.1 Newtonian Equations of Motion

The most fundamental form for the equations of motion are the standard Newton's equations. In the following we assume that the particle has no effect on the motion of the central body. Assume a frame fixed to the central mass distribution, and relative to a fixed orientation with respect to inertial space. We also assume gravitational attraction by the Sun and other bodies, and solar radiation pressure from the Sun acting on the body. Combining these different accelerations together the standard equations of motion can be stated as

$$\ddot{\boldsymbol{r}} = \frac{\partial U(\boldsymbol{C}^T \cdot \boldsymbol{r})}{\partial \boldsymbol{r}} + \sum_{i=0}^{N} \frac{\partial \mathcal{R}_{p_i}(\boldsymbol{r}, \boldsymbol{d}_i)}{\partial \boldsymbol{r}} + \frac{\partial \mathcal{R}_{SRP}(\boldsymbol{r}, \boldsymbol{d}_0)}{\partial \boldsymbol{r}} \tag{4.1}$$

where $\boldsymbol{r}$ is the spacecraft position vector relative to the small body and referred to an inertial frame (i.e., the time derivatives are relative to an inertially oriented frame), U is the gravitational potential of the body discussed in Section 2.5, $\mathcal{R}_{p_i}$ represents the perturbation from the ith attracting body, and $\mathcal{R}_{SRP}$ represents the solar radiation pressure perturbation, both of these are defined in Section 2.6. The vector $\boldsymbol{d}_0$ represents the position of the small body relative to the Sun and the vectors $\boldsymbol{d}_i$ for $i \geq 1$ represent the position of additional perturbing bodies relative to the small body. The transformation dyadic $\boldsymbol{C}^T$ takes vectors from an inertial frame to the small body-fixed frame indicating that the gravitational potential of the central body is always stated in a body-fixed frame.

It is theoretically important, at this point, to establish that this full set of equations can be restated in a Lagrangian or Hamiltonian formulation. This allows us to invoke properties of these general classes of systems later in our discussions.

D.J. Scheeres, *Orbital Motion in Strongly Perturbed Environments*, Springer Praxis Books,
DOI 10.1007/978-3-642-03256-1_4, © Springer-Verlag Berlin Heidelberg 2012

4.2 Lagrangian Form of the Equations

Starting from the Newtonian frame, it is simple to write out the general Lagrangian. Once this is given, changes in coordinates and coordinate frames are simply made with the appropriate substitution into the Lagrangian. Denote the specific kinetic energy of the spacecraft relative to the asteroid-centered frame and the full potential:

$$T = \frac{1}{2}\dot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}} \tag{4.2}$$

$$\mathcal{U} = U + \sum_{i=0}^{N} \mathcal{R}_{p_i} + \mathcal{R}_{SRP} \tag{4.3}$$

The potential is ultimately a function of position and time, and thus does not lead to a time-invariant system in general.

The Lagrangian is then formed as $L = T + \mathcal{U}$, with a positive sign as the potential is used instead of the potential energy. The Lagrange equations of motion are then [55]:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{\partial L}{\partial q_i} \; ; \; i = 1, 2, \ldots, n \tag{4.4}$$

where for the original system the q_i and $\dot{q}_i$ are the components of the $\boldsymbol{r}$ and $\dot{\boldsymbol{r}}$ vectors. Once the Lagrangian is defined, it is relatively simple to rewrite the equations of motion in an alternate frame (such as a rotating frame) or in an alternate set of coordinates, so long as the specific transformation equations can be solved for the original position and velocity vectors in terms of the new coordinates and their time derivatives.

The equations of motion as stated above are convenient for direct numerical simulation provided that the relative ephemeris of all the bodies, $\boldsymbol{d}_i$, are known and the rotational dynamics of the central body is known and specified via the transformation dyadic $\boldsymbol{C}$. This form of the equations are not, however, particularly useful in terms of understanding the system analytically, nor in interpreting the output of any numerical simulation. To address these concerns a number of different reference frames are considered from which to observe the dynamics of the system or coordinates for reporting the motion. Which frame or coordinates are chosen depends on what aspect of the problem is of interest. The main frames of interest are a frame fixed to the central body and a frame fixed to the orbit of the small body about the Sun. The former frame is more useful for surface and close proximity motions. The latter is more useful for discussing orbital mechanics perturbed by solar radiation pressure and solar gravity. The specific methods used to study these frames are significantly different, and thus the following derivations are carried out independently of each other. For each frame it is still possible, and often relevant, to carry the full effect of the small body gravity or solar perturbations.

4.2.1 Body-Fixed Frame

The equations of motion can be restated under the assumption that the particle position vector is specified relative to the central body-fixed frame, and that this body is most generally in a state of complex rotational motion. In terms of the gravitational perturbations, this formulation actually simplifies the situation as it removes the transformation from body-fixed to inertial frame that otherwise must appear in the gravitational potential. It does not simplify the external perturbations, as there must still be a transformation that maps the external perturbations from their frame of definition into the body-fixed frame. If the body's rotation state is specified as a function of time, however, these just represent the time-varying transformation $\boldsymbol{C}(t)$ discussed in the previous chapter. To make this example more specific, imagine that the central body is in a state of torque-free tumbling. This is the situation for the asteroid Toutatis, studied in [76] and Chapter 8.

If the location of the particle is $\boldsymbol{r}$ relative to an inertial frame, the particle location with respect to the body-fixed frame is then $\boldsymbol{q} = \boldsymbol{C}(t)^T \cdot \boldsymbol{r}$ and the substitution $\boldsymbol{r} = \boldsymbol{C}(t) \cdot \boldsymbol{q}$ is made everywhere. The time rate of change of the position vector can also be computed as $\dot{\boldsymbol{r}} = \boldsymbol{C}(t) \cdot (\dot{\boldsymbol{q}} + \boldsymbol{\omega} \times \boldsymbol{q})$, where $\dot{\boldsymbol{q}}$ is taken with respect to the rotating frame and is expressed in that frame. These terms can be substituted into the Lagrangian to find:

$$L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) = \frac{1}{2} \left[\dot{\boldsymbol{q}} + \boldsymbol{\omega} \times \boldsymbol{q}\right] \cdot \left[\dot{\boldsymbol{q}} + \boldsymbol{\omega} \times \boldsymbol{q}\right] + \mathcal{U}\left(\boldsymbol{C}(t) \cdot \boldsymbol{q}\right) \tag{4.5}$$

$$\mathcal{U} = U(\boldsymbol{q}) + \sum_{i=0}^{N} \mathcal{R}_{p_i}(\boldsymbol{C}(t) \cdot \boldsymbol{q}, \boldsymbol{d}_i) + \mathcal{R}_{SRP}(\boldsymbol{C}(t) \cdot \boldsymbol{q}, \boldsymbol{d}_0) \tag{4.6}$$

Application of Lagrange's equations using $\boldsymbol{q}$ and $\dot{\boldsymbol{q}}$ as the independent coordinates and velocities then yields the correct form of the equations as stated in a rotating reference frame:

$$\ddot{\boldsymbol{q}} + \dot{\boldsymbol{\omega}} \times \boldsymbol{q} + 2\boldsymbol{\omega} \times \dot{\boldsymbol{q}} + \boldsymbol{\omega} \times \boldsymbol{\omega} \times \boldsymbol{q} = \frac{\partial \mathcal{U}}{\partial \boldsymbol{q}} \tag{4.7}$$

It is interesting to note that if the central body is in a torque-free rotation state, but not rotating about its principal axes of inertia, then the angular velocity is periodic in the body-fixed state. Thus, the equations of motion have time-varying, periodic coefficients represented by the angular velocity $\boldsymbol{\omega}$.

If the exogenous perturbations are neglected the potential $\mathcal{U} = U(\boldsymbol{q})$ and is independent of the transformation dyadic $\boldsymbol{C}$. If the body rotates about a principal moment of inertia, the angular velocities $\boldsymbol{\omega}$ are constant and the $\dot{\boldsymbol{\omega}}$ term disappears. For this case the system has an integral of motion, a discussion that is left for later.

4.2.2 Orbit-Fixed Frame

The orbit-fixed frame is defined as the frame that follows the heliocentric orbit position of the small body. In this frame the Sun remains fixed in an angular

sense, although its distance from the small body will in general vary periodically as will the frame rotation rate. It is physically distinct from the body-fixed frame, however, in that the angular velocity always has a fixed direction, normal to the orbital plane.

Now denote the orbit-fixed frame with a coordinate $\boldsymbol{Q} = \boldsymbol{C}_O(t) \cdot \boldsymbol{r}$. If the inertial z axis is aligned with the heliocentric orbit angular momentum and the x-axis lies along the Sun–body position vector, the transformation matrix takes a simple form:

$$\boldsymbol{C}_O = \begin{bmatrix} \cos(\omega_0 + f) & \sin(\omega_0 + f) & 0 \\ -\sin(\omega_0 + f) & \cos(\omega_0 + f) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.8}$$

where the small body heliocentric true anomaly f is an implicit function of time. The angular velocity is aligned along the z-axis and is:

$$\boldsymbol{\Omega} = \dot{f} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{4.9}$$

where $\dot{f} = \sqrt{\frac{\mu}{p_0^3}} \left(1 + e_0 \cos(f)\right)^2$.

The transformed Lagrangian in this frame is specified as:

$$L(\boldsymbol{Q}, \dot{\boldsymbol{Q}}, t) = \frac{1}{2} \left[\dot{\boldsymbol{Q}} + \dot{f}\hat{\boldsymbol{z}} \times \boldsymbol{Q}\right] \cdot \left[\dot{\boldsymbol{Q}} + \dot{f}\hat{\boldsymbol{z}} \times \boldsymbol{Q}\right] + \mathcal{U}\left(\boldsymbol{C}_O^T \cdot \boldsymbol{Q}\right) \tag{4.10}$$

Given this form of the angular velocity, the general equations in this frame are:

$$\ddot{\boldsymbol{Q}} + \ddot{f}\hat{\boldsymbol{z}} \times \boldsymbol{Q} + 2\dot{f}\hat{\boldsymbol{z}} \times \dot{\boldsymbol{Q}} + \dot{f}^2 \hat{\boldsymbol{z}} \times \hat{\boldsymbol{z}} \times \boldsymbol{Q} = \frac{\partial \mathcal{U}}{\partial \boldsymbol{Q}} \tag{4.11}$$

Neglecting the mass distribution of the central body, i.e., by choosing $U = \mu/|\boldsymbol{Q}|$ and neglecting non-solar perturbations the equations of motion are periodic in general with period $T = 2\pi a_0^{3/2}/\sqrt{\mu_0}$. This period is very long with respect to the usual orbit period about the central body, and only becomes significant for motion far from the central body. This is consistent with the simplifying assumption where the central body mass distribution in neglected and the central body is treated as a point mass. This assumption is made in the following as it provides a good motivation to derive a significantly different set of equations for motion in this regime.

As $\dot{f}$ is a function of f, one is tempted to make the true anomaly the independent parameter for these equations. Performing this transformation alone does not yield any significant simplifications. Rather, following tradition in the elliptic restricted three-body problem [97] it is easier to simultaneously scale the vector $\boldsymbol{Q}$ by the Sun–body distance $d = p_0/(1 + e_0 \cos f)$ and change the independent parameter from time to true anomaly. Then, the transformation sequence can be stated as $\boldsymbol{Q} \to d\boldsymbol{q}$, and the time transformation can be stated as $\dot{\boldsymbol{Q}} = d\boldsymbol{Q}/dt = d\boldsymbol{Q}/df\dot{f} = \boldsymbol{Q}'\dot{f}$. This leads to the following sequence (note, do not confuse the $d\boldsymbol{q}$ notation with a

differential, it is just a factor d multiplied by $\boldsymbol{q}$):

$$\boldsymbol{Q} = d\boldsymbol{q} \tag{4.12}$$

$$\dot{\boldsymbol{Q}} = d'\boldsymbol{q}\dot{f} + d\boldsymbol{q}'\dot{f} \tag{4.13}$$

$$= \sqrt{\frac{\mu_0}{p_0}}\left[e_0 \sin f\boldsymbol{q} + (1 + e_0 \cos f)\boldsymbol{q}'\right] \tag{4.14}$$

$$\ddot{\boldsymbol{Q}} = \sqrt{\frac{\mu_0}{p_0}}\left[e_0 \cos f\boldsymbol{q} + e_0 \sin f\boldsymbol{q}' - e_0 \sin f\boldsymbol{q}' + (1 + e_0 \cos f)\boldsymbol{q}''\right]\dot{f} \tag{4.15}$$

$$= \frac{\mu_0}{p_0^2}(1 + e_0 \cos f)^3\left[\boldsymbol{q}'' + \frac{e_0 \cos f}{1 + e_0 \cos f}\boldsymbol{q}\right] \tag{4.16}$$

Substituting these into Eq. 4.11, a common factor can be extracted

$$\boldsymbol{q}'' + 2\hat{\boldsymbol{z}} \times \boldsymbol{q}' + \frac{e_0 \cos f}{1 + e_0 \cos f}\boldsymbol{q} + \hat{\boldsymbol{z}} \times \hat{\boldsymbol{z}} \times \boldsymbol{q} = \frac{p_0}{\mu_0}\frac{1}{(1 + e_0 \cos f)^2}\frac{\partial \mathcal{U}}{\partial \boldsymbol{q}} \tag{4.17}$$

Now consider the potential and its specification in this particular frame (i.e., just a point mass plus the solar perturbation). In terms of the original variables

$$\mathcal{U} = \frac{\mu}{|\boldsymbol{Q}|} + \mu_0\left[\frac{1}{|\boldsymbol{d} - \boldsymbol{Q}|} - \frac{\boldsymbol{d} \cdot \boldsymbol{Q}}{|\boldsymbol{d}|^3}\right] - \boldsymbol{Q} \cdot \frac{P_0\boldsymbol{d}}{Bd^3} \tag{4.18}$$

where the simple lumped solar radiation pressure model is assumed and the vector $\boldsymbol{d}$ points towards the small body from the Sun. Implementing the new transformation changes the potential to

$$\mathcal{U} = \frac{\mu}{d|\boldsymbol{q}|} + \frac{\mu_0}{d}\left[\frac{1}{|\hat{\boldsymbol{d}} - \boldsymbol{q}|} - \hat{\boldsymbol{d}} \cdot \boldsymbol{q}\right] - \frac{1}{d}\boldsymbol{q} \cdot \frac{P_0}{B}\hat{\boldsymbol{d}} \tag{4.19}$$

Defining $\mathcal{V} = \frac{p_0}{\mu}\frac{1}{(1+e_0 \cos f)^2}\mathcal{U}$,

$$\mathcal{V} = \frac{1}{1 + e_0 \cos f}\left[\frac{\mu/\mu_0}{|\boldsymbol{q}|} + \frac{1}{|\hat{\boldsymbol{d}} - \boldsymbol{q}|} - \hat{\boldsymbol{d}} \cdot \boldsymbol{q} - \frac{P_0/\mu_S}{B}\boldsymbol{q} \cdot \hat{\boldsymbol{d}}\right] \tag{4.20}$$

and the associated equations of motion:

$$\boldsymbol{q}'' + 2\hat{\boldsymbol{z}} \times \boldsymbol{q}' + \frac{e_0 \cos f}{1 + e_0 \cos f}\boldsymbol{q} + \hat{\boldsymbol{z}} \times \hat{\boldsymbol{z}} \times \boldsymbol{q} = \frac{\partial \mathcal{V}}{\partial \boldsymbol{q}} \tag{4.21}$$

The form of these equations is somewhat simpler than the original statement, and the periodic terms are explicitly identified. Careful attention must be given to the interpretation to the solution of these equations, as they are relative to a length scale that changes in time. Specifically, when these solutions are to be transformed into an inertial frame or into a set of osculating orbit elements, the transformations in Eqs. 4.12 and 4.14 must first be applied. Additionally, as the independent parameter is true anomaly, to switch to the time domain one must

make the transformations from true anomaly through Kepler's equation, to time. In this direction there is no need to solve Kepler's equation iteratively.

This statement of the problem has no approximations other than the neglect of the central body mass distribution and the simplified solar radiation pressure term. Thus it is similar to the elliptic restricted three-body problem, except the coordinate center is not at the center of mass. It is useful to introduce an additional approximation to this set of equations, assuming that the particle is close to the central body, or that $|\boldsymbol{q}| \ll 1$ and expanding the term $1/|\hat{\boldsymbol{d}} - \boldsymbol{q}|$ up to the second order. This is essentially the "Hill approximation" and was derived in Chapter 2. The potential $\mathcal{V}$ can then be expressed as

$$\mathcal{V} = \frac{1}{1 + e_0 \cos f} \left[\frac{\mu/\mu_0}{|\boldsymbol{q}|} + \frac{3}{2} \left(\hat{\boldsymbol{d}} \cdot \boldsymbol{q} \right)^2 - \frac{1}{2} \boldsymbol{q} \cdot \boldsymbol{q} - \frac{P_0/\mu_0}{B} \boldsymbol{q} \cdot \hat{\boldsymbol{d}} + \ldots \right] \tag{4.22}$$

To isolate the time periodic terms in the potential, extract the $1/2q^2$ for reasons that will be evident in a moment and truncate the higher-order terms. The resulting potential is $\mathcal{V}_H$, where the H stands for the Hill approximation. Evaluating the equations of motion then yield:

$$\boldsymbol{q}'' + 2\hat{\boldsymbol{z}} \times \boldsymbol{q}' + \frac{e_0 \cos f}{1 + e_0 \cos f} \boldsymbol{q} + \hat{\boldsymbol{z}} \times \hat{\boldsymbol{z}} \times \boldsymbol{q} = -\frac{1}{1 + e_0 \cos f} \boldsymbol{q} + \frac{\partial \mathcal{V}_H}{\partial \boldsymbol{q}} \tag{4.23}$$

and simplified by moving the exposed term from the right to the left:

$$\boldsymbol{q}'' + 2\hat{\boldsymbol{z}} \times \boldsymbol{q}' + \boldsymbol{q} + \hat{\boldsymbol{z}} \times \hat{\boldsymbol{z}} \times \boldsymbol{q} = \frac{\partial \mathcal{V}_H}{\partial \boldsymbol{q}} \tag{4.24}$$

Finally, noting the identity $\tilde{\hat{\boldsymbol{z}}} \cdot \tilde{\hat{\boldsymbol{z}}} = \hat{\boldsymbol{z}}\hat{\boldsymbol{z}} - \boldsymbol{U}$ leads to the final form of the equations

$$\boldsymbol{q}'' + 2\hat{\boldsymbol{z}} \times \boldsymbol{q}' + (\hat{\boldsymbol{z}} \cdot \boldsymbol{q})\hat{\boldsymbol{z}} = \frac{\partial \mathcal{V}_H}{\partial \boldsymbol{q}} \tag{4.25}$$

$$\mathcal{V}_H = \frac{1}{1 + e_0 \cos f} \left[\frac{\mu/\mu_0}{|\boldsymbol{q}|} + \frac{3}{2} \left(\hat{\boldsymbol{d}} \cdot \boldsymbol{q} \right)^2 - \frac{P_0/\mu_0}{B} \boldsymbol{q} \cdot \hat{\boldsymbol{d}} \right] \tag{4.26}$$

Thus, the periodic portion is isolated to one term on the right-hand side.

In the Hill three-body problem, it is customary to make one additional normalization to remove all parameters from the problem. This is to scale the length by the quantity $(\mu/\mu_0)^{1/3}$. Doing so does not change the form of the left-hand side of the equations, and thus can be modeled completely by defining a new potential that has been so normalized. Let us call this potential $\tilde{\mathcal{V}}_H$

$$\tilde{\mathcal{V}}_H = \frac{1}{1 + e_0 \cos f} \left[\frac{1}{|\boldsymbol{q}|} + \frac{3}{2} \left(\hat{\boldsymbol{d}} \cdot \boldsymbol{q} \right)^2 - \tilde{\beta} \boldsymbol{q} \cdot \hat{\boldsymbol{d}} \right] \tag{4.27}$$

$$\tilde{\beta} = \frac{P_0}{\mu_0 B} \left(\frac{\mu}{\mu_0} \right)^{-1/3} \tag{4.28}$$

This basic form of the equations of motion was originally derived in [171].

4.3 Hamiltonian Form of the Equations

The Lagrangian frame allows for simple restatement of the equations of motion in different coordinate frames and coordinate systems, so long as the dynamical system is specified by a set of coordinates and their time derivatives. For physical intuition and spatial geometry, this works quite well and almost all of the fundamental properties of astrodynamical systems can be expressed using a Lagrangian set of the equations of motion. There are a number of deeper properties that exist for these systems, however, that are not as easily realized in the Lagrangian format. The first step away from these is via Hamilton's equations, which replaces the idea of a coordinate and a velocity with a set of conjugate quantities that are related to each other as coordinates and their corresponding momenta.

Given a set of equations in a Lagrangian frame, it is rather simple to transform them into a Hamiltonian form. One can leave the coordinates as-is and only operate on the momenta and velocities if one wishes to make the transformation in a minimum number of steps. An efficient way to perform this transformation is via the Legendre Transformation.

Assume that a well defined Lagrangian $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$ exists from which the equations of motion can be derived using Lagrange's equations. Define the momentum, p_i, conjugate to the coordinate q_i as $p_i = \partial L / \partial \dot{q}_i$, or in vector form:

$$\boldsymbol{p} = \frac{\partial L}{\partial \dot{\boldsymbol{q}}} \tag{4.29}$$

Recalling the standard form of Lagrange's equations this becomes

$$\dot{\boldsymbol{p}} = \frac{\partial L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)}{\partial \boldsymbol{q}} \tag{4.30}$$

This does not simplify matters, however, as we wish to completely eliminate $\dot{\boldsymbol{q}}$ from the Lagrangian to find a statement of the equations of motion independent of the velocities. In principle this can be done if Eq. 4.29 can be solved for the velocity, or $\dot{\boldsymbol{q}} = f(\boldsymbol{q}, \boldsymbol{p})$. A technical condition for this is that $|\partial^2 L / \partial \dot{\boldsymbol{q}}^2| \neq 0$, which is generally satisfied for the systems considered here.

The Legendre Transformation naturally handles such substitutions and indicates the functional form of the desired expression $\dot{\boldsymbol{q}} = f(\boldsymbol{q}, \boldsymbol{p})$. To carry this out form the new functional expression:

$$H(\boldsymbol{q}, \boldsymbol{p}, t) = \boldsymbol{p} \cdot \dot{\boldsymbol{q}} - L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) \tag{4.31}$$

where the $\dot{\boldsymbol{q}}$ expressions in the above are functions of $\boldsymbol{q}$ and $\boldsymbol{p}$. Now consider variations in the new variable $\boldsymbol{p}$ while keeping $\boldsymbol{q}$ fixed:

$$H_{\boldsymbol{p}} \cdot \delta \boldsymbol{p} = \dot{\boldsymbol{q}} \cdot \delta \boldsymbol{p} + \boldsymbol{p} \cdot \delta \dot{\boldsymbol{q}} - L_{\dot{\boldsymbol{q}}} \cdot \delta \dot{\boldsymbol{q}} \tag{4.32}$$

where $\delta \dot{\boldsymbol{q}}$ can be further expressed as a variation with respect to $\boldsymbol{p}$. This is not needed, however, as the two terms dotted with $\delta \dot{\boldsymbol{q}}$ are equal and opposite (Eq. 4.29)

and cancel. This leaves the defining relationship for the transformation:

$$\dot{\boldsymbol{q}} = H_{\boldsymbol{p}} \tag{4.33}$$

Using a similar variation technique it can be shown that $H_{\boldsymbol{q}} = -L_{\boldsymbol{q}}$, leading to the complete set of equations of motion in terms of the coordinates and momenta:

$$\dot{\boldsymbol{p}} = -H_{\boldsymbol{q}} \tag{4.34}$$

Collecting the coordinates and momenta into a single state $\boldsymbol{x} = [\boldsymbol{q}; \boldsymbol{p}] \in \mathbf{R}^{2n}$ with the attendant equations of motion defines the state space form of the equations of motion

$$\dot{\boldsymbol{x}} = JH_{\boldsymbol{x}} \tag{4.35}$$

where J is defined as:

$$J = \begin{bmatrix} 0_n & I_n \\ -I_n & 0_n \end{bmatrix} \tag{4.36}$$

and 0_n is an $n \times n$ matrix of zeros and I_n is the $n \times n$ identity matrix. The matrix J plays a significant role in Hamiltonian dynamics.

A reapplication of the Legendre Transformation with the defining quantity $\dot{\boldsymbol{q}} = H_{\boldsymbol{p}}$ and assuming that $|H_{\boldsymbol{pp}}| \neq 0$ will yield the Lagrangian equations of motion. Thus the Hamiltonian and Lagrangian expressions of the equations of motion are fundamentally linked to each other. While the Legendre Transformation tells us what the new functional for deriving the equations of motion is and indicates the relationship between the parameters, one must still solve for the $\dot{\boldsymbol{q}}$ in terms of the $\boldsymbol{p}$ and make the substitutions.

The importance of the Hamiltonian form of the equations of motion is that they allow the structure of the equations of motion to be preserved through a systematic set of transformations between different sets of variables called "canonical variables." A simple definition of a canonical transformation of a Hamiltonian dynamical system is one that preserves the same structure of the equations of motion. Specifically, a transformation between old variables $\boldsymbol{q}$ and $\boldsymbol{p}$ and new variables $\boldsymbol{Q}(\boldsymbol{q}, \boldsymbol{p})$ and $\boldsymbol{P}(\boldsymbol{q}, \boldsymbol{p})$ is a canonical transformation if the transformation equations are unique and invertible and the resulting equations of motion are:

$$\dot{\boldsymbol{Q}} = \frac{\partial K(\boldsymbol{Q}, \boldsymbol{P})}{\partial \boldsymbol{P}} \tag{4.37}$$

$$\dot{\boldsymbol{P}} = -\frac{\partial K(\boldsymbol{Q}, \boldsymbol{P})}{\partial \boldsymbol{Q}} \tag{4.38}$$

where $K(\boldsymbol{Q}, \boldsymbol{P}) = H(\boldsymbol{q}(\boldsymbol{Q}, \boldsymbol{P}), \boldsymbol{p}(\boldsymbol{Q}, \boldsymbol{P}))$.

To show that a proposed transformation is canonical it is not necessary to completely reconstruct the equations of motion. Rather, there are a number of tests that can be carried out on the transformation to ascertain whether it is canonical. A particularly useful test is that if the Jacobian of the transformation is a sym-

plectic matrix, then the transformation is canonical. This test is used exclusively in this text, as it has a deeper meaning when discussing the dynamical evolution of a Hamiltonian system. The Jacobian of an arbitrary transformation is defined as:

$$M = \frac{\partial(\boldsymbol{Q}, \boldsymbol{P})}{\partial(\boldsymbol{q}, \boldsymbol{p})} \tag{4.39}$$

and is a matrix of dimension $2n \times 2n$.

For a matrix M to be symplectic it must satisfy two simple rules. It must be a square matrix of even dimension and it must satisfy the matrix identity:

$$J = M^T J M \tag{4.40}$$

The matrix J satisfies this equation and is thus symplectic itself.

Symplectic matrices form a group and have distinctive properties. These are discussed in many texts, but a particularly clear statement of their properties is found in [132], which are summarized in Appendix D. One important property is that the determinant of any symplectic matrix always equals 1, and thus a symplectic matrix is always non-singular and invertible. Next is that the inverse of a symplectic matrix can always be stated in closed form as

$$M^{-1} = -J M^T J \tag{4.41}$$

Both of these results can be directly proven from the definition of a symplectic matrix, as discussed in the appendix.

Beyond the derivation of the equations of motion in some standard forms, the main application of canonical transformations used here is to note the properties of solution flows. Specifically, the solution flow of a Hamiltonian dynamical system is, itself, a canonical transformation from the state at a specified epoch (i.e., the initial conditions) to the state at a different time. To establish this consider a Hamiltonian dynamical system evolved forward in time by a small time interval, Δt, using a Taylor series expansion in time to evolve the flow forward [56]

$$\boldsymbol{x}(t + \Delta t) = \boldsymbol{x}(t) + \dot{\boldsymbol{x}}(t)\Delta t + \frac{1}{2!}\ddot{\boldsymbol{x}}(t)\Delta t^2 + \ldots \tag{4.42}$$

For small Δt we only consider the linear terms in Δt, as the limit $\Delta t \to 0$ will be taken in a moment. Substituting the equations of motion then yields the mapping between the two states:

$$\boldsymbol{x}(t + \Delta t) = \boldsymbol{x}(t) + J H_{\boldsymbol{x}} \Delta t \tag{4.43}$$

Taking the partial of $\boldsymbol{x}(t + \Delta t)$ with respect to $\boldsymbol{x}$, denoted as Φ, yields

$$\Phi = \frac{\partial \boldsymbol{x}(t + \Delta t)}{\partial \boldsymbol{x}(t)} \tag{4.44}$$

$$= I + J H_{\boldsymbol{x}\boldsymbol{x}} \Delta t \tag{4.45}$$

Now the symplectic test is applied to Φ,

$$\Phi^T J \Phi = [I - H_{\boldsymbol{xx}} J \Delta t]\, J\, [I + J H_{\boldsymbol{xx}} \Delta t] \tag{4.46}$$
$$= [I - H_{\boldsymbol{xx}} J \Delta t]\, [J - H_{\boldsymbol{xx}} \Delta t] \tag{4.47}$$
$$= J + H_{\boldsymbol{xx}} \Delta t - H_{\boldsymbol{xx}} \Delta t + \mathcal{O}(\Delta t^2) \tag{4.48}$$
$$= J \tag{4.49}$$

The derivation is only for an infinitesimal Δt; however, canonical transformations form a group, meaning that any combination of canonical transformations in sequence results in a canonical transformation. Thus, this result can be extended to an arbitrary time interval by successive application of the transformation, meaning that the solution flow, denoted as $\boldsymbol{x}(t) = \phi(t; \boldsymbol{x}_o, t_o)$, defines a canonical transformation from $\boldsymbol{x}_o$ to $\boldsymbol{x}$.

Define the Jacobian of the solution flow as the matrix Φ, commonly termed the state transition matrix

$$\Phi(t, t_o, \boldsymbol{x}_o) = \frac{\partial \phi(t; \boldsymbol{x}_o, t_o)}{\partial \boldsymbol{x}_o} \tag{4.50}$$

Then this matrix is a symplectic matrix, a fact that becomes important later.

4.4 Lagrange Planetary and Gauss Equations

Derivation of equations of motion can be generalized beyond the structured Lagrangian and Hamiltonian approaches. In the following two very important alternate statements of the equations of motion for a particle are derived, the Lagrange Planetary Equations and the Gauss Equations. The heart of these equations of motion is the identification of a problem that can be fully integrated in terms of constants of motion, in this case the two-body problem with no additional perturbations. The inclusion of additional perturbation forces generates changes to the constants of motion, causing them to become time-varying coordinates. However, the original transformation between the constants and the solution still stands and allows the time-varying coordinates to describe the solution. The method by which such equations can be formally developed is called Variation of Parameters. Two simple derivations are provided, relying on Brouwer and Clemence for a detailed discussion of the derivations [19].

We take two different approaches to the Lagrange Planetary and Gauss equations, as each derivation and set of equations has important distinctions between them. The Lagrange Planetary equations are derived assuming that the perturbation functions arise from a potential. The Gauss equations, on the other hand, are derived assuming an arbitrary acceleration is applied to the particle. It is possible to relate these two assumptions to each other; however, for completeness independent derivations will be given for both.

4.4.1 Lagrange Planetary Equations

Consider the dynamical system defined by a position vector $\boldsymbol{r}$ and a velocity vector $\boldsymbol{v} = \dot{\boldsymbol{r}}$. Assume the system is governed by a set of equations of motion that can be decomposed as:

$$\ddot{\boldsymbol{r}} = \frac{\partial U}{\partial \boldsymbol{r}} + \frac{\partial R}{\partial \boldsymbol{r}} \tag{4.51}$$

where $U(\boldsymbol{r})$ and $R(\boldsymbol{r}, t)$. The potential function U is the main potential and R is the perturbing potential. The basic assumption is that the problem can be integrated when $R \equiv 0$ and expressed in terms of constants of motion, denoted here as $\boldsymbol{\alpha} \in \mathbf{R}^6$. Specifically, the solution of the dynamical system can be written as $\boldsymbol{r}(t) = \boldsymbol{r}(t, \boldsymbol{\alpha})$ and $\boldsymbol{v}(t) = \boldsymbol{v}(t, \boldsymbol{\alpha})$. Note that $\boldsymbol{v} = \partial \boldsymbol{r}/\partial t$ and that $\partial \boldsymbol{v}/\partial t = \partial U/\partial \boldsymbol{r}$. For example, if the main potential is $U = \mu/r$, then the $\boldsymbol{\alpha}$ could be the classical orbital elements.

The main idea behind the Variation of Constants approach is to let the constants of integration become the new variables of the dynamical system. This is a well-posed transformation so long as the $\boldsymbol{\alpha}$ comprise a full, linearly independent solution of the system, i.e., that the Jacobian $\partial \boldsymbol{\alpha}/\partial(\boldsymbol{r}, \boldsymbol{v})$ is non-singular. Then, under the perturbation term R, which need not be small, the "constants" of motion become functions of time and are no longer constant. A future state of the system $\boldsymbol{\alpha}^*$ can always be transformed to the position and velocity of the point at a new time through the transformations $\boldsymbol{r}(t) = \boldsymbol{r}(t; \boldsymbol{\alpha}^*)$ and $\boldsymbol{v}(t) = \boldsymbol{v}(t; \boldsymbol{\alpha}^*)$. The $\boldsymbol{\alpha}$ are sometimes called "osculating orbit elements" as they are the set of elements that to describe the true trajectory state in phase space at each point in time.

To derive the equations, first take the full time derivative of the position vector:

$$\dot{\boldsymbol{r}} = \frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\alpha}} \dot{\boldsymbol{\alpha}} + \frac{\partial \boldsymbol{r}}{\partial t} \tag{4.52}$$

Now note that $\dot{\boldsymbol{r}} = \boldsymbol{v}$ by definition and that $\partial \boldsymbol{r}/\partial t = \boldsymbol{v}$ by choice. Efroimsky [33] has noted that this choice is not necessary, and that by generalizing this relation one can introduce a so-called "gauge function" into the theory. In the current description the more conventional derivation is presented that assumes $\partial \boldsymbol{r}/\partial t = \boldsymbol{v}$. Due to this, the first relation for $\dot{\boldsymbol{\alpha}}$ is

$$\frac{\partial \boldsymbol{r}}{\partial \boldsymbol{\alpha}} \dot{\boldsymbol{\alpha}} = 0 \tag{4.53}$$

where $\boldsymbol{r}_{\boldsymbol{\alpha}} \in \mathbf{R}^{3\times 6}$.

Next take the full time derivative of the velocity vector

$$\dot{\boldsymbol{v}} = \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{\alpha}} \dot{\boldsymbol{\alpha}} + \frac{\partial \boldsymbol{v}}{\partial t} \tag{4.54}$$

$$= \frac{\partial U}{\partial \boldsymbol{r}} + \frac{\partial R}{\partial \boldsymbol{r}} \tag{4.55}$$

By definition $\frac{\partial \boldsymbol{v}}{\partial t} = \frac{\partial U}{\partial \boldsymbol{r}}$ and can be cancelled, leaving the second relation

$$\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{\alpha}} \dot{\boldsymbol{\alpha}} = \frac{\partial R}{\partial \boldsymbol{r}} \tag{4.56}$$

where $\boldsymbol{v}_{\boldsymbol{\alpha}} \in \mathbf{R}^{3\times 6}$.

These two relations can be combined into a single matrix equation:

$$\begin{bmatrix} \boldsymbol{r}_{\boldsymbol{\alpha}} \\ \boldsymbol{v}_{\boldsymbol{\alpha}} \end{bmatrix} \dot{\boldsymbol{\alpha}} = \begin{bmatrix} 0 \\ R_{\boldsymbol{r}}^T \end{bmatrix} \tag{4.57}$$

As the constants $\boldsymbol{\alpha}$ are assumed to provide a complete solution of the system, the matrix is invertible and can be formally solved for the equations $\dot{\boldsymbol{\alpha}}$. This is not a particularly useful approach to the problem, however, as the functions $\boldsymbol{r}$ and $\boldsymbol{v}$ must be used at every step of the way.

Lagrange introduced an elegant simplification of this problem, as explained in [19]. Pre-multiply each side of the equation by the matrix $[\boldsymbol{r}_{\boldsymbol{\alpha}}^T; \boldsymbol{v}_{\boldsymbol{\alpha}}^T] J$, where J is the symplectic matrix introduced earlier. Carrying out the multiplications yields:

$$\left[\boldsymbol{r}_{\boldsymbol{\alpha}}^T \boldsymbol{v}_{\boldsymbol{\alpha}} - \boldsymbol{v}_{\boldsymbol{\alpha}}^T \boldsymbol{r}_{\boldsymbol{\alpha}}\right] \dot{\boldsymbol{\alpha}} = \left[\boldsymbol{r}_{\boldsymbol{\alpha}}^T R_{\boldsymbol{r}}^T\right] \tag{4.58}$$

Now note a few items that will simplify these results. First, $R_{\boldsymbol{\alpha}} = R_{\boldsymbol{r}} \cdot \boldsymbol{r}_{\boldsymbol{\alpha}}$ under the assumption that R is a function of position only, thus the term on the right-hand side can be replaced by $R_{\boldsymbol{\alpha}}^T$. The implication of this is that the relation for $\boldsymbol{r}$ can be substituted as a function of $\boldsymbol{\alpha}$ into the perturbing potential and one only needs deal with a perturbing potential $R(\boldsymbol{\alpha})$.

Next note that the general term i, j in the matrix on the left-hand side can be expressed as

$$[\alpha_i, \alpha_j] = \boldsymbol{r}_{\alpha_i} \cdot \boldsymbol{v}_{\alpha_j} - \boldsymbol{v}_{\alpha_i} \cdot \boldsymbol{r}_{\alpha_j} \tag{4.59}$$

The operator $[-,-]$ is called a Lagrange bracket and serves as the inverse to the Poisson bracket. Its properties are discussed in detail in [19], and will be discussed to a limited extent here. An important feature of the Lagrange bracket is that, for a given dynamical system, it is only a function of the constants of motion $\boldsymbol{\alpha}$ that have been chosen. Indeed, it is important to note that we have not chosen the specific constants yet in this problem, and have only relied on their existence. Using the Lagrange brackets the equation can be rewritten for the variation of parameters as

$$[[\alpha_i, \alpha_j]] \dot{\boldsymbol{\alpha}} = R_{\boldsymbol{\alpha}}^T \tag{4.60}$$

where $[[\alpha_i, \alpha_j]]$ represents a matrix with elements $[\alpha_i, \alpha_j]$, these elements taking on all values of $i, j = 1, 2, \ldots, 6$. This matrix is invertible, as it is just formed by the multiplication of non-singular matrices. Carrying out this inversion then yields:

$$\dot{\boldsymbol{\alpha}} = [[\alpha_i, \alpha_j]]^{-1} R_{\boldsymbol{\alpha}}^T \tag{4.61}$$

Before the inversion of this matrix is discussed we point out a crucial property: each Lagrange bracket is constant in time. To show this, take the total time derivative of an arbitrary term in the matrix:

$$\frac{d}{dt}[\alpha_i, \alpha_j] = \boldsymbol{v}_{\alpha_i} \cdot \boldsymbol{v}_{\alpha_j} + \boldsymbol{r}_{\alpha_i} \cdot \dot{\boldsymbol{v}}_{\alpha_j} - \boldsymbol{v}_{\alpha_i} \cdot \boldsymbol{v}_{\alpha_j} - \dot{\boldsymbol{v}}_{\alpha_i} \cdot \boldsymbol{r}_{\alpha_j} \tag{4.62}$$

$$= \boldsymbol{r}_{\alpha_i} \cdot \frac{\partial}{\partial \alpha_j} [U_{\boldsymbol{r}} + R_{\boldsymbol{r}}] - \boldsymbol{r}_{\alpha_j} \cdot \frac{\partial}{\partial \alpha_i} [U_{\boldsymbol{r}} + R_{\boldsymbol{r}}] \tag{4.63}$$

$$= \boldsymbol{r}_{\alpha_i} \cdot [U_{\boldsymbol{rr}} + R_{\boldsymbol{rr}}] \cdot \boldsymbol{r}_{\alpha_j} - \boldsymbol{r}_{\alpha_j} \cdot [U_{\boldsymbol{rr}} + R_{\boldsymbol{rr}}] \cdot \boldsymbol{r}_{\alpha_i} \tag{4.64}$$

$$= 0 \tag{4.65}$$

as the dyadics $U_{\boldsymbol{rr}}$ and $R_{\boldsymbol{rr}}$ are symmetric. This is an important fact, as it implies that the matrix of Lagrange brackets need only be evaluated functionally only once, and that the resulting inversion that occurs is applicable anywhere during the subsequent perturbed motion. It will also have important implications for averaging, as it allows us to pull the averaging operator across this term.

Lagrange Planetary Equations: Standard Form

Even though the derivation has been simple up to this point, the actual computation of the Lagrange brackets for a particular choice of constants is quite onerous. For the details we refer the reader to [19]. If the $\boldsymbol{\alpha}$ are chosen to be the standard orbital elements: $(a, e, i, \omega, \Omega, \sigma)$, where $\sigma = -n\tau_o$ and τ_o is the time of periapsis passage, the following form for the equations are found [19]

$$\dot{a} = \frac{2}{na} \frac{\partial R}{\partial \sigma} \tag{4.66}$$

$$\dot{e} = \frac{1}{na^2 e} \left[(1 - e^2) \frac{\partial R}{\partial \sigma} - \sqrt{1 - e^2} \frac{\partial R}{\partial \omega} \right] \tag{4.67}$$

$$\dot{i} = \frac{1}{na^2 \sqrt{1 - e^2}} \left[\cot i \frac{\partial R}{\partial \omega} - \csc i \frac{\partial R}{\partial \Omega} \right] \tag{4.68}$$

$$\dot{\omega} = \frac{\sqrt{1 - e^2}}{na^2 e} \frac{\partial R}{\partial e} - \frac{\cot i}{na^2 \sqrt{1 - e^2}} \frac{\partial R}{\partial i} \tag{4.69}$$

$$\dot{\Omega} = \frac{\csc i}{na^2 \sqrt{1 - e^2}} \frac{\partial R}{\partial i} \tag{4.70}$$

$$\dot{\sigma} = -\frac{(1 - e^2)}{na^2 e} \frac{\partial R}{\partial e} - \frac{2}{na} \frac{\partial R}{\partial a} \tag{4.71}$$

where $n = \sqrt{\mu / a^3}$ is the mean motion. While complicated, all of the coefficients in front of the partials of R are time-invariant, even though their values change over time, driven by these ordinary differential equations.

There are many variations on these equations. Common changes are to replace the argument of periapsis with the longitude of periapsis, $\varpi = \omega + \Omega$, to replace the expression σ with some a term that is a function of the mean anomaly, and to redefine the orbit elements to eliminate the terms e and $\sin i$ in the denominators.

Delaunay Variables

There is freedom in choosing the specific constants of motion used in the variation of parameters. Thus, one can ask if there are particular choices which can simplify the equations, or place them into a standard form. In fact, it is possible to do so, and in particular it is possible to choose the $\boldsymbol{\alpha}$ so that the transformation $\boldsymbol{r}(t, \boldsymbol{\alpha}), \boldsymbol{v}(t, \boldsymbol{\alpha})$ is canonical, as the Newtonian form of the equations of motion are themselves in Hamiltonian form. Then the matrix $[\boldsymbol{r}_{\boldsymbol{\alpha}}^T; \boldsymbol{v}_{\boldsymbol{\alpha}}^T]$ is symplectic by definition and $[[\alpha_i, \alpha_j]] = J$.

One particular choice of constants that enable such a transformation are called Delaunay variables. The resulting equation of motion can be stated as

$$\dot{\sigma} = -\frac{\partial R}{\partial L} \tag{4.72}$$

$$\dot{g} = -\frac{\partial R}{\partial H} \tag{4.73}$$

$$\dot{h} = -\frac{\partial R}{\partial G} \tag{4.74}$$

$$\dot{L} = \frac{\partial R}{\partial \sigma} \tag{4.75}$$

$$\dot{H} = \frac{\partial R}{\partial h} \tag{4.76}$$

$$\dot{G} = \frac{\partial R}{\partial g} \tag{4.77}$$

where the symbols G and H and g and h are traditionally switched in the Delaunay notation, respectively, σ is defined as before, $h = \omega$, the argument of periapsis, and $g = \Omega$, the longitude of the ascending node. These are usually termed the angles or the coordinates. The conjugate momenta are all functions of the orbital elements a, e and i as follows:

$$L = \sqrt{\mu a} \tag{4.78}$$

$$H = L\sqrt{1 - e^2} \tag{4.79}$$

$$G = H \cos i \tag{4.80}$$

where L is a direct function of the energy, H equals the angular momentum magnitude, and G is the angular momentum projected on the z-axis. To cast these variables into proper Hamiltonian form, define the system Hamiltonian as $\mathcal{H} = -R$, with the implicit assumption that the two-body potential acts on the original system.

A frequent change in these variables is to introduce a modified Hamiltonian:

$$\mathcal{H} = -\frac{\mu^2}{2L^2} - R \tag{4.81}$$

and replace the epoch σ term with the mean anomaly, $l = \int n\,dt + \sigma$. The new Hamiltonian system is then:

$$\dot{l} = \frac{\partial \mathcal{H}}{\partial L} \tag{4.82}$$

$$\dot{g} = \frac{\partial \mathcal{H}}{\partial H} \tag{4.83}$$

$$\dot{h} = \frac{\partial \mathcal{H}}{\partial G} \tag{4.84}$$

$$\dot{L} = -\frac{\partial \mathcal{H}}{\partial l} \tag{4.85}$$

$$\dot{H} = -\frac{\partial \mathcal{H}}{\partial h} \tag{4.86}$$

$$\dot{G} = -\frac{\partial \mathcal{H}}{\partial g} \tag{4.87}$$

4.4.2 Gauss Equations

The Gauss equations are usually derived either from the Lagrange equations or in a direct manner. An alternate approach is given here that is sometimes useful when deriving similar equations for nonstandard coordinates or problems. Now the perturbing acceleration is represented as an explicit vector $\boldsymbol{a}$. Thus, the equations of motion become:

$$\ddot{\boldsymbol{r}} = U_{\boldsymbol{r}} + \boldsymbol{a} \tag{4.88}$$

Again, the equations with $\boldsymbol{a} \equiv 0$ have a solution expressed as a function of a set of independent constants: $\boldsymbol{r}(t, \boldsymbol{\alpha})$ and $\boldsymbol{v}(t, \boldsymbol{\alpha})$.

Now assume that these functions have been inverted to find the constants of motion as a function of the particle position and velocity at a particular time:

$$\boldsymbol{\alpha} = \boldsymbol{\alpha}(\boldsymbol{r}, \boldsymbol{v}, t) \tag{4.89}$$

When the acceleration $\boldsymbol{a} \equiv 0$ the $\boldsymbol{\alpha}$ are constant and thus $\dot{\boldsymbol{\alpha}} = \boldsymbol{\alpha}_{\boldsymbol{r}} \cdot \boldsymbol{v} + \boldsymbol{\alpha}_{\boldsymbol{v}} \cdot U_{\boldsymbol{r}} + \boldsymbol{\alpha}_t = 0$, where the substitutions $\dot{\boldsymbol{r}} = \boldsymbol{v}$ and $\dot{\boldsymbol{v}} = U_{\boldsymbol{r}}$ are used. The $\boldsymbol{\alpha}$ are again osculating elements.

Take the total time derivative of $\boldsymbol{\alpha}$ for the general case with a non-zero perturbing acceleration to find

$$\dot{\boldsymbol{\alpha}} = \frac{\partial \boldsymbol{\alpha}}{\partial \boldsymbol{r}} \cdot \dot{\boldsymbol{r}} + \frac{\partial \boldsymbol{\alpha}}{\partial \boldsymbol{v}} \cdot \dot{\boldsymbol{v}} + \frac{\partial \boldsymbol{\alpha}}{\partial t} \tag{4.90}$$

$$= \frac{\partial \boldsymbol{\alpha}}{\partial \boldsymbol{r}} \cdot \dot{\boldsymbol{r}} + \frac{\partial \boldsymbol{\alpha}}{\partial \boldsymbol{v}} \cdot [U_{\boldsymbol{r}} + \boldsymbol{a}] + \frac{\partial \boldsymbol{\alpha}}{\partial t} \tag{4.91}$$

However, the terms $\boldsymbol{\alpha}_{\boldsymbol{r}} \cdot \boldsymbol{v} + \boldsymbol{\alpha}_{\boldsymbol{v}} \cdot U_{\boldsymbol{r}} + \boldsymbol{\alpha}_t = 0$ by definition. Thus this yields the very simple, and general, form of the Gauss equations:

$$\dot{\boldsymbol{\alpha}} = \boldsymbol{\alpha}_{\boldsymbol{v}} \cdot \boldsymbol{a} \tag{4.92}$$

Thus, to compute these all one must do is define an appropriate set of integrals of motion, take their partials with respect to the velocity, and dot them with the perturbing acceleration vector.

Phrasing the Gauss equations in this form allows us to perform some derivations very easily. Consider the two-body problem with an applied acceleration and use the following integrals of motion, the two-body energy $\mathcal{E}$, the angular momentum vector $\boldsymbol{H}$, and the eccentricity (or Laplace) vector $\boldsymbol{e}$. These are defined in terms of the spacecraft state as:

$$\mathcal{E} = \frac{1}{2}\boldsymbol{v}\cdot\boldsymbol{v} - \frac{\mu}{|\boldsymbol{r}|} \tag{4.93}$$

$$\boldsymbol{H} = \boldsymbol{r}\times\boldsymbol{v} \tag{4.94}$$

$$\boldsymbol{e} = \frac{1}{\mu}\boldsymbol{v}\times\boldsymbol{r}\times\boldsymbol{v} - \frac{\boldsymbol{r}}{|\boldsymbol{r}|} \tag{4.95}$$

To apply this result one only needs take the partial of each of these with respect to the velocity vector and then dot the resulting quantity with the acceleration vector, yielding

$$\dot{\mathcal{E}} = \boldsymbol{v}\cdot\boldsymbol{a} \tag{4.96}$$

$$\dot{\boldsymbol{H}} = \tilde{\boldsymbol{r}}\cdot\boldsymbol{a} \tag{4.97}$$

$$\dot{\boldsymbol{e}} = \frac{1}{\mu}\left[\boldsymbol{a}\cdot\widetilde{(\boldsymbol{r}\times\boldsymbol{v})} + \tilde{\boldsymbol{v}}\cdot\tilde{\boldsymbol{r}}\cdot\boldsymbol{a}\right] \tag{4.98}$$

$$= \frac{1}{\mu}\left[2\boldsymbol{r}\boldsymbol{v} - \boldsymbol{v}\boldsymbol{r} - (\boldsymbol{r}\cdot\boldsymbol{v})\boldsymbol{U}\right]\cdot\boldsymbol{a} \tag{4.99}$$

Decomposing the acceleration vector into orthogonal components along the radius vector, S, along the angular momentum vector, W, and along the normal to these in the direction of travel, T, yields

$$\boldsymbol{a} = S\hat{\boldsymbol{r}} + T\hat{\boldsymbol{\theta}} + W\hat{\boldsymbol{H}} \tag{4.100}$$

Using this decomposition an explicit form for the time derivatives of these quantities is

$$\dot{\mathcal{E}} = \frac{H}{r}(\tan\gamma S + T) \tag{4.101}$$

$$\dot{\boldsymbol{H}} = r(T\hat{\boldsymbol{H}} - W\hat{\boldsymbol{\theta}}) \tag{4.102}$$

$$\dot{\boldsymbol{e}} = \frac{H}{\mu}\left[2T\hat{\boldsymbol{r}} - (S + \tan\gamma T)\hat{\boldsymbol{\theta}} - \tan\gamma W\hat{\boldsymbol{H}}\right] \tag{4.103}$$

In terms of orbital elements and integrals of motion the unit vectors have the following definitions:

$$\hat{\boldsymbol{r}} = \cos f\hat{\boldsymbol{e}} + \sin f\hat{\boldsymbol{e}}_\perp \tag{4.104}$$

$$\hat{\boldsymbol{\theta}} = -\sin f\hat{\boldsymbol{e}} + \cos f\hat{\boldsymbol{e}}_\perp \tag{4.105}$$

$$\hat{\boldsymbol{H}} = \frac{\boldsymbol{H}}{H} \tag{4.106}$$

$$\hat{\boldsymbol{e}} = \frac{\boldsymbol{e}}{e} \tag{4.107}$$

$$\hat{\boldsymbol{e}}_\perp = \hat{\boldsymbol{H}} \times \hat{\boldsymbol{e}} \tag{4.108}$$

while the radius r and $\tan\gamma$ have the usual two-body relations and are themselves functions of true anomaly f. It is interesting to note that when $e = 0$, where the Gauss equations are usually singular in argument of periapsis, that the equation for the Laplace vector is well defined, $\dot{\boldsymbol{e}}|_{e=0} = \frac{H}{\mu}\left[2T\hat{\boldsymbol{r}} - S\hat{\boldsymbol{\theta}}\right]$, and thus can serve as a non-singular coordinate. Similarly for the angular momentum in terms of inclination and longitude of the ascending node.

To derive the traditional Gauss equations from the above one only needs to express the integrals as functions of orbital elements and take time derivatives. The simplest example is for the energy, where $\mathcal{E} = -\frac{\mu}{2a}$ and $\dot{\mathcal{E}} = \frac{\mu}{2a^2}\dot{a}$ leading to

$$\dot{a} = \frac{2a^2}{\mu}\frac{H}{p}\left[e\sin fS + (1 + e\cos f)T\right] \tag{4.109}$$

which can be further simplified by substituting for H and p. Carrying out these evaluations for the classical orbital elements yields:

$$\dot{a} = \frac{2}{n\sqrt{1-e^2}}\left(e\sin fS + (1 + e\cos f)T\right) \tag{4.110}$$

$$\dot{e} = \frac{H}{\mu}\left[\sin fS + \left(\frac{e + \cos f}{1 + e\cos f} + \cos f\right)T\right] \tag{4.111}$$

$$\dot{i} = \frac{H}{\mu}\frac{\cos(\omega + f)}{1 + e\cos f}W \tag{4.112}$$

$$\dot{\omega} = \frac{H}{\mu e}\left[\frac{(2 + e\cos f)\sin f}{1 + e\cos f}T - \cos fS\right] - \cos i\dot{\Omega} \tag{4.113}$$

$$\dot{\Omega} = \frac{H}{\mu\sin i}\frac{\sin(\omega + f)}{1 + e\cos f}W \tag{4.114}$$

$$\dot{\sigma} = \frac{r}{na^2e}\left[(\cos f + e\cos^2 f - 2e)S - \sin f(2 + e\cos f)T\right] \tag{4.115}$$

A key item of contrast between the Lagrange and Gauss equations is in their coefficients. The Gauss equations have time-varying terms multiplying the accelerations involving the true anomaly, thus they must each be averaged separately.

5. Properties of Solution

This chapter discusses the properties of solutions to dynamical systems, and in particular focuses on Hamiltonian dynamical systems. The justification for this restriction is given below, but essentially arises from the fact that almost all of the relevant problems of interest to us can be transformed into this general form. We take advantage of this, as many of the fundamental properties that are of interest to us can be easily derived when the system is stated in Hamiltonian form.

5.1 Reduction to Time Invariant Hamiltonian Dynamical Systems

The reader is reminded that all of the generic equations of motion studied in the previous chapter can be transformed into a Hamiltonian formulation. Thus to study the general properties of these solutions it suffices to study the properties of Hamiltonian systems. First, note that the different formulations can all be transformed into a time-varying Hamiltonian system. Starting with the Newtonian formulation, the Hamiltonian function can be defined as the specific energy, independent of whether this is a conserved quantity. Thus $H = \frac{1}{2}\dot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}} - \mathcal{U}(\boldsymbol{r}, t)$ with the coordinates being the inertial position vectors $\boldsymbol{r}$ and the momenta being the inertial velocities $\dot{\boldsymbol{r}}$. It is simple to verify that Hamilton's equations recover the Newtonian equations of motion. Thus the Newtonian formulation is trivially in Hamiltonian form.

Next consider the Lagrangian form of the dynamics, along with a Lagrangian function $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$. As explicitly derived, one can transform these directly into a Hamiltonian form by a simple transformation in which the coordinates remain fixed and the new momenta are $\boldsymbol{p} = \frac{\partial L}{\partial \dot{\boldsymbol{q}}}$. The Hamiltonian function is then constructed as $H(\boldsymbol{q}, \boldsymbol{p}, t) = \dot{\boldsymbol{q}} \cdot \boldsymbol{p} - L$, where the $\dot{\boldsymbol{q}}$ must be explicitly solved for in terms of the coordinates and momenta.

Finally, note that the Lagrange Planetary Equations can be rewritten into a canonical form, such as the Delaunay variables, as explicitly discussed in the previous chapter. Similar transformations can also be applied to the Gauss equations,

D.J. Scheeres, *Orbital Motion in Strongly Perturbed Environments*, Springer Praxis Books,
DOI 10.1007/978-3-642-03256-1_5, © Springer-Verlag Berlin Heidelberg 2012

although it is not always possible to recast the perturbing accelerations as a perturbing potential.

Once a system is placed into a Hamiltonian form, then any application of a canonical transformation will preserve the structure of its equations of motion. Thus, in a very real sense, the specific formulation of the problem does not matter as much, especially as it relates to generic properties of these systems.

Of specific interest are dynamical systems whose Hamiltonian function can be made time invariant by transformation to a special set of coordinates. If this can be done via a canonical transformation, then the system will have an integral of motion, called the Jacobi integral, associated with it. For a time invariant Hamiltonian the Jacobi integral is just the Hamiltonian function. The number of systems for which such a transformation exists is quite limited, however, so that for most general problems of interest the Hamiltonian will always be time varying or time periodic at best.

However, any time varying Hamiltonian can always be related to a time invariant Hamiltonian system, at the cost of raising the dimension of the problem by 2. The procedure for this is quite simple. Assume a Hamiltonian system with $\boldsymbol{q} \in \mathbf{R}^n$ and $\boldsymbol{p} \in \mathbf{R}^n$, and a time varying Hamiltonian $H(\boldsymbol{q}, \boldsymbol{p}, t)$. Introduce a new independent coordinate $\tau(t)$ with its inverse $t(\tau)$, so that $dt/d\tau = t'$, where the $'$ denotes differentiation with respect to the new independent parameter τ. It is possible to take τ identically equal to the time ($t' = 1$), or one can introduce some other scaling, but the resulting expression t' should be independent of τ in order to develop a time invariant system. Formally define the new time coordinate as $q_t = t$ and the new time-momentum as p_t. The new Hamiltonian can then be constructed as $\mathcal{H}(\boldsymbol{q}, q_t, \boldsymbol{p}, p_t) = (H(\boldsymbol{q}, \boldsymbol{p}, q_t) + p_t)\, t'$. Note that this now provides the equations of motion with τ as the new independent variable. Computing Hamilton's equations of motion

$$\boldsymbol{q}' = \frac{\partial H}{\partial \boldsymbol{p}} t' \tag{5.1}$$

$$q_t' = t' \tag{5.2}$$

$$\boldsymbol{p}' = -\frac{\partial H}{\partial \boldsymbol{q}} t' \tag{5.3}$$

$$p_t' = -\frac{\partial H}{\partial q_t} t' \tag{5.4}$$

The original equations of motion for $\boldsymbol{q}$ and $\boldsymbol{p}$ are recovered with the new independent parameter. The time equation is trivially satisfied given the specified relation between τ and t. Finally, note that the time variation of the time-momentum equals the negative time partial of the Hamiltonian. Writing this out in differential form yields $dp_t = -H_t\, dt$ which can be integrated to find

$$p_t(t) - p_t(t_o) = -H(\boldsymbol{q}(t), \boldsymbol{p}(t), t) + H(\boldsymbol{q}(t_o), \boldsymbol{p}(t_o), t_o) \tag{5.5}$$

where we have shifted back to the original time notation. Rearranging exposes the new conservation relation

$$p_t(t) + H(t) = p_t(t_o) + H(t_o) \tag{5.6}$$

using an abbreviated notation. Thus, the new Hamiltonian $\mathcal{H} = p_t + H$ is formally time invariant and has a conservation relation defined for it. The time-momentum equals the negative of the original time-varying Hamiltonian plus an arbitrary constant.

Due to the existence of this result, one can formally treat any Hamiltonian system as a time invariant system. This allows us to only consider this case in these detailed discussions, although the time varying form will still be used where appropriate.

5.2 Properties of General Trajectories

Given a general solution to a Hamiltonian dynamical system there are a number of fundamental results that can be ascribed to the solution flow. While some of these properties also exist for arbitrary dynamical systems, Hamiltonian systems have additional properties that make them distinguished.

Denote the solution to a Hamiltonian system that passes through a point of phase space $\boldsymbol{x}_o$ at an epoch t_o to be $\boldsymbol{x}(t) = \phi(t; \boldsymbol{x}_o, t_o)$. Assume the ordering $\boldsymbol{x} = [\mathbf{q}, \mathbf{p}]$. Thus, the solution flow $\phi(t)$ satisfies the Hamiltonian equations of motion:

$$\frac{\partial \phi}{\partial t} = J \left. \frac{\partial H}{\partial \boldsymbol{x}} \right|_{x(t)=\phi(t)} \tag{5.7}$$

By definition, a trajectory is a canonical transformation between the state $\boldsymbol{x}_o$ and the state $\boldsymbol{x}$, with time t being an arbitrary parameter of the transformation. Thus, the Jacobian of the solution must satisfy the property that $\Phi(t, t_o; \boldsymbol{x}_o) = \partial \boldsymbol{x}(t)/\partial \boldsymbol{x}_o$ be a symplectic matrix (the functional dependence of Φ on the state will often be suppressed). This provides a few important technical results which lead to additional functional results. The basic technical result is that $|\Phi| = 1$, namely that the solution flow is never singular and can always be solved for the inverse of the flow. Denote this inverse using the solution flow description as $\boldsymbol{x}_o = \phi(t_o; \boldsymbol{x}, t)$. Further, for a linear dynamical system where the solution has the form $\boldsymbol{x}(t) = \Phi(t, t_o)\boldsymbol{x}_o$ this inversion can be carried out in closed form as

$$\Phi(t_o, t) = \Phi^{-1}(t, t_o) \tag{5.8}$$

$$\Phi^{-1}(t, t_o) = -J\Phi(t, t_o)^T J \tag{5.9}$$

from the properties of symplectic matrices (Appendix D).

5.2.1 Initial Conditions and Higher-Order Expansions

Solutions of Hamilton dynamical systems are analytic in initial conditions over finite timespans, even for chaotic systems. Thus, one can expand a nominal solution into a higher-order series in order to generate a local characterization of the solution flow. This is a familiar technique used for linear variations about a nominal trajectory, but can be expanded to higher orders. The main issue is to develop a notation that supports such higher-order expansions, as a matrix formulation cannot be unambiguously applied beyond the first order. Thus for this section an index notation is introduced, representing the state at a time t by

$$x_i(t) = \phi_i(t; t_o, \boldsymbol{x}_o) \tag{5.10}$$

$$i = 1, 2, \ldots, 2n \tag{5.11}$$

Using this notation, the partial of x_i with respect to the initial conditions $x_j(t_o)$ will be denoted as $x_{i,j}$, with the implication that the sub j index after the comma implies a partial with respect to the initial state. Specifically:

$$\frac{\partial \phi_i(t; t_o, \boldsymbol{x}_o)}{\partial x_j(t_o))} = \phi_{i,j}(t; t_o, \boldsymbol{x}_o) \tag{5.12}$$

with its natural generalizations to higher orders. For the first partial, note the correspondence between $\boldsymbol{\Phi} = \phi_{i,j}$, where i denotes the row and j denotes the column.

Using this notation the function $\phi(t; t_o, \boldsymbol{x}_o + \delta\boldsymbol{x}_o)$ can be expanded about the nominal initial condition $\boldsymbol{x}_o$ to find:

$$\begin{aligned} x_i(t) + \delta x_i(t) = \phi_i(t; t_o, \boldsymbol{x}_o) + \phi_{i,j}(t; t_o, \boldsymbol{x}_o)\delta x_j(t_o) \\ + \frac{1}{2!}\phi_{i,jk}(t; t_o, \boldsymbol{x}_o)\delta x_j(t_o)\delta x_k(t_o) + \ldots \end{aligned} \tag{5.13}$$

where the Einstein summation convention is used. The higher-order partials $\phi_{i,jk\ldots n}$ are called "state transition tensors" in [124] and can be computed analytically if a closed-form solution exists to the dynamical system, or numerically as discussed in the next chapter.

Given a nominal trajectory $x_i(t)$ these results allow us to describe motion in the neighborhood of that trajectory both forward and backwards in time. The forward solution is trivially inferred from the above and equals:

$$\delta x_i(t) = \phi_{i,j}(t; t_o, \boldsymbol{x}_o)\delta x_j(t_o) + \frac{1}{2!}\phi_{i,jk}(t; t_o, \boldsymbol{x}_o)\delta x_j(t_o)\delta x_k(t_o) + \ldots \tag{5.14}$$

With such a series it is possible to perform a reversion of series (i.e., invert the series to solve for $\boldsymbol{x}_o$), so long as the leading linear term is invertible. This is guaranteed by the symplectic nature of the state transition matrix. Carrying out this reversion

to the first two orders yields:

$$\delta x_i(t_o) = \psi_{i,j}\delta x_j(t) + \frac{1}{2!}\psi_{i,jk}\delta x_j(t)\delta x_k(t) + \dots \tag{5.15}$$

where $\psi_{i,j} = \phi_{i,j}^{-1} = -J_{il}\phi_{k,l}J_{kj}$, and $\psi_{i,jk} = -\phi_{il}^{-1}\psi_{l,mn}\phi_{mo}^{-1}\phi_{np}^{-1}$, with the higher-order terms being more complex (see [124] for a more complete discussion).

5.2.2 Solutions Analytic in a Parameter

Solutions can also be viewed as functions of free parameters in the system. Given a dynamical system defined as

$$\dot{\boldsymbol{x}} = JH_{\boldsymbol{x}}(\boldsymbol{x}, t; \mu), \tag{5.16}$$

where μ represents a parameter of the system, one can define a family of solutions as a function of this independent parameter μ. The main assumption is that the Hamiltonian function is analytic in the parameter μ, which is generally the case for physically defined problems. Examples of such parameters include masses or reduced masses of attracting bodies, physically defined parameters of a spacecraft or system that can take on different values, or even values of integrals of motion of the system. One common property that parameters have is that they are constant in time. Time-varying parameters can also be studied, under some assumptions, using the theory of adiabatic invariants (see [4]), but are not considered here.

Generally, for a specific dynamical system these parameters would not be varied, but instead would just be chosen and assigned a specific value. However, it is often convenient to allow these parameters to vary in order to understand the robustness of a particular solution to slightly different models, or to understand the global stability properties of a dynamical system across all values of a given parameter. An example of the latter would be varying the mass fraction in the restricted three-body problem from 0 to 1/2 in order to determine the stability properties of a certain class of motion.

One specific application is to the continuation of a family of periodic orbits as a parameter is varied. Consider a solution to Eq. 5.16, denoted as $\boldsymbol{x}(t) = \phi(t; \boldsymbol{x}_o, t_o, \mu)$. Such a solution can be found by numerical integration or by analytic solution. If they are obtained analytically, one can immediately describe the properties of the solution as the parameter is varied. For a numerically computed solution this description is not as easy, but can be derived directly from the equations of motion combined with a numerical integration. Consider a variation in the parameter μ:

$$\boldsymbol{x}(t) + \delta\boldsymbol{x}_\mu(t) = \phi(t; \boldsymbol{x}_o, t_o, \mu + \delta\mu) \tag{5.17}$$

$$= \phi(t; \boldsymbol{x}_o, t_o, \mu) + \frac{\partial\phi}{\partial\mu}\delta\mu + \dots \tag{5.18}$$

At the leading order the desired variation in the solution can be found

$$\delta \boldsymbol{x}_\mu(t) = \frac{\partial \phi}{\partial \mu}\delta\mu + \dots \tag{5.19}$$

and note that this variation is controlled by the partial $\phi_\mu = \frac{\partial \phi}{\partial \mu}$. Computation of this partial is most precise if the variational equation is integrated along the solution. Applying standard linearization approaches yields

$$\frac{d}{dt}\phi_\mu = JH_{xx}\phi_\mu + JH_{\boldsymbol{x}\mu} \tag{5.20}$$

where the $H_{\boldsymbol{xx}}$ and $H_{\boldsymbol{x}\mu}$ terms are evaluated along the nominal solution contemporaneously with ϕ_μ. The initial conditions for this solution is simply $\phi_\mu = 0$.

These equations are, in general, non-homogenous, time-varying differential equations. Given a state transition matrix for the system, $\Phi(t, t_o)$, a general solution for the partial derivative can be written out as

$$\phi_\mu(t; \boldsymbol{x}_o, t_o, \mu) = \int_{t_o}^{t} \Phi(t, \tau) H_{\boldsymbol{x}\mu}(\boldsymbol{x}(\tau), \tau, \mu)\, d\tau \tag{5.21}$$

From a practical perspective, it is often easiest to directly solve the differential equation Eq. 5.20 for ϕ_μ along with the nominal solution and the state transition matrix.

These computations can be generalized to higher-order systems of equations following the approaches previously outlined.

5.2.3 Eigenstructure of the State Transition Matrix

The symplecticity of $\Phi \in \mathbf{R}^{6\times 6}$ has important consequences for the structure of this matrix, which in turn have important implications for Hamiltonian dynamical systems. A matrix is defined, most generally, in terms of its eigenstructure, which can be considered to consist of its eigenvalues, right eigenvectors and left eigenvectors, and potentially generalized eigenvectors. The right eigenvector of an eigenvalue λ_i is denoted as $\boldsymbol{u}_i \in \mathbf{R}^6$ and satisfies $\Phi \boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i$. A left eigenvector of the same eigenvalue is denoted as $\boldsymbol{v}_i \in \mathbf{R}^6$ and satisfies $\Phi^T \boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i$. In general, it is possible to represent a matrix without repeated eigenvalues in terms of its eigenvalues and eigenvectors as:

$$\Phi = \sum_{i=1}^{2n} \lambda_i \boldsymbol{u}_i \boldsymbol{v}_i^T \tag{5.22}$$

$$= [U][\lambda][V]^T \tag{5.23}$$

where $[U]$ and $[V]$ are the right and left eigenvectors arranged as columns, respectively, and $[\lambda]$ is a diagonal matrix with the eigenvalues along its diagonal. The more general case of repeated eigenvalues is discussed later.

Eigenvalues

The eigenvalues of Φ can be discussed using the defining equation

$$\|\lambda I - \Phi\| = 0 \tag{5.24}$$

where $\|-\|$ denotes the determinant and I is the 6×6 identity matrix. Given that Φ is symplectic, and thus has non-zero eigenvalues, it is easy to show, using the inverse form of Φ (see Appendix D), that

$$\left\|\lambda^{-1} I - \Phi\right\| = 0 \tag{5.25}$$

where $\lambda\lambda^{-1} = 1$. Thus, every eigenvalue will have a matching inverse. For a general complex eigenvalue we can describe it as $\lambda = \rho\, e^{i\theta}$. Then its inverse has the description $\lambda^{-1} = \frac{1}{\rho}\, e^{-i\theta}$, where ρ is a real number, i is the imaginary number and e^x is the exponential function. One fundamental implication of this is that the characteristic polynomial for any symplectic Φ is symmetric, i.e., has the form:

$$\lambda^{2n} + a_1\lambda^{2n-1} + \ldots + a_m\lambda^{2n-m} + \ldots + a_m\lambda^m + \ldots + a_1\lambda + 1 = 0 \tag{5.26}$$

Thus every entry in the polynomial of form $a_m\lambda^{2n-m}$ has a counterpart $a_m\lambda^m$. Since this reciprocal relationship holds for arbitrary values of the a_i, it follows that when a symplectic matrix Φ has a unity eigenvalue, that these must also appear in pairs. In this case the eigenvectors can become more difficult to compute, and will be discussed later. Another constraint arises in the fact that Φ is a real matrix, and thus all complex eigenvalues must have a complex conjugate counterpart. Specifically, if there exists an eigenvalue λ_i which has a complex part, then $\bar{\lambda}_i$, its complex conjugate, must also be an eigenvalue.

The most general situation occurs for an eigenvalue that is complex and not of unit magnitude, $\lambda = \rho\, e^{i\theta}$, where ρ and θ are real. Then the inverse eigenvalue is $\frac{1}{\rho}\, e^{-i\theta}$, the complex conjugate of the original eigenvalue is $\rho\, e^{-i\theta}$, and the complex conjugate of the inverse is $\frac{1}{\rho}\, e^{i\theta}$, forming a set of four eigenvalues. If an eigenvalue is real, then $\lambda = \rho$ and identically equals its complex conjugate and its inverse is just the reciprocal, $\lambda^{-1} = 1/\rho$, forming a pair. If an eigenvalue has a unit magnitude, then $\lambda = e^{i\theta}$ and its complex conjugate is equal to its inverse, or $\lambda^{-1} = \bar{\lambda} = e^{-i\theta}$, again forming a pair.

The usual approach to representing the eigenvalues of a symplectic matrix is on the complex plane, with the unit circle serving to orient the values. As any free parameter of the dynamical system is varied, the eigenvalues of the associated state transition matrix will evolve on the complex plane. Due to their strict constraints, the eigenvalues can only transition between different classes by intersecting each other either on the unit circle or on the real axis.

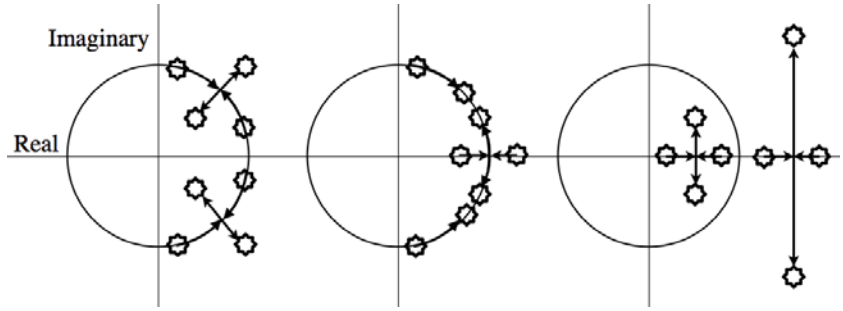


Fig. 5.1 Eigenvalues of the state transition matrix plotted on the complex plane, relative to the unit circle. The three generic types of transitions for a four-dimensional symplectic matrix are shown.

Eigenvectors

Next, given an eigenvalue of Φ, λ_i, consider its right eigenvector $\boldsymbol{u}_i$, defined by:

$$\Phi \boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i \tag{5.27}$$

Multiplying each side of the equation by Φ^{-1}, substituting the form for the inverse and simplifying yields:

$$\lambda_i^{-1} J \boldsymbol{u}_i = \Phi^T J \boldsymbol{u}_i \tag{5.28}$$

Thus, the left eigenvector of the inverse eigenvalue to λ_i is equal to the right eigenvalue of λ_i times the symplectic matrix J, or $\boldsymbol{v}_{-i} = J\boldsymbol{u}_i$, where we introduce a temporary subscript notation where $\lambda_i^{-1} = \lambda_{-i}$. Similarly the right eigenvector to λ_i^{-1}, denoted as $\boldsymbol{u}_{-i}$, serves as the generator to the left eigenvector of λ_i, or $\boldsymbol{v}_i = J\boldsymbol{u}_{-i}$.

These associated left and right eigenvectors are orthogonal to each other as $\boldsymbol{u}_i \cdot \boldsymbol{v}_{-i} = \boldsymbol{u}_i J \boldsymbol{u}_i \equiv 0$ from the skew-symmetric property of J. This can be generalized even further, borrowing from [5]. Consider an arbitrary left and right eigenvalue and eigenvector, λ_i and $\boldsymbol{u}_i$ and λ_j and $\boldsymbol{v}_j$, assuming Φ does not have repeated eigenvalues. Taking the right eigenvector definition $\Phi \boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i$, multiplying on the left by $\boldsymbol{v}_j$, and bringing all terms to one side yields the relationship:

$$(\lambda_j - \lambda_i) \boldsymbol{v}_j \cdot \boldsymbol{u}_i = 0 \tag{5.29}$$

If $i \neq j$ then $\lambda_j \neq \lambda_i$ and the left and right eigenvectors must be orthogonal to each other. For the case $i = j$ there is no explicit constraint from this result. However, in that case one knows from linear independence of the eigenvectors (for a non-repeated root) that $\boldsymbol{v}_i$ and $\boldsymbol{u}_i$ cannot be orthogonal to each other and can be properly scaled so that $\boldsymbol{v}_i \cdot \boldsymbol{u}_i = 1$. This provides the fundamental result:

$$\boldsymbol{v}_j \cdot \boldsymbol{u}_i = \delta_{ij} \tag{5.30}$$

where $\delta_{ij} = 1$ if $i = j$ and $= 0$ of $i \neq j$. This fits with the decomposition of the matrix given above, as $[U][\lambda][V]^T[U] = [U][\lambda]$ and $[V]^T[U][\lambda][V]^T = [\lambda][V]^T$.

The form is not so simple if there is a repeated eigenvalue without two linearly independent eigenvectors. For Hamiltonian systems this situation occurs generically for the state transition matrix evaluated over one period of a periodic orbit in a system which can be canonically transformed into a time-invariant system. This is discussed explicitly later when periodic orbits are considered.

5.3 Conservation Principles

5.3.1 Nonlinear Integrals

An integral of motion is any function of the state at a given epoch, and perhaps of the epoch itself, such that its value is constant as the state moves along the solution flow. Specifically, given some function $h(\boldsymbol{x}, t)$, this is an integral of motion if it is constant along the flow. For this to hold, the total time derivative of this quantity must equal zero, or

$$\frac{dh}{dt} = \frac{\partial h}{\partial \boldsymbol{x}} \cdot \dot{\boldsymbol{x}} + \frac{\partial h}{\partial t} \tag{5.31}$$

$$= \frac{\partial h}{\partial \boldsymbol{x}} \cdot J \cdot \frac{\partial H}{\partial \boldsymbol{x}} + \frac{\partial h}{\partial t} \tag{5.32}$$

$$= 0 \tag{5.33}$$

Since the first term is independent of the specific functional form in which the state is specified, assuming a canonical transformation, the time derivative can be rewritten entirely in terms of the generic Hamiltonian of the system

$$\dot{h} = \{h, H\} + \frac{\partial h}{\partial t} \tag{5.34}$$

where $\{.,.\}$ denote Poisson brackets and are defined as

$$\{h, H\} = \frac{\partial h}{\partial \boldsymbol{x}} \cdot J \cdot \frac{\partial H}{\partial \boldsymbol{x}} \tag{5.35}$$

For h to be an integral of motion the following must be true

$$\{h, H\} + \frac{\partial h}{\partial t} = 0 \tag{5.36}$$

The simplest example of this is often the Hamiltonian function itself. If a transformation can be found that places the Hamiltonian into a time invariant form, then the system in question has an integral of motion in the Hamiltonian function. This is clear as the quantity $\{H, H\}$ is a scalar and hence must be symmetric,

yet its transpose must also equal its negative as $J^T = -J$, making $\{H, H\} = 0$. Thus

$$\dot{H} = \frac{\partial H}{\partial t} \tag{5.37}$$

and only equals zero if the Hamiltonian is independent of time. If true, this is called a Jacobi integral, although we will also use the more colloquial term of "energy integral." It is not always apparent that a system has such an integral, although when it does this has important implications for the solution of the system. For instance, if the circular restricted three-body problem is posed in a Newtonian formulation it is not evident from the Hamiltonian that the system has an energy integral. It is only when a transformation into a uniformly rotating frame of the proper period is enacted that it becomes obvious. Despite this, the solutions of the Newtonian formulation have constraints on them, arising from the integral of motion, just as solutions of the rotating formulation do. It is important to note that a Jacobi integral exists for any Lagrangian system that can be transformed into a frame where the Lagrangian is time invariant. The form of this integral is just the Hamiltonian from the Legendre transformation: $\dot{\boldsymbol{q}} \cdot L_{\dot{\boldsymbol{q}}} - L$.

Another important form of a conservation principle occurs when the system can be transformed into a Hamiltonian in which some coordinates or momenta are missing. This usually only occurs for a system with a physical symmetry or one with an unreduced conservation principle. The simplest manner in which these conservation laws appear is if a specific coordinate is absent from the Hamiltonian, say q_I. Then $\dot{p}_I = -\partial H/\partial q_I = 0$, and p_I is a constant. The coordinate is not necessarily constant in this case, as H may depend on p_I. The converse case is if the Hamiltonian does not contain a certain momentum p_I, as then $\dot{q}_I = \partial H/\partial p_I = 0$, and the coordinate itself is a constant. As coordinates and momenta can always be swapped via a canonical transformation (see [56]), one can always change this situation into one where the momenta is conserved. In all cases, conservation of linear or angular momentum can generally be placed into this form.

This definition of an integral can even be applied to the Jacobi integral described above. Specifically, this tells us that if the Hamiltonian is taken as a momentum of the system, the conjugate coordinate to it will be the time. Thus, when a form of the Hamiltonian can be found that is independent of time, the Jacobi integral is recovered. The time-invariance transformation also falls into place, as in the absence of a transformation that makes the current Hamiltonian time-invariant, one can just define this time-varying quantity as the new momentum, define a new time, and trivially construct a new Hamiltonian that has a conservation law defined for it.

5.3.2 Initial Conditions as Local Integrals of Motion

Every trajectory has a full set of integrals of motion defined for it, which is just the state of the system at a specified epoch. Consider the general solution function, $\boldsymbol{x}(t) = \phi(t; \boldsymbol{x}_o, t_o)$ and its inverse $\boldsymbol{x}_o = \phi(t_o; \boldsymbol{x}, t)$. Taking the total time derivative

of $\boldsymbol{x}_o$ with respect to t yields

$$\frac{d\boldsymbol{x}_o}{dt} = \frac{\partial\phi(t_o;\boldsymbol{x}(t),t)}{\partial\boldsymbol{x}(\boldsymbol{t})}\cdot\frac{d\boldsymbol{x}}{dt} + \frac{\partial\phi(t_o;\boldsymbol{x}(t),t)}{\partial t} \tag{5.38}$$

For a time invariant system it can be shown that $\frac{\partial\phi(t_o;\boldsymbol{x}(t),t)}{\partial t} = -\dot{\boldsymbol{x}}(t_o)$, and substituting the Hamiltonian equations for the state time derivatives defines the identity

$$\Phi(t_o,t)JH_{\boldsymbol{x}}|_t - JH_{\boldsymbol{x}}|_o = 0 \tag{5.39}$$

Introducing the inverse identity $\Phi(t_o,t) = \Phi^{-1}(t,t_o) = -J\Phi(t,t_o)^TJ$ the relation can be reduced to

$$-H_{\boldsymbol{x}}|_o + \Phi^T(t,t_o)H_{\boldsymbol{x}}|_t = 0 \tag{5.40}$$

which is identically satisfied. Thus, the total time derivative of the initial conditions is identically zero and every trajectory formally has a complete set of integrals defined for it in any given state defined at an epoch. This is a fundamental result for any dynamical system. When combined with an expansion of a solution in terms of its initial conditions in the vicinity of a nominal solution, this implies that one can construct local "integrable" solutions for motion over finite timespans in the vicinity of a known solution.

However, while each set of initial conditions defines a set of integrals, the result is only local and does not translate to a global characterization in general. To do that would require the convergence of these local solutions over a large, if not global, domain. In the absence of such convergence, this result does not provide us the global characterization associated with integrable problems. Despite this limitation, such expansions and time-limited solutions are frequently sufficient for practical purposes [124].

5.3.3 Linear Integrals

The existence of an integral of motion places some additional constraints on the linearized solution flow. Assuming that $h(\boldsymbol{x},t) = C$ is an integral of the motion, note the following results and identities:

$$h(\boldsymbol{x}_o,t_o) = h(\phi(t;\boldsymbol{x}_o,t_o),t) = C \tag{5.41}$$

$$h(\boldsymbol{x}_o+\delta\boldsymbol{x},t_o) = h(\phi(t;\boldsymbol{x}_o+\delta\boldsymbol{x},t_o),t) \tag{5.42}$$

$$C + \left.\frac{\partial h}{\partial\boldsymbol{x}}\right|_{t_o}\cdot\delta\boldsymbol{x} = C + \left.\frac{\partial h}{\partial\boldsymbol{x}}\right|_{t}\cdot\Phi(t,t_o;\boldsymbol{x}_o)\cdot\delta\boldsymbol{x} \tag{5.43}$$

where higher-order terms are ignored. As this result must hold for all possible values of $\delta\boldsymbol{x}$, it follows that

$$\left.\frac{\partial h}{\partial\boldsymbol{x}}\right|_{t_o} = \left.\frac{\partial h}{\partial\boldsymbol{x}}\right|_{t}\cdot\Phi(t,t_o;\boldsymbol{x}_o) \tag{5.44}$$

The left-hand side is constant by definition, as it is evaluated at the specified epoch of the solution, and thus the right-hand side defines an associated integral of motion for the state transition matrix Φ, and was originally discussed by Poincaré [131]. This will be of particular interest when special solutions to the equations of motion are considered, such as equilibrium points and periodic orbits. It is also in similar in form to Eq. 5.40.

5.4 Equilibrium Points and Stability

Now the simplest solutions to the equations of motion are considered, those in which the dynamics at a particular point in phase space are identically zero, or $\dot{\boldsymbol{x}} = 0$. The existence of an equilibrium point often implies that the Hamiltonian function is time-invariant, although exceptions do exist, as will be seen in Chapter 12. The following will only consider time-invariant systems.

5.4.1 Form of the State Transition Matrix

Suppose that a Hamiltonian function $H(\boldsymbol{x})$ has an equilibrium point, i.e., there exists a point $\boldsymbol{x}^*$ such that $\partial H/\partial \boldsymbol{x}|_{\boldsymbol{x}^*} = 0$. Then the dynamics at this point in phase space are null, or $\dot{\boldsymbol{x}}^* = 0$, and the point ideally remains stationary, or $\phi(t; t_o, \boldsymbol{x}^*) = \boldsymbol{x}^*$. The main question then progresses to what the motion of the system is when slightly perturbed from this point in phase space. A related question is whether a given equilibrium point is stable in the sense of Lyapunov or not, and can be answered if the motion in the vicinity of the equilibrium point can be described. Formally, the solution can be characterized as $\boldsymbol{x}^* + \delta\boldsymbol{x}(t) = \phi(t; \boldsymbol{x}^* + \delta\boldsymbol{x}_o, t_o)$. Assuming that $\delta\boldsymbol{x}_o$ is an infinitesimal quantity, a Taylor Series expansion of the solution function can be performed to find $\delta\boldsymbol{x}(t) = \Phi(t, t_o) \cdot \delta\boldsymbol{x}_o$. The associated equation of motion for the state transition matrix is found to be:

$$\dot{\Phi} = J\, H_{\boldsymbol{x}\boldsymbol{x}}|_{\boldsymbol{x}^*}\, \Phi \tag{5.45}$$

where the matrix $JH_{\boldsymbol{x}\boldsymbol{x}}$ is constant as it is evaluated at the equilibrium point. Thus the solution can be expressed in closed form using the exponential matrix

$$\Phi(t, t_o) = \sum_{m=0}^{\infty} \frac{1}{m!}(t - t_o)^m [JH_{\boldsymbol{x}\boldsymbol{x}}]^m \tag{5.46}$$

$$= e^{(t-t_o)JH_{\boldsymbol{x}\boldsymbol{x}}} \tag{5.47}$$

Application of this result relies on the dynamics matrix being constant.

5.4.2 Eigenstructure of the State Transition Matrix

A few properties of the eigenvectors and eigenvalues of the state transition matrix Φ are now established. First note that if a vector $\boldsymbol{u}$ is an eigenvector of the dynamics matrix $JH_{\boldsymbol{xx}}$, then it is an eigenvector of the state transition matrix Φ. Specifically, let $JH_{\boldsymbol{xx}}\boldsymbol{u} = \sigma\boldsymbol{u}$ where σ is a complex number in general. Then, multiplying the state transition matrix by $\boldsymbol{u}$ and using the exponential equation yields

$$\Phi\boldsymbol{u} = \sum_{m=0}^{\infty} \frac{1}{m!}(t-t_o)^m [JH_{\boldsymbol{xx}}]^m \boldsymbol{u} \tag{5.48}$$

$$= \left(\sum_{m=0}^{\infty} \frac{1}{m!}(t-t_o)^m \sigma^m \right) \boldsymbol{u} \tag{5.49}$$

$$= e^{(t-t_o)\sigma}\boldsymbol{u} \tag{5.50}$$

and thus the associated eigenvalue to an eigenvector $\boldsymbol{u}$ has the form $\lambda = e^{(t-t_o)\sigma}$, linking the eigenstructure of $JH_{\boldsymbol{xx}}$ with that of Φ. From the properties of eigenvalues of a symplectic matrix, one can also infer the properties of the eigenvalues of the dynamics matrix.

Let us write $\sigma = \alpha + i\beta$, where α is the real part of the eigenvalue and β is the imaginary part, with $i = \sqrt{-1}$. Then, from the symplectic nature of Φ it is known that if $e^{(t-t_o)(\alpha+i\beta)}$ is an eigenvalue then so must its inverse, $e^{(t-t_o)(-\alpha-i\beta)}$, its complex conjugate, $e^{(t-t_o)(\alpha-i\beta)}$, and the inverse of its complex conjugate, $e^{(t-t_o)(-\alpha+i\beta)}$. Thus if $\sigma = \alpha + i\beta$ is an eigenvalue of $JH_{\boldsymbol{xx}}$ then so must $-\alpha - i\beta$, $\alpha - i\beta$, and $-\alpha + i\beta$. Note that if either α or β are zero, than the eigenvalues only come in pairs, as $\pm\alpha$ or $\pm i\beta$. For a 1-DOF system $\Phi \in \mathbf{R}^{2x2}$ and it can only have two eigenvalues, meaning that either $\alpha = 0$ or $\beta = 0$. For a 2-DOF or higher system $\Phi \in \mathbf{R}^{2n\times 2n}$, $n \geq 2$, and the more general case can occur as well. However, if the total state $2n$ is not divisible by 4, i.e., if n is odd, then at least one pair of eigenvalues must be pure real or pure imaginary.

There are, generically, three different possibilities that can occur for the eigenvalues, each of which has specific consequences for motion close to an equilibrium point. Each of these are reviewed in turn in the following.

Real Eigenvalues of $JH_{\boldsymbol{xx}}$

If σ is a real eigenvalue then there are two eigenvalues with equal magnitude and opposite signs, $\pm\alpha$. The corresponding eigenvalues of the state transition matrix are then $e^{\pm\alpha(t-t_o)}$, and there are both asymptotically stable and unstable motions in the vicinity of the equilibrium point. This is a hallmark of Hamiltonian systems and means that they cannot exhibit generic asymptotic stability. The associated eigenvectors, $\boldsymbol{u}_\pm$, are real and each eigenvalue/eigenvector pair defines a one-dimensional manifold for motion close to the equilibrium point:

$$\delta\boldsymbol{x}_\pm(t) = a\, e^{\pm\alpha(t-t_o)}\boldsymbol{u}_\pm \tag{5.51}$$

where a is an arbitrary amplitude, positive or negative, on these manifolds. The positive solution $\delta\boldsymbol{x}_+$ is unstable and will depart the equilibrium point at an exponentially increasing rate in time. This is often characterized in terms of the characteristic time of the exponential expansion, $\tau = 1/\alpha$. For example, in the Earth–Sun system the $L_{1,2}$ equilibrium points have a characteristic instability time on the order of 23 days, while in the Earth–Moon system the same points have an instability time on the order of 2 days. If the linear unstable manifold is propagated forwards in time using a numerical integration scheme, the manifold continues into a nonlinear trajectory that can stray far from the equilibrium point. Conversely, the negative solution $\delta\boldsymbol{x}_-$ is asymptotically stable as time increases and will approach the equilibrium point with the same characteristic time. To compute the nonlinear stable manifold, one starts a point on the linear stable manifold and integrates backwards in time to find where the trajectory has come from, which again can stray far from the equilibrium point. These properties of the stable and unstable manifolds, that they stray far from the equilibrium points, have been used in practical applications for transferring spacecraft to and from libration points in the Earth–Sun system and, more recently, in the Earth–Moon system. They have even been characterized as "super-highways" in the solar system, although that terminology is perhaps a bit overreaching.

Imaginary Eigenvalues of $J\boldsymbol{H}_{\boldsymbol{xx}}$

If σ is an imaginary eigenvalue then there are two eigenvalues with equal magnitudes and opposite signs, $\pm i\beta$ (where i denotes the imaginary unit in the next few paragraphs). The eigenvalues of the state transition matrix are then $e^{\pm i\beta(t-t_o)}$ and both lie on the unit circle. From Euler's formula note that the eigenvalues can also be expressed as $\cos(\beta(t-t_o)) \pm i\sin(\beta(t-t_o))$, and thus that these eigenvalues represent solutions that oscillate about the equilibrium point. This is a weak form of Lyapunov stability for a linear dynamical system, with no guarantees that when nonlinear dynamics are taken into account that the the solution will remain bounded. The question of nonlinear stability is much more difficult and few results exist. The fundamental result in this area is the Kolmogorov–Arnold–Moser Theorem [4] which places conditions on an equilibrium point, or periodic orbit, under which it can be nonlinearly stable. Conversely, Marchal has shown that a linearly stable point can at most have a nonlinear instability that is polynomial in time, and not exponential [97]. Since practical applications are generally only interested in finite time solutions, Marchal's result is taken as a motivation for accepting linear stability as practically sufficient for stability.

Since the eigenvalues are imaginary, the eigenvectors must be complex with $\boldsymbol{u}_\pm = \boldsymbol{u}_R \pm i\boldsymbol{u}_I$. Taken together, these eigenvectors define a 2-D surface around the equilibrium point filled with closed trajectories, called a center manifold. The center manifold can be expressed in terms of real numbers as

$$\delta\boldsymbol{x}_C(t) = 2a\left[\cos(\beta(t-t_o)+\varphi)\boldsymbol{u}_R - \sin(\beta(t-t_o)+\varphi)\boldsymbol{u}_I\right] \tag{5.52}$$

where a is again an arbitrary amplitude (now chosen to be positive in general) and φ is an initial phase angle. Together these two parameters trace out the two-

dimensional center manifold. This can be seen more clearly if the arbitrary phase angle is extracted to find

$$\begin{aligned}\delta\boldsymbol{x}_C(t) = {} & 2a\cos\varphi\,[\cos(\beta(t-t_o))\boldsymbol{u}_R - \sin(\beta(t-t_o))\boldsymbol{u}_I] \\ & -2a\sin\varphi\,[\sin(\beta(t-t_o))\boldsymbol{u}_R + \cos(\beta(t-t_o))\boldsymbol{u}_I]\end{aligned} \tag{5.53}$$

with the coefficients $a\cos\varphi$ and $a\sin\varphi$ serving as the two-dimensional generator of the center manifold.

The linear center manifold has a constant period of motion $2\pi/\beta$ that is independent of amplitude. In the Earth–Sun system the $L_{1,2}$ center manifolds have a period of approximately half a year, while in the Earth–Moon system they have a period of approximately 2 weeks. Nonlinear continuation of the center manifolds is more difficult than for the stable and unstable manifolds. In general, families must be computed at increasing amplitudes away from the equilibrium point using a differential correction approach. Such procedures are discussed in the following chapter when computation of periodic orbit families are covered. Using the techniques of analysis and perturbation theory it is also possible to analytically construct the center manifolds related to an equilibrium point, as has been developed rigorously to a high level of sophistication by Simo's Barcelona group (see [51]).

Complex Eigenvalues of $J\boldsymbol{H}_{\boldsymbol{xx}}$

If the dynamical system has two or more degrees of freedom, it is also possible for σ to be a complex eigenvalue with non-zero α and β. Then there are four eigenvalues associated with the system, $\pm\alpha \pm i\beta$. The eigenvalues of the state transition matrix are then $e^{\pm\alpha(t-t_o)}e^{\pm i\beta(t-t_o)}$ and trajectories consist of exponential growth and contraction in addition to oscillation, forming unstable and stable spirals emanating from the equilibrium point.

The eigenvectors are complex in general where the overbar denotes complex conjugation and with the form $\boldsymbol{u}_\pm$ and $\bar{\boldsymbol{u}}_\pm$, with the $\pm$ being associated with the sign of the real part of the eigenvalue. Each complex conjugate pair of eigenvectors define a 2-D surface around the equilibrium point filled with exponentially increasing or decreasing spirals. These manifolds can be expressed as

$$\begin{aligned}\delta\boldsymbol{x}_{\pm S}(t) = {} & 2a\,e^{\pm\alpha(t-t_o)}\,[\cos(\beta(t-t_o)+\varphi)\,(\boldsymbol{u}_\pm + \bar{\boldsymbol{u}}_\pm) \\ & +i\sin(\beta(t-t_o)+\varphi)\,(\boldsymbol{u}_\pm - \bar{\boldsymbol{u}}_\pm)]\end{aligned} \tag{5.54}$$

where again a is an arbitrary positive amplitude and φ is an initial phase angle. Each spiral traces out a two-dimensional manifold, creating a four-dimensional manifold overall when both are considered.

5.4.3 General Motion in the Vicinity of an Equilibrium Point

The generic motion in the vicinity of an equilibrium point will have components that lie on all of the available manifolds in the vicinity of the equilibrium point.

For definiteness, let us consider a 2-DOF system with a pair of stable/unstable eigenvalues and a center manifold. A general expression for motion relative to the equilibrium point is specified as

$$\begin{aligned}\delta\boldsymbol{x}(t) = a_- e^{-\alpha t}\boldsymbol{u}_- + a_+ e^{\alpha t}\boldsymbol{u}_+ + 2a_{cc}\left[\cos(\beta t)\boldsymbol{u}_R - \sin(\beta t)\boldsymbol{u}_I\right] \\ -2a_{sc}\left[\sin(\beta t)\boldsymbol{u}_R + \cos(\beta t)\boldsymbol{u}_I\right]\end{aligned} \tag{5.55}$$

As all of these eigenvalues are distinct, note the existence of a left eigenvector $\boldsymbol{v}_i$ for each eigenvalue with the property that $\boldsymbol{v}_i \cdot \boldsymbol{u}_j = \delta_{ij}$. Thus, these eigenvectors can be used to select what component of an arbitrary initial vector, $\delta\boldsymbol{x}_o$ lies on each of the manifolds. Specifically, $a_- = \boldsymbol{v}_- \cdot \delta\boldsymbol{x}_o$ and $a_+ = \boldsymbol{v}_+ \cdot \delta\boldsymbol{x}_o$. Projections onto the center manifold become more algebraically complex, and thus are more simply stated as $\boldsymbol{v}_c \cdot \delta\boldsymbol{x}_o$ and $\bar{\boldsymbol{v}}_c \cdot \delta\boldsymbol{x}_o$. Introduction of these projections allows us to take the state transition matrix definition full circle, and rewrite the motion in the vicinity of the equilibrium point in terms of the eigenstructure of that point

$$\begin{aligned}\delta\boldsymbol{x}(t) = \big[e^{-\alpha t}\boldsymbol{u}_-\boldsymbol{v}_- + e^{\alpha t}\boldsymbol{u}_+\boldsymbol{v}_+ + 2\cos(\beta t)\left(\boldsymbol{u}_R\boldsymbol{v}_R - \boldsymbol{u}_I\boldsymbol{v}_I\right) \\ -2\sin(\beta t)\left(\boldsymbol{u}_R\boldsymbol{v}_I + \boldsymbol{u}_I\boldsymbol{v}_R\right)\big] \cdot \delta\boldsymbol{x}_o\end{aligned} \tag{5.56}$$

where the products of eigenvectors are all outer, or dyad, products. This formula provides an explicit decomposition of the state transition matrix into its fundamental components and is equivalent to Eq. 5.22.

5.4.4 Constraints Due to Integrals

For the particular case of a time-invariant system and an equilibrium point, the constraint on the state transition matrix shown in Eq. 5.44 does not apply. To understand this, note that the partial of the integral, in this case $H(\boldsymbol{x})$, is identically zero when evaluated at the point $\boldsymbol{x}^*$. Thus, there is no constraint on the matrix Φ in Eq. 5.44.

If the system has an additional integral of motion, denoted as $h(\boldsymbol{x}) = C$, where C is a constant and h a scalar function (assumed time-invariant), there are additional constraints. Note the usual relationship $\dot{h} = \{h, H\} = 0$. If h is not the Hamiltonian and is a linearly independent integral, its gradient will in general not equal zero. Thus in this case the neighborhood of the equilibrium point $\boldsymbol{x}^*$ can be evaluated to find the result:

$$\left.\frac{\partial h}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}^*} \cdot \left[I - \Phi(t - t_o; \boldsymbol{x}^*)\right] = 0 \tag{5.57}$$

where the fact that the value of $h_{\boldsymbol{x}}$ is constant when evaluated at an equilibrium point is used. From this one can immediately note that Φ has a unity eigenvalue, and that the quantity $h_{\boldsymbol{x}}$ is a left-eigenvector of the matrix. This implies the presence of a second unity eigenvalue for the matrix with the right-eigenvector $Jh_{\boldsymbol{x}}$. Thus, for every additional integral of motion for a given dynamical system beyond the Hamiltonian, there exist two unity eigenvalues for the state transition matrix

evaluated in the vicinity of the equilibrium point. The presence of such additional integrals of motion can complicate the numerical search for equilibrium points, as discussed in Section 6.2, due mainly to the fact that they imply that the dynamics matrix JH_{xx} has a pair of zero eigenvalues.

5.5 Periodic Orbits and Stability

5.5.1 Definition of a Periodic Orbit

Assume a Hamiltonian system with state $\boldsymbol{x} = [q, p]$, Hamiltonian function $H(\boldsymbol{x}, t)$, and attendant equations of motion $\dot{\boldsymbol{x}} = JH_{\boldsymbol{x}}$ with solutions $\boldsymbol{x}(t) = \phi(t; \boldsymbol{x}_o, t_o)$. A solution to this set of equations is said to be periodic with period T if the state of the system exactly repeats itself after a timespan T and the equations of motion evaluated at this state repeat themselves. As with all trajectories, a periodic orbit can be uniquely defined by an initial state, here denoted as $\boldsymbol{x}_t^*$, where the t subscript indicates that it can be defined at any point along the trajectory. Thus the condition for a periodic orbit is

$$\boldsymbol{x}_t^* = \phi(t + T; \boldsymbol{x}_t^*, t) \tag{5.58}$$

$$JH_{\boldsymbol{x}}(\boldsymbol{x}_t^*, t + T) = JH_{\boldsymbol{x}}(\boldsymbol{x}_t^*, t) \tag{5.59}$$

and it can be immediately seen that for a time invariant Hamiltonian a sufficient condition for a solution to be periodic is that it repeats itself, as then the dynamics will also repeat. An equilibrium point is a special case of a periodic solution where the period of motion equals zero. For a time-varying Hamiltonian there is a more stringent condition on the dynamical system itself, namely that the gradient of the Hamiltonian repeat itself over the same timespan T. This is usually relaxed to the stipulation that the Hamiltonian itself is time periodic, meaning that there exists some constant time T such that $H(\boldsymbol{x}, t + T) = H(\boldsymbol{x}, t)$, for all time t.

For time-varying systems that do not repeat themselves, i.e., there exists no time interval T such that $H(\boldsymbol{x}, t + T) = H(\boldsymbol{x}, t)$ for all t, such systems cannot have periodic solutions in general. This is the situation for "real world" motion, as the orbits of the planets never repeat themselves exactly, etc. Still, the concept of periodic orbits in these systems is important, as planetary systems can generally be reduced to simplified models that are periodic or time invariant. If these simplified systems are "close" to the real systems, the influence of the periodic orbits often persists in the perturbed problem, meaning that trajectories in the real systems still feel the influence of the periodic orbits and their stability properties, even if those solutions do not exist in the equations at hand. The simplest way to think of their influence is that the periodic solutions exist in a lower-order model of the system, and that the non-periodic terms represent higher-order perturbations that destroy the periodicity, yet which may not entirely eliminate the influence of the first-order periodic solutions.

A motivating interest in periodic solutions is that they are defined for all time. Specifically, once $\boldsymbol{x}(t + T) = \boldsymbol{x}(t)$ is established, this solution can be generalized to

$\boldsymbol{x}(t+mT) = \boldsymbol{x}(t)$ for $m = \pm 1, \pm 2, \ldots$. Thus, knowledge of the solution over a time interval $t \in [0, T)$ is sufficient information to know the solution for all time. Except for integrable systems, this is generally the only time one can have such complete knowledge of the solution of the equations of motion.

As should be apparent from the conditions for periodicity, there is a fundamental difference between periodic orbits in time-invariant systems and in time periodic systems. For time invariant systems there are no *a priori* constraints on what possible periods of the solutions there may be. Indeed, for time-invariant systems there are usually connected families of periodic solutions that have periods which continuously cover some range of values. For every periodic orbit with a given period, there will in general always exist a continuous family of periodic orbits with different periods in its neighborhood.

For time-periodic systems, there can only be periodic orbits with periods that are multiples of the smallest period of the Hamiltonian. Thus, periodic orbits are isolated in terms of period, and as will be shown, are also isolated in initial conditions.

5.5.2 Floquet Theory

The motion in the vicinity of a periodic orbit in phase space has a special structure associated with it. In particular, while the state transition matrix expanded about a periodic orbit cannot be expressed in closed form in general (unlike for an equilibrium point), the functional form of its general solution can be found. Floquet's theorem provides such an explicit functional form for the state transition matrix relative to a periodic orbit.

The state transition matrix follows the equation:

$$\dot{\Phi}(t, t_o) = JH_{\boldsymbol{xx}}(\boldsymbol{x}(t))\Phi(t, t_o) \tag{5.60}$$

where $\boldsymbol{x}(t) = \phi(t + T; \boldsymbol{x}_o, t_o)$. Since the nominal solution is periodic with an assumed period of T, the dynamics matrix $JH_{\boldsymbol{xx}}$ is periodic with the same period. Based on this additional constraint Floquet's Theorem can be proved. The simplest statement of the theorem is that the state transition matrix evaluated relative to a periodic orbit has the standard form:

$$\Phi(t, t_o) = P(t, t_o)\, e^{M(t-t_o)} \tag{5.61}$$

where P is a non-singular, periodic matrix function of time with period T and M is a constant matrix.

Consider the state transition matrix solution over one period T, $\Phi(T)$ (where we will take $t_o = 0$ and stop carrying this term). The matrix $\Phi(T)$ is a symplectic matrix and thus its eigenvalues follow the fundamental rules for these systems. The state transition matrix of a periodic orbit taken over one period will always have two eigenvalues equal to 1, and cannot be reduced to a diagonal matrix of eigenvalues. This will be considered in more detail later, yet this fact is addressed in the following derivation. The remaining part of the matrix can be decomposed

into its left and right eigenvectors and its eigenvalues. The unity eigenvalues can at least be isolated into a Jordan block form, and will have associated left and right eigenvectors and generalized eigenvectors. This allows us to generically write the state transition matrix into a decomposed form:

$$\Phi(T) = U\Lambda V^T \tag{5.62}$$

where V is the matrix of left eigenvectors and generalized eigenvectors, U is the matrix of right eigenvectors and generalized eigenvectors, and Λ is a matrix in Jordan form with distinct eigenvalues of the state transition matrix along the diagonals and the one Jordan block for the repeated unity eigenvalues (for a time-invariant system), with the structure

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \tag{5.63}$$

The matrix Λ can be written in exponential form as $\Lambda = e^{[\alpha]T}$. The matrix $[\alpha]$ is again in Jordan form with diagonal entries associated with all the non-repeated eigvenvalues, equal to $\alpha_i = \ln(\lambda_i)/T$, and one Jordan block with zeros along the diagonal and one above the diagonal. Thus, the state transition matrix can be written as

$$\Phi(T) = U\, e^{[\alpha]T} V^T \tag{5.64}$$

This decomposition of the state transition matrix is used to motivate a new matrix function $\psi(t)$ defined as

$$\psi(t) = U\, e^{[\alpha]t} V^T \tag{5.65}$$

Note that $\psi(t)$ is the matrix solution to the time invariant matrix equation

$$\dot{\psi} = U[\alpha]V^T\psi \tag{5.66}$$

To show this recall the property of the left and right eigenvectors, namely that $V^TU = U^TV = I$. As the solution to a time invariant matrix equation, the function $\psi(t)$ has a few key properties. Of specific interest to us is the inverse relation $\psi^{-1}(t) = \psi(-t)$.

Using this new function $\psi(t)$, a decomposition of the state transition matrix can be defined

$$\Phi(t, t_o) = P(t, t_o)\psi(t - t_o) \tag{5.67}$$

or conversely a new function $P(t, t_o)$ is defined by:

$$P(t, t_o) = \Phi(t, t_o)\psi^{-1}(t - t_o) \tag{5.68}$$

For this definition it can be shown that P is a periodic function of time with period T. The proof is simple. First, note that, by definition, $P(0, 0) = I$, and furthermore

that $P(T,0) = I$, the latter due to the definition of ψ in Eq. 5.65. Next, consider the equation of motion of P,

$$\begin{aligned}\dot{P}(t,t_o) = &JH_{\boldsymbol{xx}}(t)\Phi(t,t_o)\psi^{-1}(t-t_o) \\ &-\Phi(t,t_o)\psi^{-1}(t-t_o)U[\alpha]V^T\end{aligned} \tag{5.69}$$

To establish periodicity, it is just needed to show that $\dot{P}(0,0) = \dot{P}(T,0)$, as then the solution P must be periodic since its value is equal at these two epochs. Substituting these time limits into the above expression yields the following:

$$\dot{P}(0,0) = JH_{\boldsymbol{xx}}(0) - U[\alpha]V^T \tag{5.70}$$

$$\dot{P}(T,0) = JH_{\boldsymbol{xx}}(T)\Phi(T,0)\psi^{-1}(T) - \Phi(T,0)\psi^{-1}(T)U[\alpha]V^T \tag{5.71}$$

but $H_{\boldsymbol{xx}}(T) = H_{\boldsymbol{xx}}(0)$ along a periodic orbit and $\Phi(T,0)\psi^{-1}(T) = I$, making the two expressions equal and finishing the proof.

In final form, note that the solution to the state transition matrix in the vicinity of a periodic orbit has the general form:

$$\Phi(t,t_o) = P(t,t_o)U\, e^{[\alpha](t-t_o)}V^T \tag{5.72}$$

where U and V^T are the right and left eigenvector (and generalized eigenvector) matrices of $\Phi(T)$, respectively, and $[\alpha] = \ln(\Lambda)/T$.

5.5.3 Stability of Periodic Orbits

A particular application of Floquet's Theorem is to the stability of a periodic orbit, as their stability properties provide an immediate description of the phase space flow in the vicinity of that solution. Specifically, if an orbit is stable this implies that phase flows in the vicinity of the orbit will stay close to the periodic orbit. Conversely, if an orbit is unstable, this implies that there are pathways that flow from close to the solution to "far" from the solution, and allows for the phase space to be traversed to either approach or depart from the periodic orbit.

Stability analysis is made particularly simple by using the Floquet decomposition. First, consider the propagation of the state transition matrix to an arbitrary epoch in the future, denoted as $t + mT$, assuming an initial epoch of $t_o = 0$:

$$\Phi(t+mT) = P(t)U\, e^{[\alpha](t+mT)}V^T \tag{5.73}$$

$$= P(t)U\, e^{[\alpha]t}V^T U\, e^{[\alpha]mT}V^T \tag{5.74}$$

$$= \Phi(t)U\Lambda^m V^T \tag{5.75}$$

$$= \Phi(t)\Phi(T)^m \tag{5.76}$$

The stability of the solution is clearly dominated by the eigenvalues of the state transition matrix evaluated over one orbit period. If all of the eigenvalues Λ have unity magnitude, i.e., are of the form $\lambda_j = e^{i\theta_j}$, then the matrix Λ^m will not grow under iteration and the orbit is considered to be stable. Note that this does not

constitute nonlinear stability, a much more difficult topic for which few sharp results exist, the most important being the Kolgomorov–Arnold–Moser Theorem [4]. In general, stability does imply lack of hyperbolic instability, with any instabilities that may exist only being polynomial in time at most [97]. Conversely, the existence of one eigenvalue with a magnitude not equal to unity implies instability, due to the existence of its inverse pair with magnitude less than or greater than unity, respectively.

Similar to equilibrium points, periodic orbits will also have stable, unstable and center manifolds associated with them, depending on the stability of the system. The fundamental stability result for a periodic orbit reduces to the computation of the eigenvalues and eigenvectors of the state transition matrix over one orbit period, $\Phi(T)$, which can be represented again as a set of eigenvalues $\lambda_i = e^{\alpha_i T}$ and corresponding eigenvectors, $\boldsymbol{u}_i$. While the eigenvalues of the state transition matrix, when mapped over one orbit period, are invariant with where they start within one periodic orbit, the eigenvectors will, in general, depend on where they start within one the periodic orbit. Specifically, the eigenvectors for $\Phi(T,0)\boldsymbol{u}(0) = \lambda\boldsymbol{u}(0)$ and $\Phi(T+t,t)\boldsymbol{u}(t) = \lambda\boldsymbol{u}(t)$ are different, even though the eigenvalues are the same. This arises from the relationship

$$\Phi(t+T,t) = \Phi(t,0)\Phi(T,0)\Phi^{-1}(t,0) \tag{5.77}$$

thus leading to the relationship $\boldsymbol{u}(t) = \Phi(t,0)\boldsymbol{u}(0)$ [58]. Thus, after every orbit period the eigenvectors repeat their direction, and expand or contract according to whether $|\lambda|$ is greater or less than one. For an eigenvector associated with an eigenvalue with unity magnitude, $\lambda = e^{i\theta T}$, the eigenvector will in general be rotated by a total net angle θ over one orbit period, tracing out lines on a torus surrounding the periodic orbit.

The analogy between manifolds of equilibrium points and of periodic orbits have many similarities, most of which are clearly exposed when the concept of surface of section and Poincaré map is introduced. The main difference is the dimension of the manifolds, however. For the stable and unstable eigenvalues (i.e., when λ is real and either greater or less than unity) the manifolds are now two-dimensional surfaces, essentially one-dimensional manifolds emanating from every point along a closed periodic orbit and thus forming a two-dimensional sheet. The center manifold, for a generic periodic orbit, consists of sets of quasi-periodic tori that surround the periodic orbit, forming a two-dimensional surface for a given offset amplitude. Thus, as the amplitude is varied these define three-dimensional objects. All of these manifolds can also be extended beyond the linear regime, as described previously for the equilibrium point manifolds. The procedure for extending these manifolds becomes more difficult, due to their higher dimensions. This problem has been extensively studied in the literature, using numerical techniques [118], semi-analytical techniques [109] and analytical techniques [40, 51].

5.5.4 Unity Eigenvalues for Time Invariant Systems

For periodic orbits in time-invariant systems, i.e., systems that conserve their Hamiltonians, there are always a pair of unity eigenvalues in addition to the eigenvalues discussed above, and were already alluded to in the section on Floquet Theory.

Assume that we are studying a periodic orbit in a time-invariant system, thus the Hamiltonian is a constant of motion. Again consider Eq. 5.44, now applied to the Hamiltonian of the system. Due to the periodicity of the underlying motion the partial $H_{\boldsymbol{x}}(0) = H_{\boldsymbol{x}}(T)$ and is non-zero, leading to:

$$H_{\boldsymbol{x}}(0)\left[I - \Phi(T,0)\right] = 0 \tag{5.78}$$

meaning that Φ has a unity eigenvalue and that $H_{\boldsymbol{x}}$ is a left eigenvector. By symmetry there is a second unity eigenvalue with $JH_{\boldsymbol{x}}$ as its right eigenvector. This makes sense, as a displacement along the dynamics of the initial point should just correspond to a point displaced along the periodic orbit, or $\delta\boldsymbol{x} = \dot{\boldsymbol{x}}\delta t = JH_{\boldsymbol{x}}\delta t$. Since the orbit repeats itself, this displaced point will repeat itself with the same period. The presence of these unity eigenvalues for periodic orbits in time-invariant systems complicate the solution process for finding periodic orbits.

Whenever a repeating eigenvalue is found for a matrix, it is important to understand whether the associated eigenvectors with these repeated roots are linearly independent. To evaluate this the concept of a generalized eigenvector must be introduced, discussed in some detail in [58] for astrodynamical systems. A generalized eigenvector of grade k for an eigenvalue with multiplicity m satisfies

$$\left(\Phi - \lambda I\right)^k \boldsymbol{u}_k = 0 \tag{5.79}$$

$$\left(\Phi - \lambda I\right)^{k-1} \boldsymbol{u}_k \neq 0 \tag{5.80}$$

For a periodic orbit with repeated unity eigenvalues, one can always find the generalized eigenvector of rank 1, $JH_{\boldsymbol{x}}$. Thus, the generalized eigenvector of rank 2 must satisfy

$$\left(\Phi - I\right)\boldsymbol{u}_2 = JH_{\boldsymbol{x}} \tag{5.81}$$

as $(\Phi - I)JH_{\boldsymbol{x}} = 0$. Multiplying by $\Phi^{-1} = -J\Phi^T J$ and noting that $H_{\boldsymbol{x}}$ is a left eigenvector of Φ yields

$$\left[I - \Phi^{-1}\right]\boldsymbol{u}_2 = JH_{\boldsymbol{x}} \tag{5.82}$$

Thus, subtracting the two leads to the condition that

$$\left[\frac{1}{2}\left(\Phi + \Phi^{-1}\right) - I\right]\boldsymbol{u}_2 = 0 \tag{5.83}$$

Thus, for a non-zero generalized eigenvector of rank 2 to exist the matrix $\frac{1}{2}(\Phi + \Phi^{-1}) - I$ must be singular, and $\boldsymbol{u}_2$ must lie in its null space. This matrix can be

shown to always be singular by choosing $\boldsymbol{u}_2 = JH_{\boldsymbol{x}}$ and noting that this is also a right eigenvector of Φ^{-1}, which can be shown from the symplectic property. Then $\left[\frac{1}{2}(\Phi + \Phi^{-1}) - I\right] JH_{\boldsymbol{x}} = JH_{\boldsymbol{x}} - JH_{\boldsymbol{x}} \equiv 0$, proving that the matrix must have at least one zero eigenvalue when $JH_{\boldsymbol{x}}$ is a unity eigenvector, which always occurs for time-invariant systems evaluated at a periodic orbit. Since Φ and Φ^{-1} have the same set of eigenvalues, due to the inverse property of their eigenvalues, this implies that there exists a second zero eigenvalue of this matrix as well. This generalized eigenvector must include a component along the gradient of the Hamiltonian, $H_{\boldsymbol{x}}$, however in general it also includes components along the other eigenvectors of the system [58]. That it must contain a component along $H_{\boldsymbol{x}}$ arises from the fact that this vector is orthogonal to all other eigenvectors, including $JH_{\boldsymbol{x}}$. Thus, since the remaining generalized eigenvector must complete the linear space, it must also contain a component of $H_{\boldsymbol{x}}$.

For a time-varying system, such a displacement along the orbit without modifying the time changes the dynamics and yields a solution in the vicinity of the periodic orbit and not on the periodic orbit. Thus, periodic orbits in time-periodic systems, which do not conserve the Hamiltonian, generally do not have unity eigenvalues. If they do, these only occur at specific stability bifurcation points and do not arise generically.

5.5.5 Periodic Orbit Families

A fundamental aspect of periodic orbits in time-invariant systems is that they can be continued into families by varying a parameter such as the Jacobi constant or a parameter of the dynamical system. The discussion of such families is postponed until the concept of a surface of section and Poincaré map are introduced. Once these ideas are presented, the discussion for continuation of a family becomes clearer.

5.6 General Trajectories and Stability

The two most important classes of solutions have now been discussed, equilibrium points and periodic orbits. There are important higher-dimensional generalizations of periodic orbits called quasi-periodic orbits which are an important class of solution, but are in general difficult to compute and characterize. We do not consider these classes of solutions, although they will be of future interest to the study of small body orbiters. The importance of solutions such as equilibria, periodic orbits, and quasi-periodic orbits is mainly attached to the fact that they can be specified for all time and that their stability can be unambiguously computed. Additionally, through their stable and unstable manifolds they influence wider regions of phase space. Despite all of these things, the number of such special trajectories are vanishingly small with respect to more general trajectories without such special initial conditions. Thus, it is relevant to discuss these other solutions. However, due to our not having a clear expression of these trajectories valid over all time (an

exception being for integrable systems), there are only limited things that can be stated about such trajectories.

First, recall our notation for a solution, $\boldsymbol{x}(t) = \phi(t; \boldsymbol{x}_o, t_o)$. In some special cases, namely if the state $\boldsymbol{x}_o$ lies on a manifold to an equilibrium point or a periodic orbit, then the solution may have a well-defined limit. If it lies on a stable manifold, then the trajectory is defined as $t \to \infty$, if on an unstable manifold it is defined as $t \to -\infty$. It is significant to note that the dimension of an object's manifolds will always be greater than the object in question itself. Thus, one can immediately note the outsized influence which equilibria and periodic orbits have.

If one moves away from these special cases, or has a trajectory that moves away from a special orbit either forwards or backwards in time, this leads to a general solution without a known limiting structure. The solution may still be subject to other constraints, such as having to lie on an energy surface or remain within a zero-velocity surface, but it is not necessary to assume these phase space constraints in general. Instead, all one knows is that the solution will either continue to evolve over time or may run into a physical constraint, such as impacting the surface of a small body, a planet, or the Sun. Beyond the specific path of the trajectory, which can only be ascertained by some solution procedure, the other main item of interest is determining how neighboring trajectories evolve. In fact, from a practical point of view, knowing how neighboring trajectories evolve can be more helpful in understanding a general trajectory than where the trajectory itself goes. This is primarily due to the fact that spacecraft locations in phase space are never known precisely, and thus knowledge of how the local phase volume evolves informs us as to whether neighboring solutions will remain close to each other or diverge.

The standard approach to characterizing the local dynamics about a nominal trajectory is via Lyapunov Characteristic Exponents (LCE). These are, in some sense, a generalization of the characteristic exponents found for an equilibrium point and a periodic orbit. Instead of measuring how rapidly a neighboring trajectory expands along a specified direction, an LCE instead measures the maximum rate of expansion in a region of the neighboring phase space. There are several techniques that are used to compute LCEs, or to ascertain specific properties related to them. The following is the simplest and most direct definition.

First, assume a nominal trajectory defined by an epoch state and time, $\boldsymbol{x}_o, t_o$. Linearizing about the nominal trajectory yields the state transition matrix, which takes a small variation $\delta\boldsymbol{x}_o$ into a later variation relative to the nominal trajectory: $\delta\boldsymbol{x}(t) = \Phi(t, t_o)\delta\boldsymbol{x}_o$. In principle it is always possible to choose the norm $|\delta\boldsymbol{x}_o|$ small enough to satisfy linearity assumptions that affect the deviation of $\delta\boldsymbol{x}(t)$ for any timespan $t - t_o$. Given these steps the finite time LCE is computed as

$$\chi(t) = \max_{\delta\boldsymbol{x}_o} \frac{1}{t} \ln\left(\frac{|\Phi(t, t_o)\delta\boldsymbol{x}_o|}{|\delta\boldsymbol{x}_o|}\right) \tag{5.84}$$

It is instructive to look at this function more closely.

As the logarithm is a definite function, its maximum value will occur at the maximum value of its argument. This can be rewritten by squaring the norm and dividing through by the magnitude of $\delta\boldsymbol{x}_o$ and is the same as finding $\max_{|\boldsymbol{u}|=1} \boldsymbol{u} \cdot$

$\Phi^T\Phi \cdot \boldsymbol{u}$. To solve this problem, write out a constrained Lagrangian function to find the extremal values, $\Xi = \boldsymbol{u} \cdot \Phi^T\Phi \cdot \boldsymbol{u} - \xi(\boldsymbol{u} \cdot \boldsymbol{u} - 1)$, where Ξ is a scalar function and ξ is a Lagrange multiplier. Extrema are found by solving the equations $\frac{\partial \Xi}{\partial \boldsymbol{u}} = 0$ and $\frac{\partial \Xi}{\partial \xi} = 0$. The second equation gives us $\boldsymbol{u} \cdot \boldsymbol{u} = 1$ while the first gives us $\left[\Phi^T\Phi - \xi I\right] \boldsymbol{u} = 0$. Thus, one only needs to find eigenvalues and eigenvectors of the matrix $\Phi^T\Phi$ in order to compute the extremal values. The eigenvalues of $\Phi^T\Phi$ can be computed from the eigenvalues of Φ, denoted here as $\lambda_i, i = 1, 2, \ldots, 2n$ and have the usual structure for symplectic matrices. Specifically, the eigenvalues of $\Phi^T\Phi$ equal $\xi_i = \lambda_i\bar{\lambda}_i, i = 1, 2, \ldots, 2n$. Thus, there may exist repeated unity eigenvalues if $\lambda_i = e^{i\theta}$, but these will not be defective in general and thus will have linearly independent eigenvectors (see [100] for a discussion of the general properties of these eigenvalues and eigenvectors). Evaluating the cost function Ξ at a general extremum yields $\Xi = \lambda_i\bar{\lambda}_i, i = 1, 2, \ldots, 2n$ and thus that the maximum occurs for $\max_{i\in[1,2n]} \lambda_i\bar{\lambda}_i$. Carrying this back to the original definition of $\chi(t)$,

$$\chi(t) = \frac{1}{t} \ln \max_i \sqrt{\lambda_i\bar{\lambda}_i} \tag{5.85}$$

It can shown in general that the limit of this quantity converges to a constant value as $t \to \infty$, and that it can be considered to be an integral of motion for any particular trajectory [119]. This provides the formal definition of an LCE:

$$\chi_\infty = \lim_{t\to\infty} \chi(t) \tag{5.86}$$

If $\chi_\infty > 0$, the trajectory is considered to be unstable as neighboring trajectories will deviate exponentially from the given trajectory. Conversely, if $\chi_\infty = 0$ this implies that neighboring solutions are bounded, in a linear sense, although it does not guarantee nonlinear boundedness. The practical computation of the LCE is difficult, it being impossible to determine a general trajectory for all time. There are a wealth of practical computation schemes and related definitions which are discussed in the reference [42].

It is instructive to apply this definition to an equilibrium point. For the state transition matrix evaluated at an equilibrium point, the maximum eigenvalue $\lambda_i\bar{\lambda}_i$ will have the generic value $e^{2\alpha t}$ where α is the largest real part of the eigenvalues of $JH_{\boldsymbol{xx}}$ evaluated at the equilibrium point. Thus, it can be immediately seen that $\chi_\infty = \alpha$, which fits precisely with one's expectations. In particular, if the equilibrium point is stable then all of the eigenvalues satisfy $\lambda_i\bar{\lambda}_i = 1$ and $\chi(t) \equiv 0$. Similar results are found if applied to a periodic orbit.

5.7 Surfaces of Section and Poincaré Maps

Finally, the important concepts of surface of section maps and Poincaré maps are introduced. Although these two are usually related to each other, it instructive to separate our discussion of them. Following their appropriate definition families of periodic orbits and additional properties of these solutions can be discussed.

5.7.1 Surface of Section Maps

For time-invariant dynamical systems it is often convenient to reduce the dimensionality of the system via a "surface of section." A surface of section is some geometrically defined surface or condition in phase space which, when a trajectory crosses this surface or passes through this condition, the precise point of intersection is obtained and reported in some coordinate frame. Usually the trajectory must cross the surface or condition in a proscribed direction. The simplest example of a surface of section is some coordinate axis in the configuration space, say the x-axis. Then, whenever a trajectory crosses through the value $y = 0$ the values of all of its other coordinates and momenta (or velocities) are reported. For a simple orbit traveling about the origin in the x–y plane there should be at least two x-crossings per orbit – however, by restricting the surface of section to only be reported (to only exist) when the trajectory crosses in a certain direction, say for $\dot{y} > 0$, then there could only be one surface of section crossing per orbit. A surface of section need not be defined only in the configuration space, although these are often the most convenient and commonly used surfaces. For example, periapsis passages can be used as a surface of section, defined as the condition $\dot{r} = \boldsymbol{r} \cdot \dot{\boldsymbol{r}}/r = 0$ with the additional condition that $\ddot{r} > 0$.

To set this up in a more generic fashion, let the surface of section be defined by a scalar equation $S(\boldsymbol{x}) = 0$, with $\boldsymbol{x} \in \mathbf{R}^{2n}$ (see Fig. 5.2). In the above examples this function would equal x_2 or $\dot{\boldsymbol{r}} \cdot \boldsymbol{r}/r$. To be well-defined, the solution flow to be observed using the surface of section must be transversal to the surface of section, meaning that the trajectory is non-tangent to the surface. This can be enforced by requiring $\frac{\partial S}{\partial \boldsymbol{x}} \cdot \dot{\boldsymbol{x}} > 0$, where a preferred direction of crossing has been defined. The solution flow relative to a well-defined surface of section will repeatedly cross through the surface after each orbit and will not have a grazing, or tangential, intersection with it. A given trajectory being transverse to the surface of section at one crossing is no guarantee that it will be transversal to the surface at a future crossing, or even that it will ever cross the surface again. It is possible for a trajectory to initially have transversal crossings of a surface of section yet later evolve to the point where there are grazing, or perhaps no, additional crossings. An example of the latter is if a trajectory with a periapsis surface of section escapes from the system in question – it will never have another periapsis crossing and hence will never appear on the surface again. Due to this, the choice of surface of section is often very important in analyzing a dynamical system, as a poor choice may lead to no intersections. Thus, the selection of surfaces of section is highly dependent on the problem and the particular type of flow that is being investigated.

Surfaces of section are usually only set up in time invariant dynamical systems, as then, whenever a trajectory crosses through the same surface of section coordinates again, the motion, by definition, defines a periodic orbit. For a time-varying dynamical system, the surface of section would have time-varying dynamics across its surface and does not yield any systematic simplification. It is interesting to consider a time periodic dynamical system, i.e, one with a Hamiltonian where $H(\boldsymbol{x}, t) = H(\boldsymbol{x}, t + T)$, that is transformed into its time invariant form with its attendant increase in dimensions. For such a time invariant system the natural surface

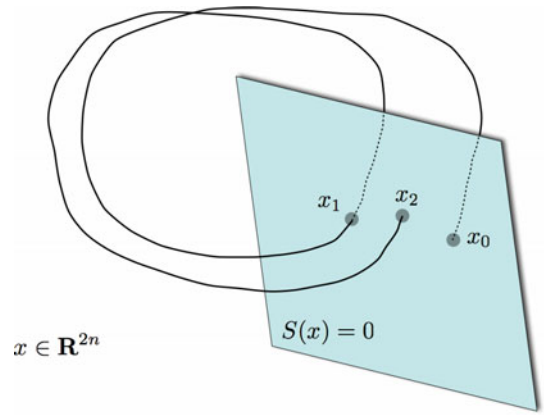


Fig. 5.2 Illustration of a surface of section.

of section to choose is the time coordinate. Such a surface would consist of reporting the dynamics whenever the time coordinate equals $t_o + nT$, $n = 0, \pm 1, \pm 2, \ldots$. By definition, the solution flow will always be transverse to this surface of section. This particular application of a surface of section is called a stroboscopic map and has been widely applied to time-periodic systems.

Let us take the periapsis passage in a two-body orbit as an example, using a Lagrangian formulation. Then $S(\boldsymbol{x}) = \dot{r} = \frac{\boldsymbol{r}\cdot\dot{\boldsymbol{r}}}{r}$ and $S_{\boldsymbol{x}} = \left[\frac{\dot{\boldsymbol{r}}}{r}; \frac{\boldsymbol{r}}{r}\right]$, where application of the result $\dot{r} = 0$ simplifies the partial with respect to position. The transversality condition then becomes $S_{\boldsymbol{x}} \cdot \dot{\boldsymbol{x}} = \frac{\dot{\boldsymbol{r}}\cdot\dot{\boldsymbol{r}}}{r} + \frac{\boldsymbol{r}\cdot\ddot{\boldsymbol{r}}}{r} > 0$. If the acceleration in the two-body problem is substituted as an example, $\ddot{\boldsymbol{r}} = -\mu\boldsymbol{r}/r^3$, this leads to $S_{\boldsymbol{x}} \cdot \dot{\boldsymbol{x}} = \frac{1}{r}\left[\dot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}} - \frac{\mu}{r}\right]$. Thus the surface is only crossed when this is greater than zero, implying that the local speed is greater than local circular speed, a prerequisite to have a periapsis passage.

The main advantage of the surface of section is that it allows us to replace a continuous time solution with a discrete map, effectively removing the time from the system. Reducing the trajectory to these surface crossings allows us to define a mapping function $G(\boldsymbol{x})$. Assume an initial point on the surface of section is chosen, $\boldsymbol{x}_0$, at a time t_0, such that $S(\boldsymbol{x}_0) = 0$ and the transversality condition is satisfied. Then the mapping to the next surface crossing can be represented as $\boldsymbol{x}_1 = G(\boldsymbol{x}_0) = \left\{\boldsymbol{x}(t_1) | S(\phi(t_1, t_0, \boldsymbol{x}_0)) = 0,\, S_{\boldsymbol{x}} \cdot \dot{\boldsymbol{x}}|_{t_1} > 0, t_1 > t_0\right\}$. This mapping implicitly defines a passage time t_1, constrained in general to be the first passage of the surface of section after t_0, although the exact value of this time is not of specific interest in terms of the surface of section map. Repeating this procedure, it is possible to propagate a single trajectory through many surface crossings, denoted as:

$$\begin{aligned} \boldsymbol{x}_{i+1} &= G(\boldsymbol{x}_i) && (5.87)\\ &= G(G(\boldsymbol{x}_{i-1})) && (5.88)\\ &= G^2(\boldsymbol{x}_{i-1}) && (5.89)\\ &= G^{i+1}(\boldsymbol{x}_0) && (5.90) \end{aligned}$$

This notation still carries around a full $2n$ state vector, although technically this can be reduced by one dimension to a $2n-1$-dimensional state vector through the constraint $S(\boldsymbol{x}_i) = 0$. It is important to note that in the computation of a surface of section, to compute subsequent passages it is necessary to integrate the entire trajectory forward (or backwards) in time. For a stroboscopic map in a time periodic dynamical system the definition of the surface of section is $\boldsymbol{x}_n = \phi(t_o + nT; \boldsymbol{x}_o, t_o)$.

5.7.2 Poincaré Maps

As the surface of section map was only defined for a time invariant system, the discrete surface of section mapping defined above has a Jacobi integral of motion. Specifically,

$$H(\boldsymbol{x}_0) = H(G^i(\boldsymbol{x}_0)) = C_0 \tag{5.91}$$

As the Hamiltonian is a relatively simple algebraic function it is also possible to eliminate an additional coordinate or momenta from consideration, reducing the dimensionality of the state by two (one for the surface of section and the other for the energy) to a $2n-2$-dimensional system. Generally, one removes a velocity or momentum component using the constant Hamiltonian, while one removes a coordinate using a surface of section. This is not a fundamental restriction, however. Once the additional state is removed from the map, the value of the Hamiltonian along a specific trajectory becomes a new parameter of the system and the resulting system is defined as a Poincaré map.

To formalize this, assume that two states are removed from the original $2n$-dimensional state vector $\boldsymbol{x}$, one each from the conditions $S(\boldsymbol{x}) = 0$ and $H(\boldsymbol{x}) = C$. For a given dynamical system it is important, of course, to systematically remove the same states at each crossing. Let us define the so-reduced state vector as $\boldsymbol{y} \in \mathbf{R}^{2n-2}$. It is important to remember that there is a unique state $\boldsymbol{x}$ associated with each reduced state $\boldsymbol{y}$, surface of section function S, and energy value C. Redefine the mapping function, now explicitly specifying the energy level that is being considered, with the form $g(\boldsymbol{y}; C)$. This discrete mapping function defines a Poincaré map, a mapping function that returns a trajectory back to a surface of section and maps every initial point chosen on that section with the same value of the Hamiltonian. Our dynamics are then:

$$\boldsymbol{y}_{i+1} = g(\boldsymbol{y}_i; C) \tag{5.92}$$

$$= g^{i+1}(\boldsymbol{y}_0; C) \tag{5.93}$$

The properties of the Poincaré map reduction have been studied extensively in the literature for a variety of applications. Our main application of this map will be in the computation of periodic orbits in time-invariant systems and in understanding the stability of such periodic motions. The map has also been extensively applied to the study of dynamical systems, an application not explored here.

It should be noted that if a dynamical system has additional integrals of motion, it is possible to further reduce the dimensionality of the Poincaré map. Perhaps the main benefit of the map is that it allows one to reduce the dimensionality of the system without having to go through the laborious process of eliminating integrals of motion from the full dynamical system. It can be assumed the reduction of the dynamical system to a Poincaré map has explicitly removed a coordinate–momentum conjugate pair. If dealing with a Lagrangian system, it suffices to remove a coordinate value and its time derivative, as when transformed into a Hamiltonian system this will also correspond to a coordinate–momentum pair being removed.

Systematic application of a Poincaré map to a collection of initial conditions implies that these conditions are chosen with the same Hamiltonian value, although the trajectories will be distinct in general. Mapping such collections repeatedly through the map then allows for the exploration of a dynamical system through an inspection of $2n-2$ dimensions, instead of the usual $2n$ dimensions. For 2-DOF systems this is tractable, as the resulting map is only two-dimensional. For higher-degree-of-freedom systems the utility of this approach is not as great, as the map itself lives in a four-dimensional or higher space.

5.7.3 Linearized Poincaré Map

Notions of local motion can be carried over to the Poincaré map. Assume a nonlinear map between points $\boldsymbol{y}_0$ and $\boldsymbol{y}_1 = g(\boldsymbol{y}_0; C)$. Consider an initial variation of $\boldsymbol{y}_0$ such that $\delta\boldsymbol{y}_0$ lies in the surface of section and conserves the Hamiltonian. Specifically, this means that $\frac{\partial S}{\partial \boldsymbol{y}_0} \cdot \delta\boldsymbol{y}_0 = 0$ and $\frac{\partial H}{\partial \boldsymbol{y}_0} \cdot \delta\boldsymbol{y}_0 = 0$. This small variation will also be mapped back to the surface of section and will naturally conserve energy, defining $\boldsymbol{y}_1 + \delta\boldsymbol{y}_1 = g(\boldsymbol{y}_0; C) + \Phi_{0,1}\delta\boldsymbol{y}_0$, where the reduced state transition matrix $\Phi_{0,1}$ is defined as the $(2n-2) \times (2n-2)$ matrix that linearly maps the Poincaré map back to itself in the vicinity of a nominal trajectory. Note the following identities $\frac{\partial S}{\partial \boldsymbol{y}_1} \cdot \delta\boldsymbol{y}_1 = 0$ and $\frac{\partial H}{\partial \boldsymbol{y}_1} \cdot \delta\boldsymbol{y}_1 = 0$. The practical computation of this linearized map is discussed later, at this point one only needs to note its existence.

This has defined a mapping of local variations about a nominal trajectory

$$\delta\boldsymbol{y}_1 = \Phi_{1,0}\delta\boldsymbol{y}_0 \tag{5.94}$$

which naturally generalizes to the iterated map

$$\begin{aligned} \delta\boldsymbol{y}_{n+1} &= \Phi_{n+1,n}\delta\boldsymbol{y}_n && (5.95)\\ &= (\Pi_{i=0}^{n}\Phi_{i+1,i})\,\delta\boldsymbol{y}_0 && (5.96) \end{aligned}$$

where the state transition matrix $\Phi_{i+1,i}$ depends on the associated nominal trajectory between the points $\boldsymbol{y}_i$ and $\boldsymbol{y}_{i+1} = g(\boldsymbol{y}_i; C)$, and is defined for a specific surface of section and energy value. If the reduction to the nonlinear Poincaré map is performed by systematic reduction of a coordinate–momentum pair, the resulting system is Hamiltonian and hence the linearized Poincaré map will be a symplectic matrix, and thus will inherit all the relevant properties defined for these maps.

5.7.4 Periodic Orbits and Stability

Now, let us reconsider periodic orbits in light of the Poincaré map. Note that a periodic orbit will correspond to a fixed point in a Poincaré map, denoted as $\boldsymbol{y}^*$. It is defined as

$$\boldsymbol{y}^* = g(\boldsymbol{y}^*; C) \tag{5.97}$$

To evaluate the stability of this fixed point, consider a small variation in the neighborhood of its initial conditions and its expansion through the linearized map

$$\begin{aligned} \boldsymbol{y}^* + \delta\boldsymbol{y}_1 &= g(\boldsymbol{y}^* + \delta\boldsymbol{y}_0; C) && (5.98)\\ &= g(\boldsymbol{y}^*; C) + \Phi^*_{1,0}\delta\boldsymbol{y}_0 + \dots && (5.99)\\ \delta\boldsymbol{y}_1 &= \Phi^*_{1,0}\delta\boldsymbol{y}_0 + \dots && (5.100) \end{aligned}$$

The linear map about a fixed point is independent of which iteration it is being applied to, and in particular $\Phi^*_{1,0} = \Phi^*_{i+1,i} = \Phi_M$. The linearized Poincaré map Φ_M is defined as the monodromy matrix, and it describes the stability of the periodic orbit in question. To note this, recall from Floquet theory that the eigenvalues of the state transition matrix mapped over one period were independent of where the matrix was evaluated along the orbit. Thus, the eigenvalues that the monodromy marix inherits from the full state transition matrix are independent of where the surface of section is chosen along the periodic orbit, and iterates of the monodromy matrix will then describe local motion on the surface of section from crossing to crossing

$$\delta\boldsymbol{y}_n = \Phi^n_M\delta\boldsymbol{y}_0 \tag{5.101}$$

Before the eigenstructure of Φ_M is discussed, note that the reduction process to a constant energy surface and to a surface of section will remove the two generic unity eigenvalues present in the full state transition matrix. From Floquet theory, the monodromy matrix has the general form $e^{M'T}$, where M' is now a constant matrix of size $(2n-2)\times(2n-2)$ and generically does not have zero eigenvalues.

Consider the eigenvalue and right eigenvector pairs of Φ_M, $(\lambda_i, \boldsymbol{u}_i, i = 1, 2, \dots, 2n-2)$. Again, associated with each λ_i is its complex conjugate, its inverse, and the inverse of its complex conjugate. If λ_i is real, then only its inverse must also exist, and if the magnitude of λ_i is unity, then only its complex conjugate must exist. The interpretation and analysis of the stable, unstable and center manifolds associated with the fixed point are all analogous to the discussion for the structure in the vicinity of an equilibrium point, except now instead of being continuous in time the dynamics occur at discrete mappings.

Stability of Periodic Orbits in a 2-DOF problem

Consider the monodromy matrix computed for a periodic orbit in a 2-DOF problem. In this case it is a simple 2×2 matrix, and thus it is possible to provide a tractable

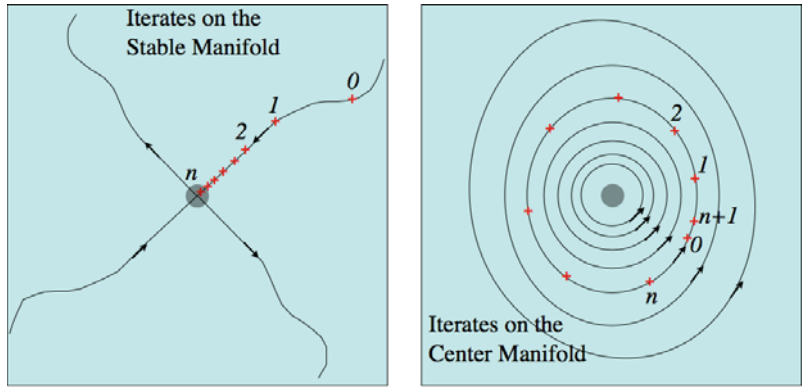


Fig. 5.3 Illustration of a stable and unstable fixed point, showing the manifolds and iterates.

discussion of its stability. The general form of the monodromy matrix will be

$$\Phi_M = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{5.102}$$

with the constraint that $ad - bc = 1$, which for 2×2 matrices is the necessary and sufficient condition for them to be symplectic. The eigenvalues of this matrix can be specifically computed from

$$\lambda_{\pm} = \frac{1}{2}(a+d) \pm \frac{1}{2}\sqrt{(a+d)^2 - 4} \tag{5.103}$$

and $\lambda_{\pm}\lambda_{\mp} = 1$. If $|a+d| < 2$ the periodic orbit is linearly stable and $\bar{\lambda}_{\pm} = \lambda_{\mp}$. If instead $|a+d| > 2$ the eigenvalues are both real with one greater than 1 in magnitude and the other less than 1 in magnitude, and $\lambda_{\pm} = 1/\lambda_{\mp}$. Thus, in this case the orbit is unstable in general, and has one stable and one unstable manifold. In the limiting case when $a + d = 2$, the eigenvalues equal 1 and are repeated, and usually represent a transition of the eigenvalues from being on the unit circle to being off the unit circle, or vice versa. If $a + d = -2$, the eigenvalues equal -1 and are repeated. This also represents a possible transition of the eigenvalues from stability to instability, or vice versa. The implications of these different transitions are discussed below. Note that periodic orbits in 2-DOF systems cannot exhibit complex instability, unlike equilibrium points in 2-DOF systems.

Stability of Periodic Orbits in a 3-DOF problem

Consider the monodromy matrix computed for a periodic orbit in a 3-DOF problem. In this case it is a 4×4 matrix, and again it is possible to provide a tractable discussion of its stability. The general form of the monodromy matrix will be

$$\Phi_M = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ j & k & l & m \\ n & p & q & r \end{bmatrix} \tag{5.104}$$

with the symbols i and o being reserved. The characteristic equation can be expressed as

$$0 = \lambda^4 + \alpha\lambda^3 + \beta\lambda^2 + \alpha\lambda + 1 \tag{5.105}$$

$$\alpha = -(a + f + l + r) \tag{5.106}$$

$$\begin{aligned} \beta = {} & af - be + al - cj + ar - dn \\ & + fl - gk + fr - hp + lr - mq \end{aligned} \tag{5.107}$$

Defining $\gamma = \lambda + \lambda^{-1}$ it is possible to show that γ satisfies the equivalent characteristic equation

$$\gamma^2 + \alpha\gamma + \beta - 2 = 0 \tag{5.108}$$

For stability it can be shown that γ must be a real number and satisfy $-2 \leq \gamma \leq 2$. These conditions will hold if and only if the following inequalities hold

$$8|\alpha| - 16 \leq 4\beta - 8 \leq \alpha^2 \leq 16 \tag{5.109}$$

This defines a stability plot in the α, β space, showing the different possible stability states for a periodic orbit (see Fig. 5.4).

5.8 Periodic Orbit Families

Given the Poincaré map definition and the identification of the monodromy matrix it becomes tractable to discuss the continuation of periodic orbits into families in a more complete fashion. A number of different topics are discussed in the following, including the generic isolation of periodic orbits at fixed values of Jacobi energy, the possible stability transitions along a period orbit family as its energy changes, and the intersections of periodic orbits with other families of the same and higher periods.

5.8.1 Isolation of Periodic Orbits at Constant Jacobi Values

First determine the conditions for another periodic orbit to exist in the vicinity of a given periodic orbit at a fixed value of energy. Given a fixed point $\boldsymbol{y}^*$ of a Poincaré map

$$\boldsymbol{y}^* = g(\boldsymbol{y}^*; C) \tag{5.110}$$

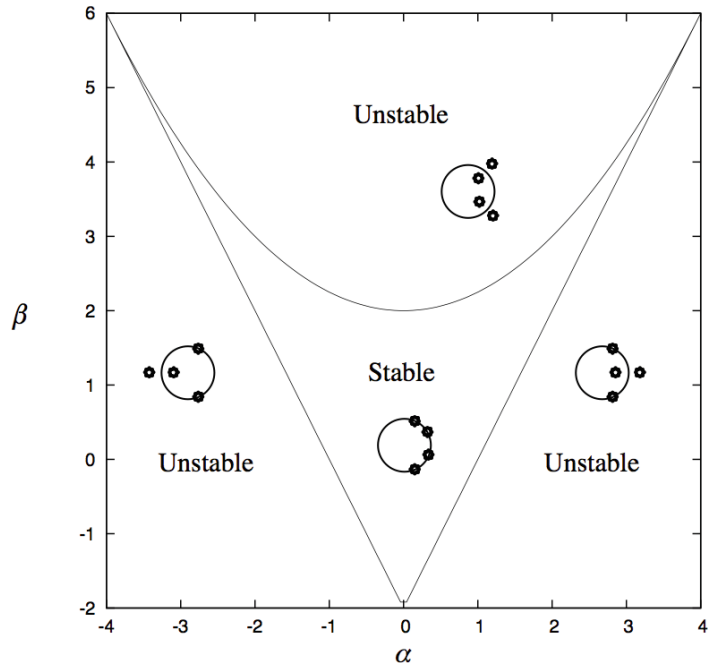


Fig. 5.4 Stability diagram for periodic orbits in a time invariant 3-DOF system as a function of α vs. β.

consider a variation in the state while keeping the energy fixed. Assume the variation is small enough to expand the fixed point using the monodromy matrix, and tacitly assume that the new point maps back into itself

$$\boldsymbol{y}^* + \delta\boldsymbol{y}^* = g(\boldsymbol{y}^*; C) + \Phi_M \delta\boldsymbol{y}^* \tag{5.111}$$

Simplifying the condition becomes

$$[\Phi_M - I]\, \delta\boldsymbol{y}^* = 0 \tag{5.112}$$

Recall that Φ_M has had its unity eigenvalues removed and thus it generically will not have any unity eigenvalues except at exceptional points. If there are unity eigenvalues, then there will be neighboring periodic orbits with their relative location to $\boldsymbol{y}^*$ located along the null space of $\Phi_M - I$. If, as is expected for the generic case, there are no unity eigenvalues of Φ_M, then the periodic orbit is isolated at the fixed value of energy.

This is generally the case for the stroboscopic Poincaré map defined for time-periodic systems. In fact, for these systems where the conserved Hamiltonian equals the time-momentum plus the Hamiltonian, one can show that the value of this Hamiltonian is in general invariant with the state and is always equal to the same constant. This implies that periodic orbits in stroboscopic systems are generically

isolated, unless there exists another integral of motion that introduces additional unity eigenvalues.

5.8.2 Continuation of Periodic Orbits with Jacobi Energy

Given the generic isolation of periodic orbits at a given energy, their continuation is explored with a variation in the energy. Now in addition to varying from the fixed point we also allow the Hamiltonian value C to vary by δC, leading to the condition

$$[\Phi_M - I]\,\delta \boldsymbol{y}^* = \left.\frac{\partial g}{\partial C}\right|_* \delta C \tag{5.113}$$

where the partial of the Poincaré map with respect to energy can be computed relatively easily, as is discussed in the next chapter. Assuming that Φ_M has no unity eigenvalues the new, neighboring periodic orbit is defined by

$$\boldsymbol{y}^* + [\Phi_M - I]^{-1}\, g_C^* \delta C \tag{5.114}$$

and maps into itself at the linear level. An algorithm for using the linearized Poincaré map to nonlinearly converge on a fixed point is discussed in the next chapter. If, instead, Φ_M has unity eigenvalues then this implies that there exists a neighboring periodic orbit with the same value of energy. This is not a generic situation and will generally only apply for an exceptional periodic orbit.

Thus, a family of periodic orbits can be defined that passes through $\boldsymbol{y}^*$ at a value of C and generally continues for larger and smaller values of C. This can be represented as a mapping from C to $\boldsymbol{y}$ as $\boldsymbol{y} = \boldsymbol{h}(C)$, where $\boldsymbol{h} \in \mathbf{R}^{2(n-1)}$. The graph of $\boldsymbol{h}$ then defines the family, where the tangent vector to the curve is defined by $[\Phi_M - I]^{-1}\, g_C$.

5.8.3 Stability Transitions along a Periodic Orbit Family

As a family of periodic orbits is traced out, the eigenvalues of the associated monodromy matrix are expected to vary as well. To emphasize this we now show the monodromy matrix as a function of energy, $\Phi_M(C)$, assuming that there exists a family of fixed points defined by $\boldsymbol{y} = \boldsymbol{h}(C)$.

Let us consider a 2-DOF problem first, as there are only two eigenvalues for such a system. If the family is stable at a point $\boldsymbol{h}(C_0)$, with eigenvalues $e^{\pm i\theta}$ (where i is the imaginary unit) where the rotation angle $\theta \neq 0, \pi$, then the local family will always be a stable fixed point as well. The only way the system can transition to an unstable orbit is for the eigenvalues to meet at $\theta = 0, \pi$ and transition to the positive or negative real axis, respectively. The converse situation holds if the fixed point is unstable at a given value of C_0. The two different possible transitions have significantly different implications.

Intersection along $\theta = 0$

First consider the implications if $\theta \to 0$. Drawing on an analysis of such stability transitions for symmetric 2-DOF periodic orbits by Hénon [66], one can distinguish two cases. For both cases it can be assumed that Φ_M has unity eigenvalues so $\|\Phi_M - I\| = 0$ and the matrix $[\Phi_M - I]$ cannot be uniquely inverted. Then the solutions that are consistent with this condition can be explored

$$[\Phi_M - I]\,\delta \boldsymbol{y} = \boldsymbol{g}_C \delta C \tag{5.115}$$

In this case one can always choose a $\delta \boldsymbol{y}$ along an eigenvector of unity such that the left-hand side equals zero.

Local Extremum If either term of $\boldsymbol{g}_C \neq 0$, then the right-hand side can only be made to equal zero if $\delta C = 0$. This implies that the current family is forced to change along a direction defined by the null space of $\Phi_M - I$ with no change in energy. This generically corresponds to the periodic family curve $\boldsymbol{h}(C)$ being at an extremum with respect to C and, in some sense, turning back on itself in phase space. See Fig. 5.5 left for a graphical representation. Following such a transition there will exist two branches of the periodic orbit family at a given value of energy, one will be stable and the other unstable. Such reversals in a periodic orbit family are relatively common, and essentially represent a bifurcation with the family appearing or disappearing as the energy C is varied.

Family Intersection If, on the other hand, $|\boldsymbol{g}_C| \to 0$ as $\theta \to 0$, then this represents an intersection of the current family with another family with the same value of energy, C. For this to be defined these two fixed points must physically cross through the same point in phase space at the same value of C. At such a crossing of families, they will generally exchange stability, as shown in Fig. 5.5 right. The vanishing of the $\boldsymbol{g}_C$ terms is a hallmark of such an intersection, as in the vicinity of the intersection there are multiple solutions which the linear conditions cannot accommodate. If the continuation conditions were developed to second order or higher (depending the the number of intersections), then there should be a definitive set conditions that can be solved for.

Intersection along $\theta = \pi$

In this case the eigenvalues collide along the negative real axis at $\lambda = -1$. At such a bifurcation the matrix $\Phi_M - I$ has full rank and is invertible. Thus, the family seems to seamlessly transition from stable to unstable, or vice versa, at such a transition. What is really going on here is a bit more interesting, however. If the double-iterate of the Poincaré map is considered, $\boldsymbol{g}^2(\boldsymbol{y}^*)$, this is still a fixed point but now its monodromy matrix is Φ_M^2 and, at the stability transition point, will have a pair of unity eigenvalues. This is an example of a period-doubling bifurcation, where a

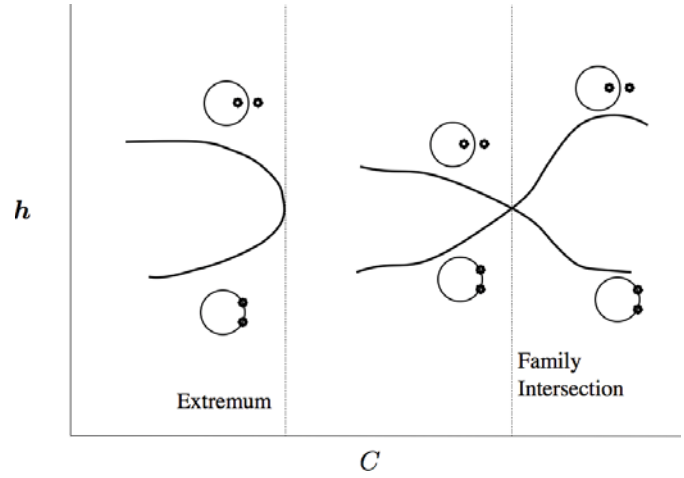


Fig. 5.5 Local extremum of a periodic orbit family (left) and an intersection of two periodic orbit families (right).

pair of fixed points of twice the period are born, inheriting the stability type of the single-period orbit that undergoes its transition. The resulting family may either persist or may intersect with the single-period family again at a different value of C, providing the single family with its same initial stability type. Figure 5.6 shows the generic situation.

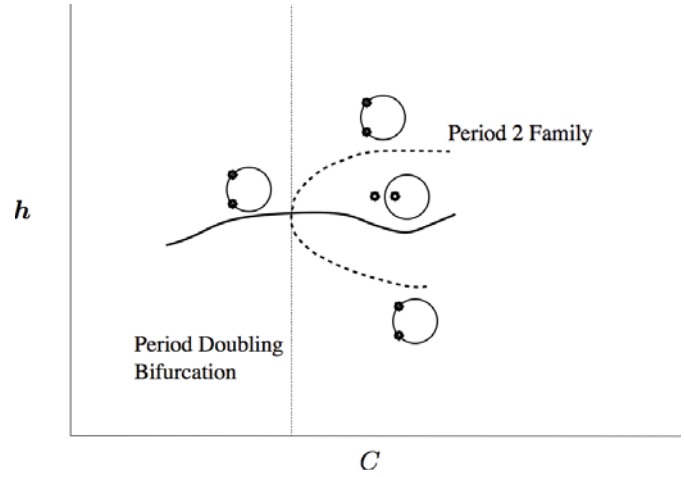


Fig. 5.6 Period doubling stability bifurcation.

Higher-period Intersections

Passage of θ through π corresponds to the birth of a period-2 family. Similarly passage through special values of $\theta = 2\pi m/n$ may result in n-period families to bifurcate. Whether such higher-period fixed points bifurcate or not, it is important to note that the base fixed point will not undergo any stability transitions for these bifurcations, and indeed will seemingly remain unaware of whether such bifurcations are occurring.

Higher-dimension Systems

Similar reasoning applies to systems with three or more degrees of freedom, except there exists one additional type of intersection. Here, it is possible for two sets of eigenvalues to intersect along the unit circle and bifurcate into an unstable periodic orbit with a set of two-dimensional stable and unstable spiral manifolds. Since these stability bifurcations do not occur along the real axis, continuation of the family is not a problem. It is interesting to note that there are rigorous conditions for when such an intersection will result in a stability bifurcation. Specifically, Krein's Theorem states that if the pair of colliding eigenvalues on the unit circle are traveling in the same direction, then a stability bifurcation will not occur. If they are traveling in the opposite direction, then a bifurcation may occur [5].

6. Solution and Characterization Methods

In the following we discuss various methods by which the equations of motion can be solved to find the trajectory of a particle or a solution flow in the neighborhood of a trajectory. We make an implicit assumption that the equations we wish to solve are non-integrable, meaning that there exists no closed-from solution. In many situations we will rely on a closed-form solution of a simpler, non-perturbed system to help generate a solution to a more general problem. As the given problems are non-integrable, we are always limited in that our solutions and solution methods are approximate, and only represent the true solution up to some degree of precision over a finite timespan. Some exceptions to these occur for the special solutions of dynamical systems, namely (relative) equilibria and periodic orbits. A relative equilibria is an exact solution to a system, and in general does not rely on a solution of the simpler, non-perturbed problem. A periodic orbit is not exact, as it has a limit on the precision (except in the rare cases where a completely convergent series is found in closed form), although there is no limit on the timespan for the solution, that being infinite. All other solutions have limitations on both precision and time.

Our goal in constructing solutions is not to find single trajectories of arbitrary precision – that can already be done using numerical techniques. Rather, our goal is to represent a solution and the neighborhood about that solution, since that is the most important aspect of our dynamical understanding of motion.

In the following we will sometimes not specifically assume any structure for the equations of motion, due to the generality of many solution techniques, but can assume a generic formulation of the equations of motion as $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, t)$ where $\boldsymbol{x} \in \mathbf{R}^{2n}$ and $\boldsymbol{f} : \mathbf{R}^{2n} \times \mathbf{R} \to \mathbf{R}^{2n}$. In other instances we will include an added term representing perturbations from an integrable problem. Exceptions to this convention will be noted. We will also use the Hamiltonian formalism when appropriate.

D.J. Scheeres, *Orbital Motion in Strongly Perturbed Environments*, Springer Praxis Books,
DOI 10.1007/978-3-642-03256-1_6, © Springer-Verlag Berlin Heidelberg 2012

6.1 Numerical Integration

The most direct, and in many ways simplest and most precise, method to solve for the motion of a particle given the equations of motion is to numerically integrate the trajectory. A numerical integration takes a current value of the state and the dynamics function, which must be computable as a function of the state, and generates a discrete map in time to a subsequent (or previous) state. For a given numerical integrator there is always an error associated with this map; however, this error can be driven to a small value by appropriate choice of the time step and algorithm. The field of numerical analysis is far too comprehensive to be easily summarized, but for astrodynamics applications there are a few main approaches that are used, described in brief later. The main point is that the field is sufficiently advanced so that there is no reason to doubt the stated numerical precision of a solution, given the caveats associated with the particular numerical scheme. Thus, one can consider the output of a numerical integrator to be a valid solution to the differential equations, to within some degree of precision. It should not be mistaken with the actual solution, which will have additional properties that the numerical solution may not enjoy, such as symplecticity and analyticity, yet for a given initial state it is an accurate representation of the solution. As such, these solutions can be used to test analytical results, to provide precise predictions or reconstructions of a particular trajectory, and can be used as a tool for analysis.

Technically, the solution that comes out of a numerical integrator is a series of discrete states $\boldsymbol{x}_i$ at given time steps, t_i, where $i = 0, 1, \ldots, N$, t_0 is the initial time, and t_N is the final time (where $t_N < t_0$ can occur without ambiguity). Most numerical routines also provide a means by which to interpolate between the states $\boldsymbol{x}_i$ and $\boldsymbol{x}_{i+1}$, sometimes using neighboring states to improve the interpolation. This provides the illusion of a continuous solution, whereas the interpolated points will in general have lower precision than the discrete points output by the system, and may not share the fundamental properties of the solutions as they are not directly constructed from the equations of motion. It is important to note that the total timespan of a given map, $t_N - t_0$, need not be an integer number of some fundamental time step, as almost all integrators have the capability to adjust the time step to achieve any desired end point in time. Because of this a numerical solution can be represented as if it were a continuous time solution, $\boldsymbol{x}(t) = \boldsymbol{\phi}(t; \boldsymbol{x}_o, t_o)$, where the solution is also subject to the chosen parameters of the system. This allows us to represent a numerical solution precisely in the same way as an analytical solution would be represented.

6.1.1 Lagrangian, Hamiltonian and Orbit Element Computations

If the system is defined by a Lagrangian in terms of coordinates and their time rate of change, $L(\boldsymbol{r}, \dot{\boldsymbol{r}}, t)$, then the corresponding equations of motion are found by solving Lagrange's equation $\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\boldsymbol{r}}}\right) - \frac{\partial L}{\partial \boldsymbol{r}} = 0$ for the acceleration, resulting in a set of equations $\ddot{\boldsymbol{r}} = \boldsymbol{F}(\boldsymbol{r}, \dot{\boldsymbol{r}}, t)$. Then make the usual transformation of a second-order differential equation into a first-order by defining $\boldsymbol{x} = [\boldsymbol{r}, \boldsymbol{v}]$, where $\dot{\boldsymbol{r}} = \boldsymbol{v}$, and

$\boldsymbol{f} = [\boldsymbol{v}, \boldsymbol{F}(\boldsymbol{r}, \boldsymbol{v}, t)]$. Hamilton's equations with a Hamiltonian function $H(\boldsymbol{q}, \boldsymbol{p}, t)$ are already in the appropriate form, as if given $\boldsymbol{x} = [\boldsymbol{q}; \boldsymbol{p}]$ the equations $\dot{\boldsymbol{x}} = JH_{\boldsymbol{x}}$ follow, where $\boldsymbol{f}(\boldsymbol{x}, t) = JH_{\boldsymbol{x}}$.

We also consider the orbit element forms of the equations that arise from the Lagrange and Gauss equations. In these one can generically assemble the orbit elements into a single state vector, for the classical orbit elements this could be $\boldsymbol{x} = [a, e, i, \omega, \Omega, \tau_o]$. Then both sets of equations can be immediately written into general form. In fact, there is a bit more structure to the equations than that. Lagrange's planetary equations have the form $\dot{\boldsymbol{x}} = F_L(\boldsymbol{x})\frac{\partial R}{\partial \boldsymbol{x}}$, where F_L is a 6×6 matrix that is a nonlinear function of the orbit elements but not time, and $R(\boldsymbol{x})$ is the perturbing potential. Similarly, the Gauss equations take the form $\dot{\boldsymbol{x}} = F_G(\boldsymbol{x}, t)\boldsymbol{a}$, where F_G is now a 6×3 matrix that is a function of the state $\boldsymbol{x}$ and time, while $\boldsymbol{a} \in \mathbf{R}^3$ is the applied acceleration. Once in these forms the equations of motion can be directly passed to a generic numerical integration scheme.

6.1.2 State Transition Matrix Computations

Note that the state transition matrix is a crucial component of the theory. The state transition matrix can be analytically solved for only in the case where the nominal solution is an equilibrium point of a time-invariant system. For any other solution one must numerically integrate to find $\Phi(t, t_o)$. Already discussed is the differential equation which Φ satisfies, which is always derived from the fundamental equations of motion as

$$\dot{\Phi}(t, t_o) = \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right|_t \Phi(t, t_o) \tag{6.1}$$

$$\Phi(t_o, t_o) = I \tag{6.2}$$

where the solution of the nominal trajectory $\boldsymbol{\phi}(t; \boldsymbol{x}_o, t_o)$ must be substituted into the dynamics matrix

$$A(t) = \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right|_t$$

at each point of time. It is often convenient to integrate both the state $\boldsymbol{x}(t)$ and the state transition matrix simultaneously, thus evolving the dynamics associated with a point in phase space and the linear description of the dynamics about that point.

Depending on what form of the equations of motion are being integrated, the state transition matrix may not be symplectic. However, for many cases the state transition matrix can be simply transformed into a symplectic form. For example consider a Lagrangian framework, where

$$\Phi(t, t_o) = \frac{\partial(\boldsymbol{r}(t), \boldsymbol{v}(t))}{\partial(\boldsymbol{r}_o, \boldsymbol{v}_o)}$$

Unless the position and velocity are relative to an inertial frame, this will not be a symplectic matrix. Recall the simple Legendre Transformation, where the current

position vector can be taken as the Hamiltonian coordinate $\boldsymbol{q}(t) = \boldsymbol{r}(t)$ and the momentum is defined as the partial of the Lagrangian with respect to the velocity, $\boldsymbol{p}(t) = \frac{\partial L}{\partial \boldsymbol{v}}$. Using these definitions one can express a symplectic state transition matrix in terms of the Lagrangian variables as

$$\frac{\partial(\boldsymbol{q}(t), \boldsymbol{p}(t))}{(\boldsymbol{q}_o, \boldsymbol{p}_o)} = \begin{bmatrix} I & 0 \\ \left.\dfrac{\partial^2 L}{\partial \boldsymbol{v} \partial \boldsymbol{r}}\right|_t & \left.\dfrac{\partial^2 L}{\partial \boldsymbol{v}^2}\right|_t \end{bmatrix} \Phi(t, t_o) \begin{bmatrix} I & 0 \\ \left.\dfrac{\partial^2 L}{\partial \boldsymbol{v} \partial \boldsymbol{r}}\right|_{t_o} & \left.\dfrac{\partial^2 L}{\partial \boldsymbol{v}^2}\right|_{t_o} \end{bmatrix}^{-1} \tag{6.3}$$

where the state transition matrix Φ is computed for the Lagrangian variables.

6.1.3 Common Integrators

The field of numerical integration continues to be an area of innovation and research advancement. Thus in the following only a few different types of numerical integration routines are mentioned, leaving the interested reader to decide for themselves which one to use. Each tends to have advantages that suit themselves for different applications. The day when one single approach to numerical integration will be used is still apparently in the distant future.

Variable Step Predictor–Corrector Routines

Perhaps the most widely used numerical integrators are those that simultaneously develop a prediction and an estimate of the current error in the numerical integration. Based on the predicted error and the desired tolerance in the computation the time step can be adjusted to be longer or shorter, varying this as the solution is being computed. A common formulation for these integrators are the Runga–Kutta methods, and they are widely available in pre-packaged computational environments. These methods are often quite useful for astrodynamics problems where there can be relatively long periods of small perturbations interspersed with short periods of more intense interactions between bodies. When numerically integrating high-eccentricity orbits this occurs quite frequently, with longer time steps being acceptable through apoapsis and short time steps through periapsis.

Fixed-Step Integrators

Another school of thought is willing to forgo speed for greater precision in computation. Along these lines are the extremely accurate fixed-step integration methods such as Burlisch–Stoer. Use of these algorithms in the literature is usually associated with computations that strive to achieve machine precision over long timespans [52, 53].

Symplectic or Variational Integrators

More recently, the development and application of symplectic or variational integrators has gained much attention. These are integration methods where the discrete time steps are chosen so that they conserve some fundamental property of the dynamical system. For instance, symplectic integrators have been developed for the integration of Hamiltonian systems and preserve the symplectic structure of the underlying equations of motion. These approaches have been further generalized recently to conserve other quantities that might be associated with the geometry of motion or the underlying dynamical system. When symplecticity of the state transition matrix is important it is possible to implement an integration scheme for just that matrix which will ensure symplecticity independent of how the nominal trajectory is integrated [184]. A useful review of numerical integrations, with a definite bias toward variational integrators, is provided in [60].

6.1.4 Verification of a Numerical Integrator's Performance

More important than what numerical integration routine is used is that the researcher performs verification and validation tests of their numerical integrations to ensure that the algorithms have been implemented correctly and that adequate error tolerances have been chosen. For astrodynamical problems, which usually are dealing with a system perturbed from two-body motion, an effective way to verify integration performance is to directly integrate the two-body problem. As this problem is integrable, it is possible to transform the system at each point into orbit elements and verify that they remain conserved for stressful integration cases such as high-eccentricity orbits. For systems that have a Jacobi integral defined for the full system, it is usually also good practice to confirm that the integration scheme keeps the integral constant. These are not proofs of the integrator's accuracy, but are effective ways to verify performance.

For integrations of the state transition matrix it is also useful to perform validations of the numerical accuracy of its computation. For all dynamical systems without dissipation, the determinant of the state transition matrix should generically equal 1, and this can be used as a check. Additionally, for a nominally symplectic matrix it is possible to check whether the matrix verifies the symplectic condition, and view deviation from this either as an indication of poor error tolerances or a mistake in the implementation.

6.2 Computation of Equilibrium Solutions

6.2.1 General Algorithm

Given a dynamical system reduced to Hamiltonian form, $\dot{\boldsymbol{x}} = JH_{\boldsymbol{x}}$, an equilibrium solution is a particular state of the system, $\boldsymbol{x}^*$, which causes the dynamics function to be identically zero. The defining equation is:

$$H_{\boldsymbol{x}}(\boldsymbol{x}^*) = \mathbf{0} \tag{6.4}$$

where the Hamiltonian is assumed to be time-invariant and hence is a constant of the motion. If the equation is written as a Lagrangian system, the equilibrium condition can be reduced to a set of algebraic equations of the form

$$\frac{\partial L(\boldsymbol{r}^*, \mathbf{0})}{\partial \boldsymbol{r}} = \mathbf{0}, \tag{6.5}$$

where this is a three-dimensional equation instead of the six-dimensional Hamiltonian form. Thus, for a specific problem, it is often easier to solve the Lagrangian form for zero accelerations in the configuration space, as the equilibrium condition for the velocities is trivially satisfied.

When the dynamical system has no integrals of motion other than the Jacobi integral (for a time invariant Hamiltonian or Lagrangian), then equilibrium points are in general non-degenerate, or $|H^*_{\boldsymbol{xx}}| \neq 0$. This implies that a Newton–Raphson method can be applied to compute the equilibrium points. Specifically, given a test value for an equilibrium point $\boldsymbol{x}_i$, search for the correction $\delta\boldsymbol{x}_i$ that will satisfy the condition: $H_{\boldsymbol{x}}(\boldsymbol{x}_i + \delta\boldsymbol{x}_i) = \mathbf{0}$. If "close" to the solution (i.e., if a reasonable guess for the point is provided) it is possible to expand the function in a Taylor Series: $H_{\boldsymbol{x}}(\boldsymbol{x}_i) + H_{\boldsymbol{xx}}(\boldsymbol{x}_i) \cdot \delta\boldsymbol{x}_i + \ldots = \mathbf{0}$. Solution of this equation at the linear order provides the update:

$$\delta\boldsymbol{x}_i = -H_{\boldsymbol{xx}}(\boldsymbol{x}_i)^{-1} \cdot H_{\boldsymbol{x}}(\boldsymbol{x}_i) \tag{6.6}$$

and a new test solution is computed as $\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \delta\boldsymbol{x}_i$ and the procedure is iterated until the equilibrium condition is satisfied to an acceptable degree of precision.

6.2.2 Complications from Additional Integrals

The presence of an integral of motion in addition to the Jacobi integral complicates this procedure and ensures that the above algorithm will fail. As seen earlier in Chapter 5 the state transition matrix expanded about an equilibrium solution has a pair of unity eigenvalues when an additional integral of motion exists. Specifically, for an integral of motion linearly independent from the Hamiltonian, $h(\boldsymbol{x}) = C_h$, the state transition matrix evaluated at an equilibrium $\boldsymbol{x}^*$ satisfies the following condition: $h_{\boldsymbol{x}}(\boldsymbol{x}^*) \cdot [I - \Phi(t, t_o; \boldsymbol{x}^*)] = 0$. Thus, $h_{\boldsymbol{x}}$ is a left eigenvector of Φ with a unity eigenvalue. The symmetry of Hamiltonian systems ensures that there exists another unity eigenvalue associated with the matrix with a right eigenvector $Jh_{\boldsymbol{x}}(\boldsymbol{x}^*)$. As the associated linear system is time-invariant the explicit solution exists

$$\Phi(t, t_o; \boldsymbol{x}^*) = e^{JH^*_{\boldsymbol{xx}}(t - t_o)} \tag{6.7}$$

$$= I + (t - t_o)JH^*_{\boldsymbol{xx}} + \frac{1}{2}(t - t_o)^2 \left[JH^*_{\boldsymbol{xx}}\right]^2 + \ldots \tag{6.8}$$

Substitution of this form of the solution into the eigenvalue equations yields the following conditions:

$$H^*_{\boldsymbol{xx}} \cdot J \cdot h^*_{\boldsymbol{x}} = 0 \tag{6.9}$$

Thus, the Hamiltonian is degenerate at an equilibrium with two zero eigenvalues and any numerical technique using the Newton–Raphson method to compute the point will become ill-defined as the test solution comes close to the equilibrium.

This happens as associated with every equilibrium point of a system with an additional integral is a family of neighboring equilibrium points at the same value of energy and same value of the integral $h(\boldsymbol{x})$. Specifically, suppose an equilibrium solution is given such that $H^*_{\boldsymbol{x}} = \mathbf{0}$ and $h(\boldsymbol{x}^*) = C_h$. Consider an arbitrary variation from this point $\delta\boldsymbol{x}$, the question is whether there exists a neighboring point that will satisfy both the equilibrium condition and have the same value of the integral h. If so, then there should be a solution satisfying the two conditions $H^*_{\boldsymbol{xx}} \cdot \delta\boldsymbol{x} = 0$ and $h^*_{\boldsymbol{x}} \cdot \delta\boldsymbol{x} = 0$. As the Hamiltonian is degenerate at the equilibrium point the first condition is satisfied by the eigenvector $\delta\boldsymbol{x} = Jh^*_{\boldsymbol{x}}$, meaning that there is another equilibrium point lying in this direction. This vector direction also satisfies $h^*_{\boldsymbol{x}} \cdot Jh^*_{\boldsymbol{x}} = 0$ trivially. Thus, there is not a unique point to which the algorithm can converge, and hence the procedure becomes ill-defined.

This demonstration also suggests a specific approach that can be used to construct a more robust computation algorithm. Instead of searching over the full state direction $\delta\boldsymbol{x}$, one can restrict the search to the direction $\delta\boldsymbol{x}_\perp$ defined such that $\delta\boldsymbol{x}_\perp \cdot Jh_{\boldsymbol{x}}(\boldsymbol{x}) = 0$, i.e., not allowing variations along the line of continuous equilibria. Practically speaking, this equation can be solved in terms of one variation direction to define a $2n-1$ space in which to search. It is also possible to make a further reduction, by choosing to eliminate a second direction by further constraining $\delta\boldsymbol{x}_\perp$ by the condition $h_{\boldsymbol{x}} \cdot \delta\boldsymbol{x}_\perp = 0$, maintaining a constant value of h. However, this presupposes that a value of h has been chosen that allows an equilibrium to exist, something that is not necessarily true.

An explicit example of this form of degeneracy can be given. Consider the relative two-body problem in a rotating reference frame with the Hamiltonian function $H = \frac{1}{2}\left(P_r^2 + \frac{1}{r^2}P_\theta^2\right) - \omega P_\theta - \frac{\mu}{r}$, where ω is the rotation rate. Due to the absence of θ the momentum P_θ is an integral of motion. The second, associated integral of motion for θ exists by quadratures once the solution for r and P_r is found. Applying the equilibrium condition yields

$$H_{\boldsymbol{x}} = \begin{bmatrix} -\frac{1}{r^3}P_\theta^2 + \frac{\mu}{r^2} \\ 0 \\ P_r \\ \frac{1}{r^2}P_\theta - \omega \end{bmatrix} \tag{6.10}$$

$$H_{\boldsymbol{xx}} = \begin{bmatrix} \frac{3}{r^4}P_\theta^2 - \frac{2\mu}{r^3} & 0 & 0 & -\frac{2}{r^3}P_\theta \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{2}{r^3}P_\theta & 0 & 0 & \frac{1}{r^2} \end{bmatrix} \tag{6.11}$$

It can be verified that the matrix $H_{\boldsymbol{xx}}$ has two zero eigenvalues. The gradient of h is $h_{\boldsymbol{x}} = [0, 0, 0, 1]$ and $Jh_{\boldsymbol{x}} = [0, 1, 0, 0]$. If the search is constrained in the direction normal to $Jh_{\boldsymbol{x}}$ this implies that θ is fixed. If the search in the direction $h_{\boldsymbol{x}}$ is constrained, this implies that P_θ is fixed. Doing so provides us with two equations which can be solved to find $P_r = 0$ and $r = P_\theta^2/\mu$, with an ancillary condition

$\frac{1}{r^2}P_\theta - \omega = 0$. Note that an improper choice of P_θ may yield an inconsistency with one of these equations. One can either use the final equation to define the proper value of P_θ, something which is trivial for the given simple problem, or not be restricted to searching for it at a specified value of P_θ to begin with, which allows one to simultaneously solve for r, P_r and P_θ. For more general systems where the solutions are not as easily found, it is preferred to solve for the value of the integral simultaneously with the solution of the other states. Then the full solution is found to be: $r^* = \left(\mu/\omega^2\right)^{1/3}$, $P_r^* = 0$, $P_\theta^* = \left(\mu^2/\omega\right)^{1/3}$.

6.2.3 Stability Computation

Once an equilibrium point of the system is found, then the stability of that point can be determined by evaluating the eigenvalues of the matrix $JH^*_{\boldsymbol{xx}}$, evaluated at the equilibrium point. Only in the simplest cases can the eigenvalues and eigenvectors of this matrix be determined analytically. It is noted that even in the restricted three-body problem, the Euler equilibrium solutions must be found numerically for a precise evaluation of them. Thus, it is usually easiest to use a computational toolbox to evaluate the eigenvalues and eigenvectors of the dynamics matrix at the equilibrium point. Once these are found the methods outlined in the previous chapter for describing motion in the vicinity of the equilibrium point can be applied.

6.3 Computation of Periodic Orbits

6.3.1 General Algorithm

The traditional approach to analytical or numerical computation of periodic orbits relies heavily on the symmetric properties of the forces in the system [70, 65, 18]. Examples of this include the computation of periodic orbits in the restricted three-body problem and the Hill problem, best exemplified by the seminal work of Hénon. As a direct application to asteroids, this approach has been used to compute symmetric periodic orbits about tri-axial ellipsoids [153]. For application to "real" objects, however, there are no symmetries in the force fields and the periodic orbit computation must be generalized to a form that does not rely on such symmetries. The algorithm of Deprit and Henrard [31] views periodic orbits in terms of their intrinsic differential geometric properties and enables such computations, as applied by Lara [86] and others. In the following we present a more direct approach that we find to be more intuitive and which only requires the computation of the state transition matrix.

For the computation of a periodic orbit it is necessary to solve the equation:

$$\boldsymbol{x} = \phi(T; \boldsymbol{x}, 0) \tag{6.12}$$

for a given period T. The standard approach is to assume an initial state $\boldsymbol{x}_0$ that is sufficiently close to, but does not solve, the periodicity condition. Often this

initial estimate can be found by applying simple two-body theory to the perturbed problem. Then assume that a small correction, $\delta\boldsymbol{x}$, may solve the problem

$$\boldsymbol{x}_0 + \delta\boldsymbol{x} = \phi(T; \boldsymbol{x}_0 + \delta\boldsymbol{x}, 0) \tag{6.13}$$

$$= \phi(T; \boldsymbol{x}_0, 0) + \Phi(T, 0)\delta\boldsymbol{x} + \ldots \tag{6.14}$$

If the higher-order terms are ignored this equation can be rearranged to

$$[I - \Phi(T, 0)]\,\delta\boldsymbol{x} = \phi(T; \boldsymbol{x}_0, 0) - \boldsymbol{x}_0 \tag{6.15}$$

where the right-hand side is not exactly zero. At this point, prior to a periodic orbit having been converged on, the matrix on the left is generally non-singular and can be inverted to yield an estimate for the correction. Then the trial solution can be updated as $\boldsymbol{x}_1 = \boldsymbol{x}_0 + \delta\boldsymbol{x}$.

Normally this process would be repeated until the mismatch $|\phi(T, 0, \boldsymbol{x}_i) - \boldsymbol{x}_i| < \epsilon$. If the system is time periodic, then the matrix $\Phi(T, 0)$ will, generically, not have unity eigenvalues, even when the periodic orbit has been converged upon, and is itself the monodromy matrix for that orbit. Thus, in this sense, computation of periodic orbits in time periodic systems are much easier than in time-invariant systems, although generation of the appropriate initial conditions can be difficult. First, the period of the motion must be equal to an integer multiple of the system period and thus is immediately specified. Second, a Newton–Raphson method using the full state transition matrix works well to converge upon a periodic orbit, assuming the initial estimate is chosen sufficiently close to a periodic motion. For perturbed orbital problems, it is often sufficient to choose the motion that would be periodic in the two-body problem as a first estimate for the algorithm.

For a time-invariant system there are additional challenges. First, the period of the orbit is a free parameter. Second, and more importantly, when the solution tends towards a periodic orbit, two eigenvalues of Φ tend towards unity and, if at a periodic orbit, would equal unity. Thus, as the target state is approached the matrix $[I - \Phi]$ becomes singular and the algorithm diverges. One remedy would be to choose the pseudo-inverse of the singular matrix. An alternate, and constructive approach, is to remove the offending unity eigenvalues by restricting the linear variation to a Poincaré map. The advantage of the latter approach is that it defines the monodromy matrix of the periodic orbit once computed. This approach, first detailed in [175], is presented in the following.

6.3.2 Reduction of the State Transition Matrix to a Linear Poincaré Map

The reduction is easiest to derive for a Lagrangian system, but can be generalized for a Hamiltonian system as well. In the following assume that the full Lagrangian state is $\boldsymbol{x} = [\boldsymbol{r}, \dot{\boldsymbol{r}}]$, that the Lagrangian is time-invariant, $L(\boldsymbol{r}, \dot{\boldsymbol{r}})$, and thus a Jacobi integral exists which can be evaluated as the Hamiltonian of the system, $H(\boldsymbol{x}) = \dot{\boldsymbol{r}} \cdot \frac{\partial L}{\partial \dot{\boldsymbol{r}}} - L$, albeit with positions and velocities instead of coordinates and momenta. Assume a simple surface of section, chosen to equal a coordinate

value $S(\boldsymbol{x}) = r_I = 0$, where $I = 1, 2,$ or 3. Transversality in this case means that $S_{\boldsymbol{x}} \cdot \dot{\boldsymbol{x}} = \dot{r}_I \neq 0$. Using the simple example of the surface of section being $r_I = 0$, its associated speed v_I is then eliminated from the Jacobi integral. Note that for most Lagrangian systems encountered in astrodynamics $\frac{\partial H}{\partial v_I} = v_I$ and is non-zero at the surface of section by definition, and thus it is usually simple to eliminate this term. Then the reduced state is $\boldsymbol{y} = [r_J, r_K, \dot{r}_J, \dot{r}_K]$ where $J, K \neq I$ and do not equal each other.

The procedure starts with choosing an initial state $\boldsymbol{x}_0$ such that $S(\boldsymbol{x}_0) = 0$ and $H(\boldsymbol{x}_0) = C$, a specified value of energy. The trajectory is then propagated until a time t_1 when $S(\boldsymbol{x}(t_1)) = 0$. The state transition matrix is propagated simultaneously to find $\Phi(t_1, t_0)$, where this will map deviations in the initial state to deviations in the final state at time t_1, or $\delta\boldsymbol{x}(t_1) = \Phi(t_1, t_0)\delta\boldsymbol{x}_0$. The initial deviation is $\delta\boldsymbol{x}_0 = [\delta r_J, \delta r_K, \delta r_I, \delta v_J, \delta v_K, \delta v_I]$. It is trivial to choose $\delta r_I = 0$, thus keeping the initial state on the surface of section. The energy of the deviated trajectory must also be restricted to a constant value, or $H(\boldsymbol{x}_0 + \delta\boldsymbol{x}_0) = H(\boldsymbol{x}_0)$. Expanding the Hamiltonian and simplifying yields the condition at the linear order $H_{\boldsymbol{x}}|_0 \cdot \delta\boldsymbol{x}_0 = 0$. Expanding this in terms of the reduced state results in $H_{\boldsymbol{y}}\big|_0 \cdot \delta\boldsymbol{y}_0 + H_{v_I}|_0\, \delta v_I = 0$, which can be solved for the necessary variation in the initial velocity v_I to maintain a constant value of energy

$$\delta v_I = -\frac{1}{H_{v_I}|_0}\, H_{\boldsymbol{y}}\big|_0 \cdot \delta\boldsymbol{y}_0 \tag{6.16}$$

Now a linear mapping can be constructed that takes a variation in the reduced state, $\delta\boldsymbol{y}_0$, and maps it into a variation in the initial state $\delta\boldsymbol{x}_0$ that can be mapped forward using the full state transition matrix

$$\delta\boldsymbol{x}_0 = (P_0 + P_H)\, \delta\boldsymbol{y}_0 \tag{6.17}$$

where P_0 and P_H are $2n \times (2n-2)$ matrices in general that enforce the initial projection onto the surface of section and constant energy surface. Specifically the matrices have the following form for the assumed configuration

$$P_0 = \left[\begin{array}{c|c} I_{n-1,n-1} & 0_{n-1,n-1} \\ \hline 0_{1,n-1} & 0_{1,n-1} \\ \hline 0_{n-1,n-1} & I_{n-1,n-1} \\ \hline 0_{1,n-1} & 0_{1,n-1} \end{array}\right] \tag{6.18}$$

$$P_H = -\hat{\boldsymbol{w}}_{I+n}\, H_{\boldsymbol{y}}\big|_0 \tag{6.19}$$

Note that the matrices in P_0 all have their subscripts defined for both the identity matrices and the zero matrices. In the equation for P_H for the example considered, $\hat{\boldsymbol{w}}_{I+n}$ is the zero vector with a 1 in the $I+n$th row and in general is the unit vector along the eliminated velocity component. Then P_H is formed by taking the outer product of the two vectors, creating a matrix of size $2n \times 2(n-1)$.

This vector is then applied to the state transition matrix to find

$$\delta\boldsymbol{x}(t_1) = \Phi(t_1, t_0)\, (P_0 + P_H)\, \delta\boldsymbol{y}_0$$

The state deviation at t_1 automatically satisfies energy conservation, or $H_{\boldsymbol{y}}|_1 \cdot \delta\boldsymbol{x}(t_1) = 0$, and thus the value of $\delta v_I(t_1)$ can be ignored. The same does not hold for the surface of section constraint, however, and $S(\boldsymbol{x}(t_1)+\delta\boldsymbol{x}(t_1)) = S_{\boldsymbol{x}}|_1 \cdot \delta\boldsymbol{x}(t_1) \neq 0$ in general. In fact, for an arbitrary variation of $\delta\boldsymbol{y}_0$ there will only be a line of initial conditions that have the deviated state at t_1 on the surface of section, with the rest of the states either having already passed through the surface or not having reached it yet (see Fig. 6.1). This occurs generically as the time from one surface of section passage to the next is not constant over linear variations, but in general will change with the initial state.

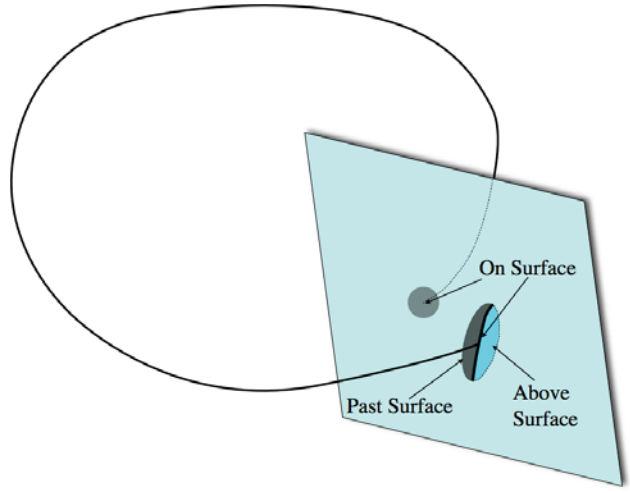


Fig. 6.1 First return map of a solution flow to a surface of section. The initial set is on the surface but the first return at the nominal time only intersects the surface along a line, with the other half of the flow having passed through the surface already and the other not having passed through yet.

To address this allow the time of passage to vary in order to force the deviated trajectory to cross the surface of section

$$S(\boldsymbol{x}(t_1+\delta t)+\delta\boldsymbol{x}(t_1)) = S(\boldsymbol{x}(t_1)) + S_{\boldsymbol{x}}|_1 \cdot \left[\frac{\partial\boldsymbol{x}(t_1)}{\partial t}\delta t + \delta\boldsymbol{x}(t_1)\right] \tag{6.20}$$

Restricting to linear terms, equating this to zero, and noting that $\partial\boldsymbol{x}(t_1)/\partial t = \dot{\boldsymbol{x}}(t_1)$ results in the condition

$$S_{\boldsymbol{x}}|_1 \cdot \dot{\boldsymbol{x}}(t_1)\delta t + S_{\boldsymbol{x}}|_1 \cdot \delta\boldsymbol{x}(t_1) = 0 \tag{6.21}$$

which can be solved for the necessary time variation as a function of the deviated final state

$$\delta t = -\frac{1}{S_{\boldsymbol{x}}|_1 \cdot \dot{\boldsymbol{x}}(t_1)}\, S_{\boldsymbol{x}}|_1 \cdot \delta\boldsymbol{x}(t_1) \tag{6.22}$$

Given the fundamental transversality assumption, $S_{\boldsymbol{x}}|_1 \cdot \dot{\boldsymbol{x}}(t_1) = v_I(t_1) \neq 0$.

Now the final deviated state can be adjusted to ensure that it lies on the surface of section

$$\delta \boldsymbol{x}_1 = \delta \boldsymbol{x}(t_1) + \dot{\boldsymbol{x}}(t_1)\delta t \tag{6.23}$$

$$= \left[I - \frac{1}{S_{\boldsymbol{x}}|_1 \cdot \dot{\boldsymbol{x}}(t_1)} \dot{\boldsymbol{x}}(t_1)\ S_{\boldsymbol{x}}|_1 \right] \cdot \delta \boldsymbol{x}(t_1) \tag{6.24}$$

$$= P_S \cdot \delta \boldsymbol{x}(t_1) \tag{6.25}$$

Note that the final deviation $\delta \boldsymbol{x}_1$ has been defined to denote the deviation that lies on the surface of section. Also, note that the mapping matrix P_S is a $2n \times 2n$ matrix and is full rank in general. Finally, note that the term $\dot{\boldsymbol{x}}(t_1)\ S_{\boldsymbol{x}}|_1$ is an outer product of two vectors.

Thus, the final deviation $\delta \boldsymbol{x}_1$ now both conserves energy and lies completely on the surface of section. Thus, for the final step this vector can be projected onto the reduced state to find the final deviation of the linearized Poincaré map

$$\delta \boldsymbol{y}_1 = P_0^T P_S \Phi(t_1, t_0)\,(P_0 + P_H)\,\delta \boldsymbol{y}_0 \tag{6.26}$$

This has explicitly defined a reduced linear map from the surface of section to the surface of section in the vicinity of the nominal trajectory

$$\Phi_{1,0} = P_0^T P_S \Phi(t_1, t_0)\,(P_0 + P_H) \tag{6.27}$$

At this point, if $\Phi_{1,0}$ is required in a symplectic form, the linear transformation discussed previously from a Lagrangian system to a Hamiltonian can be implemented (Eq. 6.3). It is essential to note that, if evaluated at a fixed point, the matrix $\Phi_{1,0} = \Phi_M$ and will have both of the generic unity eigenvalues removed from it.

6.3.3 General Algorithm Revisited

Having a properly reduced linear Poincaré map the computation scheme can be reconsidered, now restricted to the reduced state. Consider the nonlinear Poincaré map from an initial point $\boldsymbol{y}_0$ and fixed energy C to the next crossing

$$\boldsymbol{y}_1 = \boldsymbol{g}(\boldsymbol{y}_0, C) \tag{6.28}$$

Now, consider a deviation in the initial state that makes this into a fixed point map

$$\boldsymbol{y}_0 + \delta \boldsymbol{y}_0 = \boldsymbol{g}(\boldsymbol{y}_0 + \delta \boldsymbol{y}_0, C) \tag{6.29}$$

$$= \boldsymbol{y}_1 + \Phi_{1,0}\delta \boldsymbol{y}_0 + \ldots \tag{6.30}$$

assuming the algorithm to compute $\Phi_{1,0}$. The initial deviation is chosen to satisfy this equation, leading to

$$\delta \boldsymbol{y}_0 = \left[I - \Phi_{1,0}\right]^{-1} (\boldsymbol{y}_1 - \boldsymbol{y}_0) \tag{6.31}$$

and the test solution is iterated as $\boldsymbol{y}_1 = \boldsymbol{y}_0 + \delta \boldsymbol{y}_0$ and the process repeated. Now, if the initial state is within the basin of attraction of the fixed point, the process will converge and, as discussed previously, the matrix $I - \Phi_{1,0}$ will generically remain non-singular and as $\boldsymbol{y}_0 \to \boldsymbol{y}^*$ then $\Phi_{1,0} \to \Phi_M$, the monodromy matrix.

6.3.4 Stability Computation

Once the fixed point is converged to a specified accuracy the monodromy matrix Φ_M is also automatically defined by reduction of the original state transition matrix. Stability of the resulting fixed point is then determined by computing the eigenvectors and eigenvalues of Φ_M. Again, it is usually easiest to perform such computations using a standard software package.

6.3.5 Families of Periodic Orbits

Having computed a periodic orbit at a particular value of a parameter or energy value, the question arises as to how this single object can be generalized to a larger family of such solutions, generated as the parameter value is changed. From the earlier discussion on family continuation, the first estimate of the new fixed point at energy $C + \delta C$ is given as $\delta \boldsymbol{y}^* + \left[I - \Phi_M\right]^{-1} \boldsymbol{g}_C^* \delta C$. The remaining issue is how to compute

$$\boldsymbol{g}_C^* = \frac{\partial \boldsymbol{g}(\boldsymbol{y}^*, C)}{\partial C}$$

For the example system from above, and for the reduced state in general, there is one coordinate which is eliminated and which has a non-zero gradient along the Hamiltonian. Thus

$$\delta C = \frac{\partial H}{\partial v_I} \delta v_I$$

where the velocity coordinate is chosen to ensure that this gradient is non-zero. As the gradient term is a scalar, it can be solved directly for the necessary variation in the velocity component for a given change in energy,

$$\delta v_I = \frac{1}{H_{v_I}} \delta C$$

Next, note that the velocity v_I can be varied independently of the reduced state variation, which can remain zero in general. Thus, the deviation in the final reduced state is equal to

$$\delta \boldsymbol{y}_1 = P_0^T P_S \Phi(t_1, t_0) \hat{\boldsymbol{w}}_{I+3} \frac{\delta C}{H_{v_I}} \tag{6.32}$$

where the unit vector $\hat{\boldsymbol{w}}_{I+3}$ is as described before. From this the ratio $\delta \boldsymbol{y}_1/\delta C$ can be directly computed to find the necessary partial

$$\boldsymbol{g}_C^* = P_0^T P_S \Phi(t_1, t_0) \hat{\boldsymbol{w}}_{I+3} \frac{1}{H_{v_I}} \tag{6.33}$$

It is also possible to continue periodic orbit families in parameters of the equations of motion. The procedure is largely the same as outlined here for tracing out a family as a function of energy, except now the partial of the state with respect to the parameter must be determined. In Chapter 5 the defining differential equation for this partial is described, and following its computation it can also be projected to the surface of section using the same methodology as outlined here.

6.4 Semi-Analytical Solutions: Higher-Order Solution Expansions

An alternate solution approach is to take a nominal, numerical solution of the system as the known solution and expand about it using an assumed form for the relative solution. This is exactly what is done when a linear expansion of an orbit is performed; in the following this approach is generalized and continued to higher orders in a systematic fashion. The following is largely borrowed from the derivation given in [124].

Assume a standard form for the equations of motion, $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, t)$, and that $\boldsymbol{x} \in \boldsymbol{R}^{2n}$ and $\boldsymbol{f} : \mathbf{R}^{2n} \times \mathbf{R} \to \mathbf{R}^{2n}$. Then the solution of these equations can be represented as $\boldsymbol{x}(t) = \boldsymbol{\phi}(t; \boldsymbol{x}_o, t_o)$ and $\boldsymbol{\phi}$ is analytic in the initial conditions for $|t| < \infty$. It is not necessary to assume any special conditions on the solution flow, such that it be a periodic orbit, be an equilibrium point or lie on a manifold. All of these conditions can be modeled by this theory, however.

The goal is to expand the solution as a Taylor Series in terms of its initial conditions, with the coefficients being time-varying functions. Introduce tensor notation to specify the equations of motion and solution as

$$\dot{x}_i = f_i(\boldsymbol{x}, t) \tag{6.34}$$

$$x_i(t) = \phi_i(t; \boldsymbol{x}_o, t_o) \tag{6.35}$$

$$i = 1, 2, \ldots, 2n \tag{6.36}$$

where $\boldsymbol{x} = (x_1, x_2, \ldots, x_{2n})$ is kept as a convenient way to denote the entire state. For the expansion introduce a variation about the initial conditions, $\boldsymbol{x}_o + \delta \boldsymbol{x}_o$ which is introduced in the solution function. Carrying out the expansion for $\boldsymbol{\phi}(t; \boldsymbol{x}_o + \delta \boldsymbol{x}_o, t_o) - \boldsymbol{\phi}(t; \boldsymbol{x}_o)$ leads to

$$\delta x_i(t) = \sum_{p=1}^{\infty} \frac{1}{p!} \phi_{i, j_1 j_2 \ldots j_p}(t, t_o) \delta x_{j_1}^o \delta x_{j_2}^o \cdots \delta x_{j_p}^o \tag{6.37}$$

where the Einstein summation convention is assumed for the repeated indices in the above summation, i.e.,

$$\phi_{i,j}\delta x_j^o = \sum_{j=1}^{2n} \phi_{i,j}\delta x_j^o \quad \text{and} \quad \phi_{i,j_1 j_2 \ldots j_p} = \frac{\partial^p \phi_i(t; \boldsymbol{x}_o, t_o)}{\partial x_{j_1}^o \partial x_{j_2}^o \cdots \partial x_{j_p}^o}$$

Note that the series is formal, but due to the properties of ordinary differential equations it is expected to converge for small enough $\delta \boldsymbol{x}_o$ when evaluated over a finite time, providing an explicit equation for $\delta x_i(t)$ as a function of δx_j^0.

Similarly, expand the equations of motion evaluated at a time t in terms of the above $\delta x_i(t)$, for $\boldsymbol{f}(\boldsymbol{x}(t) + \delta \boldsymbol{x}(t), t) - \boldsymbol{f}(\boldsymbol{x}(t), t)$ to find

$$\delta \dot{x}_i(t) = \sum_{p=1}^{\infty} \frac{1}{p!} f_{i,j_1 j_2 \ldots j_p}(\boldsymbol{x}(t), t) \delta x_{j_1}(t) \delta x_{j_2}(t) \cdots \delta x_{j_p}(t) \tag{6.38}$$

In a final step substitute Eq. 6.37 into Eq. 6.38 to derive expressions for the time derivatives of the coefficients $\phi_{i,j_1 j_2 \ldots j_p}(t, t_o)$ so they can be posed as solutions to a series of equations of motion. Now balance orders of the initial conditions, realizing that they are the constants of motion for this solution procedure and that the resulting expansions must balance at each order. Carrying this out recovers the expressions in [124], which are listed through third order as

$$\dot{\phi}_{i,j} = f_{i,\alpha}(t)\phi_{\alpha,j} \tag{6.39}$$

$$\dot{\phi}_{i,jk} = f_{i,\alpha}(t)\phi_{\alpha,jk} + f_{i,\alpha\beta}(t)\phi_{\alpha,j}\phi_{\beta,k} \tag{6.40}$$

$$\begin{aligned} \dot{\phi}_{i,jkl} = {} & f_{i,\alpha}(t)\phi_{\alpha,jkl} + f_{i,\alpha\beta}(t)\left(\phi_{\alpha,j}\phi_{\beta,kl} + \phi_{\alpha,jk}\phi_{\beta,l} + \phi_{\alpha,jl}\phi_{\beta,k}\right) \\ & + f_{i,\alpha\beta\gamma}(t)\phi_{\alpha,i}\phi_{\beta,j}\phi_{\gamma,k} \end{aligned} \tag{6.41}$$

The first equation is just the State Transition Matrix, while the higher-order equations are generalizations of this, and thus are called the State Transition Tensors (STTs). The convergence properties of these solutions are tested in [124] and shown to be quite effective in modeling the nonlinear dynamics relative to a nominal trajectory. All of the higher-order partials are symmetric in their indices, i.e., $\phi_{i,\alpha\beta\gamma}(t, t_o)$ and $f_{i,\alpha\beta\gamma}(t)$ are symmetric in α, β and γ in general.

Having the differential equations defined for the STTs all that is needed are their initial conditions to carry out numerical integrations. These are quite trivial

$$\phi_{i,j}(t_o, t_o) = \delta_{ij} \tag{6.42}$$

$$\phi_{i,j_1 j_2 \ldots j_p}(t_o, t_o) = 0 \text{ for } p \geq 2 \tag{6.43}$$

It is most efficient to integrate the nominal trajectory $\phi_i(t; \boldsymbol{x}_o, t_o)$ in tandem with the STTs, as the entire set of equations then form a closed system. The computational cost of this approach is not trivial. Including the nominal state the number of differential equations to compute for a first-order expansion is $2n(2n+1)$, at second order is $2n\left(2n^2 + 3n + 1\right)$, accounting for the symmetric in the higher-order partials, and continues to scale as $(2n)^4$ at third order, etc.

Depending on the specific form of the equations of motion many of the dynamics terms $f_{i,j_1 j_2 \ldots j_p}$ may be identically zero above a certain index value. For example, if the system is modeled using Newton's Equations, $\boldsymbol{f}(\boldsymbol{x}, t) = [\boldsymbol{v}; U_{\boldsymbol{r}}(\boldsymbol{r})]$, the higher-order expansions would follow

$$f_{i,i+3} = 1 \text{ for } i = 1, 2, 3 \tag{6.44}$$

$$f_{i+3,k_1,k_2,\ldots,k_p} = U_{x_i,x_{k_1},x_{k_2},\ldots,x_{k_p}} \text{ for } i, k_i = 1, 2, 3 \tag{6.45}$$

$$f_{i,j_1 j_2 \ldots j_p} = 0 \text{ for all other } j_p \text{ for } i = 1, 2, 3 \tag{6.46}$$

and

$$f_{i,j_1 j_2 \ldots j_p} = 0 \text{ for } i, j_i = 4, 5, 6 \tag{6.47}$$

6.5 Analytic Solutions

There is a long history of analytical solution methods that arises out of Celestial Mechanics. By an analytic solution we mean that the solution to the equations of motion can be expressed in terms of analytical functions expanded in some parameter about a known solution. The use of analytical solutions are several. First, an analytical solution evaluated to a high degree of precision essentially replaces a numerical integration with a closed-form function evaluation. This is the fundamental and original motivation behind development of these solutions, and it has been developed to extremely high levels of precision in the study of motions in the restricted three-body problem [51]. The existence of an analytic solution of this form enables the structure of motion in these problems to be studied in detail.

Further motivations include the development of approximate solutions valid over extremely long timespans, and is considered further in the following section when we discuss secular solutions. Development of analytical solutions enable much more difficult problems to be considered, such as mapping orbit uncertainty forward in time. Specifically, given an analytical solution, it becomes possible to solve the diffusionless Fokker–Planck–Kolmogorov Equation to find the probability density function of particle as a function of time. Additionally, it becomes simple to explore the neighborhood of a solution given an analytical approximation to it, enabling applications to navigation and orbit determination techniques.

Two approaches are discussed that develop analytical solutions by very different methods. The first applies analytic continuation and follows the classical development by Moulton. In this approach the solution is built up at each order by solving a quadrature. The other approach is the Lie–Deprit algorithm, which has a basis in von Zeipel's method. This method enforces the condition that the solution is constructed so that it is a canonical solution, and hence preserves the Hamiltonian constraints of the original problem statement. We do not discuss this method in detail, but instead refer to some appropriate texts for its exposition.

6.5.1 Moulton's Method of Analytic Continuation

The following is modified from Moulton [111], consult that text for a more rigorous description of the convergence of the series involved, plus a more in-depth description of this general solution method and its variations.

Assume a general differential equation of the form:

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, t) + \epsilon \boldsymbol{g}(\boldsymbol{x}, t) \tag{6.48}$$

where the problem can be completely solved when $\epsilon = 0$. Denote the general solution to the problem as $\boldsymbol{x}(t) = \boldsymbol{\phi}(t; \boldsymbol{x}_o, t_o, \epsilon)$, where ϵ is retained as a parameter. Denote the known, analytical solution when $\epsilon = 0$ as $\boldsymbol{x}^{(0)} = \boldsymbol{\phi}(t; \boldsymbol{x}_o, t_o, 0)$ and note that this is not necessarily a Hamiltonian system.

By hypothesis, the solutions to this full system of equations are analytic in ϵ, and thus the solution can be expanded as an analytic function of ϵ. The term ϵ need not be a physically defined parameter, and can just represent the fact that the order of magnitude of $|\boldsymbol{g}| \ll |\boldsymbol{f}|$ in general. Specifically one can expand the general solution into a Taylor series about the known solution

$$\boldsymbol{x}(t) = \boldsymbol{x}^{(0)} + \epsilon \boldsymbol{x}^{(1)} + \frac{1}{2!}\epsilon^2 \boldsymbol{x}^{(2)} + \dots \tag{6.49}$$

We do not try to express what the radius of convergence of this series is, but note that for an analytic and well-behaved dynamical solution over a finite timespan that such a radius of convergence will exist.

The derivation of the method is simple, one just substitutes the above general form of the solution into the full equations of motion, expands them in orders of ϵ, and then balances terms at each order of epsilon, which is where analyticity is assumed. Skipping some of the involved details of the expansion, carrying it out and comparing terms at orders ϵ^0, ϵ^1 and ϵ^2 yields the results:

$$\epsilon^0 : \ \dot{\boldsymbol{x}}^{(0)} = \boldsymbol{f}(\boldsymbol{x}^{(0)}, t) \tag{6.50}$$

$$\epsilon^1 : \ \dot{\boldsymbol{x}}^{(1)} = \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right|_0 \boldsymbol{x}^{(1)} + \boldsymbol{g}(\boldsymbol{x}^{(0)}, t) \tag{6.51}$$

$$\epsilon^2 : \ \dot{\boldsymbol{x}}^{(2)} = \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right|_0 \boldsymbol{x}^{(2)} + 2 \left.\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{x}}\right|_0 \boldsymbol{x}^{(1)} + \boldsymbol{p}(\boldsymbol{x}^{(0)}, \boldsymbol{x}^{(1)}, t) \tag{6.52}$$

and at the kth level has the same general form:

$$\begin{aligned} \epsilon^k : \ \dot{\boldsymbol{x}}^{(k)} = & \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right|_0 \boldsymbol{x}^{(k)} + k \left.\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{x}}\right|_0 \boldsymbol{x}^{(k-1)} \\ & + \boldsymbol{p}(\boldsymbol{x}^{(0)}, \boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(k-1)}, t) \end{aligned} \tag{6.53}$$

where the $\boldsymbol{p}$ are functions that arise from the higher-order partials of the equations of motion and are all functions of the expansion at lower orders. These are worked

out to the second and sometimes higher order in Moulton, but can also be computed using a symbolic manipulator.

Then, assuming that the solution $\boldsymbol{x}^{(0)}(t)$ is known in closed form, each order equation only involves linear terms at that order with all nonlinear terms appearing at lower order, meaning that the problem can be solved by quadratures once a solution for the homogeneous linear term is found. For the case where the solution $\boldsymbol{x}^{(0)}$ is constant, such as occurs for orbit element formulations, the linear equations can be solved in closed form with the exponential matrix, and if $\boldsymbol{f} \equiv 0$ the entire solution process can be reduced to quadratures.

Let us specifically consider the case when $\boldsymbol{f} \equiv 0$, then the solution to these equations can be expressed in a concise, recursive form. First note that the solution at order 0 is just the initial conditions, $\boldsymbol{x}^{(0)} = \boldsymbol{x}_o$, implying that the initial conditions for all higher-order terms are identically zero, $\boldsymbol{x}^{(k)}(t_o) = 0$, $k \geq 1$. Thus the solutions have the form:

$$\boldsymbol{x}^{(1)}(t, t_o) = \int_{t_o}^{t} \boldsymbol{g}(\boldsymbol{x}_o, \tau)\, d\tau \tag{6.54}$$

$$\boldsymbol{x}^{(2)}(t, t_o) = \int_{t_o}^{t} 2 \left.\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{x}}\right|_o \boldsymbol{x}^{(1)}(\tau, t_o)\, d\tau \tag{6.55}$$

$$\begin{aligned} \boldsymbol{x}^{(k)}(t, t_o) = \int_{t_o}^{t} \Bigg[k \left.\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{x}}\right|_0 \boldsymbol{x}^{(k-1)}(\tau, t_o) \\ + \boldsymbol{p}(\boldsymbol{x}^{(0)}, \boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(k-2)}, \tau, t_o) \Bigg]\, d\tau \end{aligned} \tag{6.56}$$

While implementation of this method can be cumbersome beyond the first order, it provides a systematic method to describe these higher-order dynamics.

6.5.2 Generating Function Approaches

These approaches generally rely on the canonical form of the equations of motion and state the problem as a general expansion of a generating function, defined for all orders of perturbation. At each order of expansion one finds partial differential equations that must be solved. These partial differential equations are usually simplified by eliminating fast variables, retaining only the more slowly varying terms. Then the solution procedure is advanced to the next order, systematically building up a higher-order solution. The von Zeipel method is the first widely applied version of this solution technique, and is detailed in [19]. An important innovation for this theory was introduced by Deprit, through the use of Lie transforms, which eliminated the need to perform analytical operations on the von Zeipel solution and enabled the direct solution of the initial conditions. A detailed discussion of the Lie–Deprit method is given in [13]. The manner in which these theories were applied took fundamental advantage of the periodic motion of the unperturbed two-body system to construct generating functions that eliminated such periodic angles at subsequently higher orders. In this way, these analytical expansion theories essen-

tially perform high-order averaging of a dynamical system and, at leading order, provide similar results to those discussed next for the mean motion derivations.

Despite the rigor and advantages of these methods, we do not present the theory in detail, as this would require significant additional discussions that stray from the core purpose of this text. It must be noted that these expansion techniques have not been widely applied to the problem of small-body orbiters to date, except to the motion about planetary satellites [90], and may represent a new and fruitful application of this theory. Possible drawbacks include the relative strength of the perturbations, which may require high-order expansions in order to fully capture the effects analytically.

6.6 Mean Motion Derivations

A particularly effective method for the approximate evaluation of the dynamics of a system is through the use of averaged equations. Once derived, such averaged equations can themselves be integrated numerically, or can often be solved in closed form. The Lie–Deprit method taken at the first order alone provides what can be considered to be an "averaged" form of the equations of motion, as it provides a systematic way of removing the fast variables while retaining the slower variables of the system. This approach has the advantage of allowing the solutions to be expanded to higher orders, as has been investigated in numerous studies. However, as indicated above this methodology has considerable theoretical overhead and is beyond the scope of the current text, which is focused on more direct and practical approaches for describing orbital motion about small bodies.

The method presented below, in contrast, develops the averaged equations of motion at first order and discusses properties of these solutions. These averaged equations of motion capture the secular evolution of the system and can either be numerically integrated or in some cases solved in closed form. The advantage of this approach is that it is possible to easily capture the qualitative effect of perturbations over long timespans with significantly reduced computational requirements. Similar approaches have been systematically developed to higher orders by Laskar [92] for investigating the long-term stability of the solar system.

These methods are generally applied to orbit element formulations such as the Lagrange Planetary Equations or the Gauss Equations. Thus, the dynamical system of interest for this approach has a simplified form as compared to the general system for Moulton's analytical theory

$$\dot{\boldsymbol{x}} = \epsilon \boldsymbol{g}(\boldsymbol{x}, t) \tag{6.57}$$

where when $\epsilon = 0$ the resulting equation is trivially solved by specifying some set of initial conditions, $\boldsymbol{x}_0$, which one can think of as being the orbit elements of an unperturbed satellite in motion about a point mass body. In developing the acceleration perturbations $\boldsymbol{g}(\boldsymbol{x}_0, t)$, even though the nominal solution consists of constants it is usually necessary to map this solution into a Cartesian frame which specifies the motion of the satellite relative to the system as a function of time.

This is generally where the time-variation in the equations of motion arises from. The following development makes the tacit assumption that, all else being equal, the resulting perturbing acceleration $\boldsymbol{g}(\boldsymbol{x}_0, t)$ will be periodic or quasi-periodic in time t. In the application section there are explicit examples of these perturbations and the relevant assumptions are discussed again at that time.

6.6.1 First-Order Solutions

Now apply Moulton's theory to first order to find the approximate evolution of the system as a function of time. If the perturbation is relatively small, then these changes may not be dramatic, although they will systematically change the state of the system. Writing out the general expression for the first-order solution, ignoring the parameter ϵ, yields

$$\dot{\boldsymbol{x}}_1 = \boldsymbol{g}(\boldsymbol{x}_0, t) \tag{6.58}$$

where the analytic continuation results are used, meaning that $\boldsymbol{x}_0$ is the unperturbed motion. For orbital systems, as discussed above, one can assume that $\boldsymbol{g}(\boldsymbol{x}_0, t+T) = \boldsymbol{g}(\boldsymbol{x}_0, t)$, where T is generally the orbit period of the unperturbed motion or is related to it. There are often other time-varying quantities in these equations, in computing the periodicity one often assumes that these times are held fixed, and may be averaged over later under some mild conditions. The solution for $\boldsymbol{x}_1$ can be explicitly written out as

$$\boldsymbol{x}_1(t) - \boldsymbol{x}_1(0) = \int_0^t \boldsymbol{g}(\boldsymbol{x}_0, \tau)\, d\tau \tag{6.59}$$

where for convenience we take $t_o = 0$. Then over a period $t + T$

$$\boldsymbol{x}_1(t+T) - \boldsymbol{x}_1(0) = \int_0^{t+T} \boldsymbol{g}(\boldsymbol{x}_0, \tau)\, d\tau \tag{6.60}$$

$$= \int_0^T \boldsymbol{g}(\boldsymbol{x}_0, \tau) d\tau + \int_T^{T+t} \boldsymbol{g}(\boldsymbol{x}_0, \tau)\, d\tau \tag{6.61}$$

$$= \boldsymbol{x}_1(T) + \int_0^t \boldsymbol{g}(\boldsymbol{x}_0, \tau)\, d\tau \tag{6.62}$$

leading to

$$\boldsymbol{x}_1(t+T) - \boldsymbol{x}_1(t) = \boldsymbol{x}_1(T) \tag{6.63}$$

Thus, at the level of the approximation a propagation n orbits into the future yields:

$$\boldsymbol{x}_1(t+nT) = \boldsymbol{x}_1(t) + n\boldsymbol{x}_1(T) \tag{6.64}$$

meaning that, to the leading order, the long-term effects can be captured by understanding the change in $\boldsymbol{x}_1$ over one orbit period of the nominal system. Indeed, it now is possible to take the time average of $\boldsymbol{x}_1$ to understand the predicted long-term behavior:

$$\bar{\boldsymbol{x}}_1 = \lim_{t\to\infty} \frac{1}{t} \int_0^t \boldsymbol{x}_1(\tau)\, d\tau \tag{6.65}$$

$$= \lim_{n\to\infty} \frac{1}{t' + nT} \left[\int_0^{t'} \boldsymbol{x}_1(\tau)\, d\tau + n \int_0^T \boldsymbol{x}_1(\tau)\, d\tau \right] \tag{6.66}$$

$$= \frac{1}{T} \int_0^T \boldsymbol{x}_1(\tau)\, d\tau \tag{6.67}$$

Thus, to initially understand the long-term evolution of the system one only needs to study its evolution over one cycle using the approximate map

$$\boldsymbol{x}(t + nT) = \boldsymbol{x}_0 + n\boldsymbol{x}_1(T) + \int_0^t \boldsymbol{g}(\boldsymbol{x}_0, \tau)\, d\tau \tag{6.68}$$

where $t \leq T$. Such extrapolation of this result is inaccurate, however, since the integral $\boldsymbol{x}_1(T)$ depends strongly on the nominal value $\boldsymbol{x}_0$, and if the full solution is diverging from this initial value the integral will become less and less relevant. This can be fixed by replacing the simple extrapolation of the analytical solution with a more sophisticated mapping approach, updating the nominal solution at every step in time. To represent this dependence a more precise notation is introduced, $\boldsymbol{x}_1(T; \boldsymbol{x}_0)$ which shows that the perturbed solution depends explicitly on the nominal state.

Then the sequence in which the nominal solution is updated using $\boldsymbol{x}_1$ is

$$\boldsymbol{x}(T) = \boldsymbol{x}_0 + \boldsymbol{x}_1(T; \boldsymbol{x}_0) \tag{6.69}$$

$$\boldsymbol{x}(2T) = \boldsymbol{x}(T) + \boldsymbol{x}_1(T; \boldsymbol{x}(T)) \tag{6.70}$$

$$\boldsymbol{x}((n+1)T) = \boldsymbol{x}(nT) + \boldsymbol{x}_1(T; \boldsymbol{x}(nT) \tag{6.71}$$

This directly gives the expected change in the state over one orbit period at an arbitrary step n as

$$\boldsymbol{x}((n+1)T) - \boldsymbol{x}(nT) = \boldsymbol{x}_1(T; \boldsymbol{x}(nT)) \tag{6.72}$$

or

$$\Delta\boldsymbol{x}(nT) = \boldsymbol{x}_1(T; \boldsymbol{x}(nT)) \tag{6.73}$$

This can be interpreted as a net rate of change of the state over one orbit period

$$\frac{\Delta \boldsymbol{x}(nT)}{T} = \frac{1}{T}\boldsymbol{x}_1(T; \boldsymbol{x}(nT)) \tag{6.74}$$

$$= \frac{1}{T}\int_0^T \boldsymbol{g}(\boldsymbol{x}(nT), \tau)\, d\tau \tag{6.75}$$

$$= \bar{\boldsymbol{g}}(\boldsymbol{x}(nT)) \tag{6.76}$$

and this is interpreted as the average rate of change in the state, where now that rate of change can shift from orbit to orbit. Abstracting this to a continuous time system yields:

$$\dot{\boldsymbol{x}} = \bar{\boldsymbol{g}}(\boldsymbol{x}) \tag{6.77}$$

where the function $\bar{\boldsymbol{g}}$ has had its time variation due to orbital motion removed and is computed as

$$\bar{\boldsymbol{g}}(\boldsymbol{x}) = \frac{1}{2\pi}\int_0^{2\pi} \boldsymbol{g}(\boldsymbol{x}, M)\, dM \tag{6.78}$$

where the time is replaced with the mean motion, $M = nt$, and the period by $T = 2\pi/n$, essentially assuming that all averaging will be over orbital motion. In this sense this process is equivalent to the first-order Lie Series equations where the "fast" variable has been eliminated. From the analytic continuation side this represents the average rate of change of the system and defines a new nonlinear set of equations for the propagation of the orbit state.

6.6.2 Short-Period Terms

Now consider the solution over one time period of the unperturbed problem. Then the actual solution should be represented with two terms, a secular (or averaged) term represented above, which varies linearly with time over one orbit period, and a periodic term

$$\boldsymbol{x}(t) = \boldsymbol{x}_o + \bar{\boldsymbol{g}}t + \boldsymbol{x}_p(t) \tag{6.79}$$

for $0 \leq t \leq T$. Note that it is not needed to keep $\bar{\boldsymbol{g}}$ constant, but this eases the derivation. Then the periodic part of the solution $\boldsymbol{x}_p(t)$ repeats itself over one orbit, yielding the proper change in the state over one period. Taking the time derivative of this relationship yields:

$$\boldsymbol{g}(\boldsymbol{x}, t) = \bar{\boldsymbol{g}}(\boldsymbol{x}) + \dot{\boldsymbol{x}}_p(t) \tag{6.80}$$

which provides an explicit differential equation for the periodic term:

$$\dot{\boldsymbol{x}}_p = \boldsymbol{g}(\boldsymbol{x}_o, t) - \bar{\boldsymbol{g}}(\boldsymbol{x}_o) \tag{6.81}$$

where only the constant terms inside of the dynamics equations are retained, essentially substituting the nominal initial condition $\boldsymbol{x}_o$ for the full solution. Then a quadrature can be carried out for $\boldsymbol{x}_p$ to find

$$\boldsymbol{x}_p(t) = \int_0^t \boldsymbol{g}(\boldsymbol{x}_o, \tau)\, d\tau - t\bar{\boldsymbol{g}}(\boldsymbol{x}_o) \tag{6.82}$$

This provides an estimate of the periodic solution relative to the secular solution within each time interval, and will be explicitly used in the following.

The solutions to the secular equations given above represent the average evolution of the dynamical system, but comparison between numerical integrations and these secular equations, when starting from the same initial conditions, will in general show a divergence between the two. While this may be due to neglected second-order effects for a strongly perturbed system, often this results because of an inconsistent choice of initial conditions. Carefully considering the original derivation of the mean state $\bar{\boldsymbol{x}}$, it is important to note that the limit $t \to \infty$ was taken. In the later derivation of the secular rates of change of these elements, however, we necessarily limited ourselves to a single cycle of the nominal system. Returning to the decomposition of motion into a mean part and a periodic part, it becomes clear that over a single cycle there may be a non-trivial contribution to the average from the periodic part which does not grow in time, yet which can shift the mean value of the state starting from some initial condition $\boldsymbol{x}_o$. Thus, taking the average of a solution decomposed as Eq. 6.79 over a single cycle yields:

$$\bar{\boldsymbol{x}} = \boldsymbol{x}_o + \frac{1}{2}\bar{\boldsymbol{g}}T + \bar{\boldsymbol{x}}_p \tag{6.83}$$

The average of the periodic part is non-zero and equals:

$$\bar{\boldsymbol{x}}_p = \frac{1}{T}\int_0^T \left[\int_0^{\tau'} \boldsymbol{g}(\boldsymbol{x}_o, \tau)\, d\tau - \tau'\bar{\boldsymbol{g}}(\boldsymbol{x}_o)\right] d\tau' \tag{6.84}$$

$$= \frac{1}{T}\int_0^T \int_0^{\tau'} \boldsymbol{g}(\boldsymbol{x}_o, \tau)\, d\tau\, d\tau' - \frac{T}{2}\bar{\boldsymbol{g}}(\boldsymbol{x}_o) \tag{6.85}$$

So given an initial condition for the dynamical system, $\boldsymbol{x}_o$, which is used in the numerical integrations, the average value of the state over a single period is offset from this and equals:

$$\bar{\boldsymbol{x}} = \boldsymbol{x}_o + \frac{1}{T}\int_0^T \int_0^{\tau'} \boldsymbol{g}(\boldsymbol{x}_o, \tau)\, d\tau\, d\tau' \tag{6.86}$$

The interpretation of this result is that given an initial condition for a state $\boldsymbol{x}_o$ the secular equations should be initialized at the above value of $\bar{\boldsymbol{x}}$ for the mean equations to track the true evolution more closely. This small correction can often reduce deviation between secular and true solutions by a significant amount.

6.6.3 Multiple Averaging Time Scales

For some problems of interest, there may be multiple time scales or frequencies over which relevant dynamical motion occurs. Examples include planetary satellite orbiters, where the orbit period of the spacecraft is often much smaller than the orbit period of the planetary satellite (see Chapter 17). Another example would be orbital dynamics about a slowly rotating asteroid, as again the orbit period may be much shorter than the rotation period [72].

In these situations the two time scales are defined by an orbital rate, n, and the motion of external bodies in the system with their own angular rate N. If $N/n \ll 1$ it is acceptable to hold N constant while an initial average is taken over n. Then the remaining system has a time-varying term associated with N, in general periodic. This term can also be averaged over, usually yielding a greatly simplified set of equations. The usual stipulation is that the time scales between the two periods are sufficiently far apart. The degree of separation that is sufficient should be studied on a case-by-case basis.

6.6.4 Averaging the Lagrange Planetary Equations

A special mention must be made of averaging applied to the Lagrange Planetary Equations. These equations have the general form $\boldsymbol{g}(\boldsymbol{x}, t) = G(\boldsymbol{x})R_{\boldsymbol{x}}(\boldsymbol{x}, t)$, where $G(\boldsymbol{x})$ is a matrix and does not contain time explicitly. Thus, averaging applied to these equations can be drawn all the way back to the perturbing potential $R(\boldsymbol{x}, t)$, greatly simplifying the derivation of the mean equations. The average potential is defined as

$$\bar{R}(\boldsymbol{x}) = \frac{1}{2\pi}\int_0^{2\pi} R(\boldsymbol{x}, M)\, dM \tag{6.87}$$

where the averaging is generally expressed in terms of mean anomaly. The resulting secular equations are then $\dot{\bar{\boldsymbol{x}}} = G(\boldsymbol{x})\bar{R}_{\boldsymbol{x}}(\bar{\boldsymbol{x}})$. One main consequence of this approach is that the mean anomaly is removed from the perturbation potential and thus the semi-major axis is constant in the averaged motion and serves as an integral of motion for these equations. If the original system has a Jacobi integral, then that function is still conserved after the averaging and defines an integral of motion distinct from the semi-major axis.

If the singly averaged potential is a function of a second periodic term of period T, or $R(\boldsymbol{x}, t, M) = R(\boldsymbol{x}, t + T, M)$, and sufficient distance between the two time scales exists, another averaging may be performed to yield:

$$\bar{\bar{R}}(\boldsymbol{x}) = \frac{1}{T2\pi}\int_0^{T}\int_0^{2\pi} R(\boldsymbol{x}, \tau, M)\, dM\, d\tau \tag{6.88}$$

Specific applications of this in will be made in Chapter 17.

6.7 Discrete Orbit Updates

When a dynamical system has a time-varying quantity, an averaging analysis generally holds this time-varying quantity stationary while performing its quadrature over the motions of a satellite. For many systems this is a reasonable approximation which allows the equations to be simplified. Important situations can arise, however, where this is not a good approximation. Specific examples usually include systems where the time variation is of similar magnitude as the orbital variation in time, such as a spacecraft orbiting close to a rotating asteroid. For these systems it is not appropriate to hold the asteroid stationary during the averaging process as significant interactions between the two systems result, yielding large changes in quantities such as the energy or angular momentum.

It is desirable to capture these effects analytically or semi-analytically, as these approaches generally provide more insight into the factors of the system which control the changing orbit. To analyze these effects it is useful to state the dynamical system in terms of the Lagrange Planetary Equations. Assume that the perturbation function has an explicit time variation, $R(\boldsymbol{x}, t)$. If the system has significant interaction over one orbit, then averaging the system can mask these interactions. To capture them consider the Moulton analytic theory to the first order, expressed as a quadrature:

$$\boldsymbol{x}_1(t) = \int_{t_o}^{t} \boldsymbol{f}(\boldsymbol{x}_0(\tau), \tau)\, d\tau \tag{6.89}$$

where both the orbital motion and the physical system are allowed to vary simultaneously. If this is performed over one orbit then one finds explicit formula for the change in the state over this interaction.

$$\Delta\boldsymbol{x} = \int_{t_o}^{t_o+T} \boldsymbol{f}(\boldsymbol{x}_0(\tau), \tau)\, d\tau \tag{6.90}$$

where T is the orbit period. Formally, it is then possible to update the nominal state and re-integrate over a given time-span, leading to a discrete map approximation [130].

These quadratures generally cannot be evaluated in closed form, essentially due to the explicit dependence of the two-body solution on true anomaly and the dependance of the physical system on time. For an analytic expression of the potential, however, it is usually possible to isolate these quadratures as a function of semi-major axis and eccentricity alone, allowing the dependance of the other orbital elements to remain clear. Specific applications of this approach are presented in Chapter 7. Other examples of this approach can be found in [187, 129].

6.8 Phase Space Constraints

The final, and in some sense most precise, constraints that can be placed on a solution arise from the topology of conserved quantities. The most common approach to developing a hard constraint on a trajectory is the use of zero-velocity surfaces. These arise in systems with a Jacobi integral and allow one to specifically map out regions where it is impossible for a particle to enter, as a function of the Jacobi integral value.

For systems where a Jacobi integral exists the generic form that these integrals take is

$$J(\boldsymbol{r}, \boldsymbol{v}) = \frac{1}{2}v^2 - V(\mathbf{r}) \tag{6.91}$$

where V is the potential and $v = |\boldsymbol{v}|$ is the speed relative to the frame in which all time variation in the system is removed (usually a uniformly rotating frame). The concept is quite simple: for a given value of the Jacobi integral, C, the inequality $v^2 \geq 0$ must hold, leading to the more general constraint

$$C + V(\boldsymbol{r}) \geq 0 \tag{6.92}$$

Thus, this provides explicit constraints on the position of the particle, independent of the speed, for a given constant C, which is $V(\boldsymbol{r}) \geq -C$.

A simple example of this is found in the two-body problem and its energy integral, $E = 1/2v^2 - \mu/r$. Now note that the potential is negative definite, with the resulting inequality being expressed as $\mu/r \geq -E$. Next consider the different possibilities, that $E < 0$, $E = 0$ and $E > 0$. First consider $E = 0$, yielding the inequality $\mu/r \geq 0$. This provides no constraint on the system, as the quantity on the left is always positive and only approaches 0 as $r \to \infty$. Similarly for the case $E > 0$, which leads to the inequality $\mu/r \geq 0 \geq -E$. Thus, in these situations there are no constraints placed on the motion, even though additional constraints can be developed for the special case of the two-body problem. Use of zero-velocity curves generally has such null results for specific values of the constant. This is due to the potentials usually being definite in sign, meaning that the inequality will often be trivially true.

More interesting for the two-body problem is the case of negative energy, $E < 0$, as the inequality then becomes $\mu/r \geq -E > 0$, meaning that there are values of r which violate the inequality. Specifically, the system is restricted to $r \leq \mu/|E|$, providing an upper bound on the distance that the solution can get from the origin. While these results are somewhat trivial for the two-body problem, they also apply for more complex systems that do not have specific solutions, such as the restricted three-body problem.

Generally, the presence of additional integrals of motion allow for additional restrictions on the motion, without having to solve for the motion specifically. Although often these restrictions must be combined with some other integral, such as the Jacobi or energy integrals. A good example of this can be found when some component of the angular momentum is conserved. Let's assume that the system

conserves angular momentum with a relationship of the form $h = r^2\dot{\theta} = rv\cos\gamma$, and that it has a Jacobi integral of the form $v^2 = 2C + 2V(\boldsymbol{r})$. The presence of such an additional relation allows us to eliminate the speed, v, and transform the angular momentum integral to $h = r\sqrt{2C + 2V(\boldsymbol{r})}\cos\gamma$. Note that the previous zero-velocity constraint can assure us that only regions where $V + C \geq 0$ are considered. If $h \neq 0$, then the sign of the expression cannot change in general and it can be noted that $\cos\gamma \neq 0$ unless $r \to \infty$, and thus that γ lies in the interval $(-\pi/2, \pi/2)$. Now, let us assume that the trajectory has a closest or furthest approach to the origin, i.e., a periapsis or apoapsis passage. Geometrically it is known that $\gamma = 0$ at these points which leads to the relation $h = r_{apse}\sqrt{2C + 2V(\boldsymbol{r}_{apse})}$, implying that there are unique closest and furthest passages, dependent on the number of roots these expressions have. This was seen explicitly in the two-body example given in Chapter 3 where there was a resultant quadratic equation related to this constraint.

Part III
Applications to Asteroids, Comets and Planetary Satellites

Applied Analysis and Case Studies

We now analyze a number of specific problems that involve motion about small bodies, applying the models and techniques from the previous chapters to these strongly perturbed systems. Our approach is to consider some specific bodies and describe the application of our previously stated analysis techniques to them. Accordingly, this chapter represents a series of case studies for specific bodies or specific classes of bodies. However, the descriptions are made as general as possible in order to ensure that these analyses can be readily transported to other systems.

In Chapters 7 to 10 we focus on orbital dynamics about asteroids, only accounting for their mass distribution and morphology. In Chapter 11 we consider controlled motion in the vicinity of an asteroid. Chapters 12 to 14 focus on the case when solar radiation pressure is the dominant perturbation. Chapter 15 discusses the interaction of a spacecraft with radial cometary outgassing fields. Chapters 16 to 18 discuss the orbital dynamics about a planetary satellite.

7. Uniformly Rotating Bodies: Asteroid 433 Eros

The fundamental problem of orbital dynamics about small bodies is typified by the interaction of the spacecraft with a non-spherical body that is rotating uniformly in space. This was the problem presented to NASA's Near Earth Asteroid Rendezvous (NEAR) mission when it had its rendezvous with asteroid 433 Eros. Eros has a highly irregular shape (see Fig. 1.5), is uniformly rotating about its maximum moment of inertia, and is massive enough so that solar radiation pressure and solar gravity play a relatively minor role in perturbing a spacecraft trajectory. Thus the analysis of this problem can restrict itself solely to the interaction of a spacecraft and uniformly rotating mass distribution. This chapter is based on an analysis performed prior to rendezvous with the asteroid in 2000, based on a model from a flyby of the asteroid in late 1998 [175] and also recounts some pre-mission planning analyses [154]. Thus, the gravity field values, total mass, and asteroid shape are not as accurate as currently available models, but making such changes will not alter the results and conclusions significantly. Detailed descriptions of the actual mission dynamics and Eros shape model are found in [107, 83].

7.1 Model of 433 Eros

The necessary parameters and models for this analysis are the total mass of the asteroid, the rotation period of the asteroid, and the shape of the body (as we will apply a constant density approximation to find its gravitational field) (Table 7.1). The asteroid orbit about the Sun is not relevant, as we will be neglecting solar radiation pressure and the solar tide.

The shape model of Eros used in this analysis was obtained from the NEAR imaging science team [185] and has a 5-degree surface resolution. It has a total estimated volume of approximately 3,000 km^3 and a mean (volumetric) radius of 8.97 km. This shape was significantly refined following the NEAR spacecraft rendezvous with the asteroid [47]. The shape model has been transformed into a triangular plate model with 2,432 vertices and 4,860 plates, centered at its volumetric center and oriented along its principal axes of inertia (assuming constant density).

D.J. Scheeres, *Orbital Motion in Strongly Perturbed Environments*, Springer Praxis Books,
DOI 10.1007/978-3-642-03256-1_7, © Springer-Verlag Berlin Heidelberg 2012

Table 7.1 Basic Eros parameters used for this analysis

Parameter	Symbol	Value	Units
Gravitational Parameter	μ	5×10^{-4}	km^3/s^2
Rotation Period	T_r	5.27	hours
Rotation Rate	ϖ	3.3118×10^{-4}	radians/s
Oblateness Gravity Coefficient	$R_o^2 C_{20}$	-26.755	km^2
Ellipticity Gravity Coefficient	$R_o^2 C_{22}$	12.752	km^2

For gravitational calculations outside of the circumscribing sphere of the asteroid, a gravity field expanded up to degree and order 16 derived from the constant density shape is used [191], while when close to the body the polygonal gravitational field algorithm of Werner, described previously, is used. Analytical calculations are restricted to the low degree and order gravity field coefficients.

7.2 Equations of Motion

There are two forms of the equations of motion that are useful for this analysis, a Lagrangian dynamics form and a Lagrange Planetary Equations form. Both of these are used and briefly restated for the specific case of a uniformly rotating body with no external perturbations.

7.2.1 Lagrangian Dynamics

First consider the Lagrangian dynamics written in the frame rotating with the asteroid. In this coordinate system the central gravity field does not change with time, and since the rotational rate of the asteroid is constant, the equations of motion are time-invariant. The Lagrangian in the body-fixed frame is

$$L(\boldsymbol{r}, \dot{\boldsymbol{r}}) = \frac{1}{2}\dot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}} + \frac{1}{2}\varpi^2 \boldsymbol{r} \cdot \tilde{\hat{\boldsymbol{z}}} \cdot \tilde{\hat{\boldsymbol{z}}} \cdot \boldsymbol{r} - U(\boldsymbol{r}) \tag{7.1}$$

We assume that the principal moments of inertia have been aligned with the coordinate axes of the system, with the $\hat{\boldsymbol{z}}$-axis aligned with the maximum moment of inertia. Hence the rotational dynamics of the body consist of uniform rotation about this axis.

In the body-fixed reference frame the equations of motion are:

$$\ddot{x} - 2\varpi\dot{y} = \varpi^2 x + U_x \tag{7.2}$$

$$\ddot{y} + 2\varpi\dot{x} = \varpi^2 y + U_y \tag{7.3}$$

$$\ddot{z} = U_z \tag{7.4}$$

where ϖ is the rotation rate computed from the asteroid rotation period and U is the gravitational potential. Since these equations are time invariant they have a

Jacobi integral

$$J = \frac{1}{2}\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) - \frac{1}{2}\varpi^2\left(x^2 + y^2\right) - U(x, y, z) \tag{7.5}$$

7.2.2 Lagrange Planetary Equations

An alternate description of the orbit dynamics is found using the Lagrange Planetary Equations. For our computations it is more useful to use the canonical form of the Lagrange equations given in Chapter 4, revised to express the change in orbit Keplerian energy, E, total orbit angular momentum, H, and orbit angular momentum projected onto the z-axis, G, as a function of the gravitational perturbations acting on them.

$$\frac{dE}{dt} = \frac{\partial R}{\partial t} \tag{7.6}$$

$$\frac{dH}{dt} = \frac{\partial R}{\partial \omega} \tag{7.7}$$

$$\frac{dG}{dt} = \frac{\partial R}{\partial \Omega} \tag{7.8}$$

where:

$$E = -\frac{\mu}{2a} \tag{7.9}$$

$$H = \sqrt{\mu a(1 - e^2)} \tag{7.10}$$

$$G = H\cos i \tag{7.11}$$

$$R = U - \frac{\mu}{r} \tag{7.12}$$

and a, e, i, ω and Ω are the osculating orbital elements. Also note the orbit parameter $p = a(1 - e^2)$ which is used when convenient. Finally, denote the true anomaly by f and the mean anomaly at epoch term as σ.

For the additional equations describing the dynamics of the argument of periapsis, longitude of the ascending node, and mean epoch the classical form of the Lagrange Planetary Equations are used.

$$\dot{\omega} = \frac{\sqrt{1 - e^2}}{na^2 e}\frac{\partial R}{\partial e} - \frac{\cot i}{na^2\sqrt{1 - e^2}}\frac{\partial R}{\partial i} \tag{7.13}$$

$$\dot{\Omega} = \frac{\csc i}{na^2\sqrt{1 - e^2}}\frac{\partial R}{\partial i} \tag{7.14}$$

$$\dot{\sigma} = -\frac{(1 - e^2)}{na^2 e}\frac{\partial R}{\partial e} - \frac{2}{na}\frac{\partial R}{\partial a} \tag{7.15}$$

It is important to restate the Jacobi integral in terms of the variables considered here. In this system the integral takes on a particularly simple form when so evaluated:

$$J = E - \varpi G - R \tag{7.16}$$

7.3 Analytic Characterization

First the application of approximate, analytical analysis to characterize spacecraft orbital dynamics about a uniformly rotating body are described. The results from such an analysis are not as exact as numerical computations, but have general application as several key parameters can be left in functional form.

7.3.1 Perturbations from Asteroid Oblateness

The effect of the C_{20} gravity term, or the oblateness, on a spacecraft orbit has been studied extensively in connection with the Earth orbiter problem (see [19]). To accurately capture this effect analytically requires the inclusion of higher-order gravity terms, such as C_{30} and C_{40}. Recall the general forms for these potential expansions

$$R_{20} = \frac{\mu R_o^2 C_{20}}{r^3}\left(1 - \frac{3}{2}\cos^2\delta\right) \tag{7.17}$$

$$R_{30} = \frac{\mu R_o^3 C_{30}}{2r^4}\sin\delta\left(5\sin^2\delta - 3\right) \tag{7.18}$$

$$R_{40} = \frac{\mu R_o^4 C_{40}}{8r^5}\left[35\sin^4\delta - 40\sin^2\delta + 3\right] \tag{7.19}$$

where $\sin\delta = \sin i \sin(\omega + f)$. The averaging of these potential yield the secular potentials

$$\overline{R}_{20} = \frac{\mu R_o^2 C_{20}}{2a^3(1-e^2)^{3/2}}\left(\frac{3}{2}\sin^2 i - 1\right) \tag{7.20}$$

$$\overline{R}_{30} = \frac{3\mu R_o^3 C_{30}}{2a^4(1-e^2)^{5/2}}\sin i \sin\omega e\left(\frac{5}{4}\sin^2 i - 1\right) \tag{7.21}$$

$$\overline{R}_{40} = \frac{3\mu R_o^4 C_{40}}{8a^5(1-e^2)^{7/2}} \left[\left(\frac{35}{8}\sin^4 i - 5\sin^2 i + 1\right)\left(1 + \frac{3}{2}e^2\right) + \frac{5}{4}\sin^2 i\cos(2\omega)\left(3 - \frac{7}{2}\sin^2 i\right)e^2\right] \tag{7.22}$$

Substituting $\overline{R}_{20}$ into the equations for the secular rates in argument of periapsis, longitude of ascending node and mean anomaly yields

$$\frac{d\overline{\omega}}{dt} = \frac{3nC_{20}}{2p^2}\left(\frac{5}{2}\sin^2 i - 2\right) \tag{7.23}$$

$$\frac{d\overline{\Omega}}{dt} = \frac{3nC_{20}}{2p^2}\cos i \tag{7.24}$$

$$\overline{n} = n\left[1 + \frac{3C_{20}\sqrt{1-e^2}}{2p^2}\left(\frac{3}{2}\sin^2 i - 1\right)\right] \tag{7.25}$$

The significant item to note is that the secular rates of these angles can become very large when orbiting close to asteroids. The leading frequency of these variations is $A_{20} = 3nC_{20}/(2p^2)$ and for Eros has the value

$$A_{20} = -85.6 \frac{1}{\tilde{a}^{7/2}(1-e^2)^2} \text{ degrees/hour} \tag{7.26}$$

where $\tilde{a}$ is the spacecraft semi-major axis normalized by the mean asteroid radius (8.97 km for Eros) and e is the orbital eccentricity. The corresponding frequency for an Earth orbiter is 0.415 degrees/hour, over two orders of magnitude smaller. Thus, for operations within several mean radii of the asteroid the spacecraft orbit will be subject to large secular rates in its orbit plane orientation and in-plane orientation. The large rate clearly shows how orbital dynamics close to an asteroid are significantly more perturbed than planetary orbiters, and implies that assumptions made classically for planetary orbiters may not always apply.

Once an orbit is chosen that is "safe" from most destabilizing influences, such as a retrograde orbit, these large secular rates must still be accounted for and cannot, except in a few specific orbit geometries, be directly controlled. The impact and control of these effects were evaluated and incorporated into the NEAR orbital mission plan [154].

7.3.2 Analytical Stability Analysis of Polar Orbits

Polar orbits and the global coverage they provide were of particular interest to the NEAR mission and will be to any asteroid mapping experiment. In fact, the NEAR mission had an extended period in a polar orbit and there was interest in decreasing the orbit radius as much as possible to gain additional signal in the gravitational measurements. During the mission it was determined numerically that polar orbits had a lower limit on radius that, if violated, soon led to unstable (escaping) orbital motion.

It is, in fact, possible to better understand this lower limit using classical analytical and averaging techniques combined with more modern characterizations found for asteroid orbiters. To better understand the dynamics in this particular case, the combined effects of the C_{20}, C_{30}, and C_{40} gravity coefficients on a polar orbit must be considered.

Of particular interest to us is the evolution of the eccentricity, which follows the basic equation:

$$\dot{e} = -\frac{\sqrt{1-e^2}}{na^2e}\frac{\partial R}{\partial \omega} \tag{7.27}$$

$$e' = -\frac{r^2}{\mu a e}\frac{\partial R}{\partial \omega} \tag{7.28}$$

where e' denotes differentiation with respect to orbit true anomaly. Recall the familiar result from the standard averaging analysis of C_{20} that the argument of

periapsis has a secular component defined as:

$$\dot{w} = \frac{3nr_o^2 C_{20}}{2a^2(1-e^2)^2}\left(\frac{5}{2}\sin^2 i - 2\right) \tag{7.29}$$

meaning that the argument of periapsis will have a constant rate of change on average, and that the eccentricity will have no secular evolution due to this particular gravity term.

What is observed numerically for orbiters near Eros, however, is a large transient oscillation in eccentricity which causes the orbit periapsis to occasionally dip closer to or farther from the asteroid. As established previously in [175] (and discussed later in this chapter) it is the orbit periapsis and eccentricity that controls the coupling of the asteroid gravity field to the orbit dynamics, and that periapsis passages at a close enough distance can rapidly destabilize motion. Thus it is of interest to characterize analytically the expected variations in eccentricity due to the gravity field perturbations. This can be done by inspecting the transient variations in eccentricity due to C_{20} and the long-period variations in eccentricity due to C_{30} and C_{40} combined with the secular motion of the argument of periapsis. In the following the assumption that the orbit is near-circular is made, and thus that higher orders of eccentricity can be ignored.

The differential equation for eccentricity with respect to C_{20} can be written as:

$$\begin{aligned} e' = \frac{3}{2}\frac{C_{20}\sin^2 i}{a^2}\big[&\sin(2\omega)\left(\cos(3f)+\cos(f)\right) \\ &+\cos(2\omega)\left(\sin(3f)+\sin(f)\right)+\mathcal{O}(e)\big] \end{aligned} \tag{7.30}$$

which can be integrated with respect to true anomaly to yield:

$$\begin{aligned} \Delta e_{20} = \frac{3}{2}\frac{C_{20}\sin^2 i}{a^2} & \\ \Bigg[&\sin(2\omega)\left(\frac{1}{3}\sin(3f)+\sin(f)\right) \\ &+\cos(2\omega)\left(\frac{1}{3}(1-\cos(3f))+1-\cos(f)\right)+\mathcal{O}(e)\Bigg] \end{aligned} \tag{7.31}$$

This gives the amplitude of the short-period eccentricity fluctuation about its mean value, which is found by taking the average of the above equation over one orbit, or

$$\overline{\Delta e}_{20} = \frac{1}{2\pi}\int_0^{2\pi} \Delta e\, dM \tag{7.32}$$

$$= \frac{2r_o^2 C_{20}}{a^2}\sin^2 i\cos(2\omega)+\mathcal{O}(e) \tag{7.33}$$

Thus, the total fluctuation in eccentricity due to C_{20} is computed as $\overline{\Delta e}_{20} + \Delta e_{20}$, for which a reasonable bound can be found:

$$\overline{\Delta e}_{20} + \Delta e_{20} \leq \frac{10}{3} \frac{r_o^2 C_{20}}{a^2} \sin^2 i \tag{7.34}$$

This upper limit does not, however, agree with numerical integration of an orbit in the full gravitational field, indicating that higher-order terms must be included in the analysis.

To analyze the contribution of C_{30} to the variation of e consider the differential equation for this term

$$\dot{e}_{30} = -\frac{3}{2} \frac{n r_o^3 C_{30}}{a^3 (1-e^2)^2} \left(\frac{5}{4} \sin^2 i - 1 \right) \sin i \cos \omega \tag{7.35}$$

The presence of $\cos \omega$ in the equation is important as it will have a secular change due to C_{20}. Thus, we change the independent variable to ω to find:

$$\frac{de_{30}}{d\omega} = \dot{e}_{30} / \dot{\omega}_{20} \tag{7.36}$$

$$= -\frac{1}{2} \frac{r_o^3 C_{30}}{a r_o^2 C_{20}} \sin i \cos \omega \tag{7.37}$$

which can be immediately integrated to find:

$$\Delta e_{30} = -\frac{1}{2} \frac{r_o^3 C_{30}}{a r_o^2 C_{20}} \sin i \sin \omega \tag{7.38}$$

By combining an analysis of the effect of C_{20} and C_{30} on a near-circular, polar orbit the following relation for the maximum value of eccentricity over one circulation period of the argument of periapsis is found

$$e_{max} \sim \frac{10}{3} \frac{r_o^2 |C_{20}|}{a^2} + \frac{r_o^3 |C_{30}|}{a r_o^2 |C_{20}|} \tag{7.39}$$

This relation predicts that the maximum excursion in eccentricity is 0.125 for a 35-km orbit, 0.1 for a 40-km orbit, and 0.07 for a 50-km orbit. The equation is plotted in Fig. 7.1 for the values of Eros. The period of the oscillation is computed from secular C_{20} theory for the circulation in argument of periapsis:

$$T_\omega = \frac{8\pi}{3} \frac{a^{3.5}}{\sqrt{\mu} r_o^2 |C_{20}|} \tag{7.40}$$

and is equal to 41 days in a 35-km orbit, 62 days in a 40-km orbit, and 142 days in a 50-km orbit. This long-period oscillation in eccentricity does not seem to affect the long-term stability of the orbit until the semi-major axis drops below 35 km. Agreement between these approximate limits and numerical integrations for long-term stable orbits seem to agree well; see Fig. 7.2.

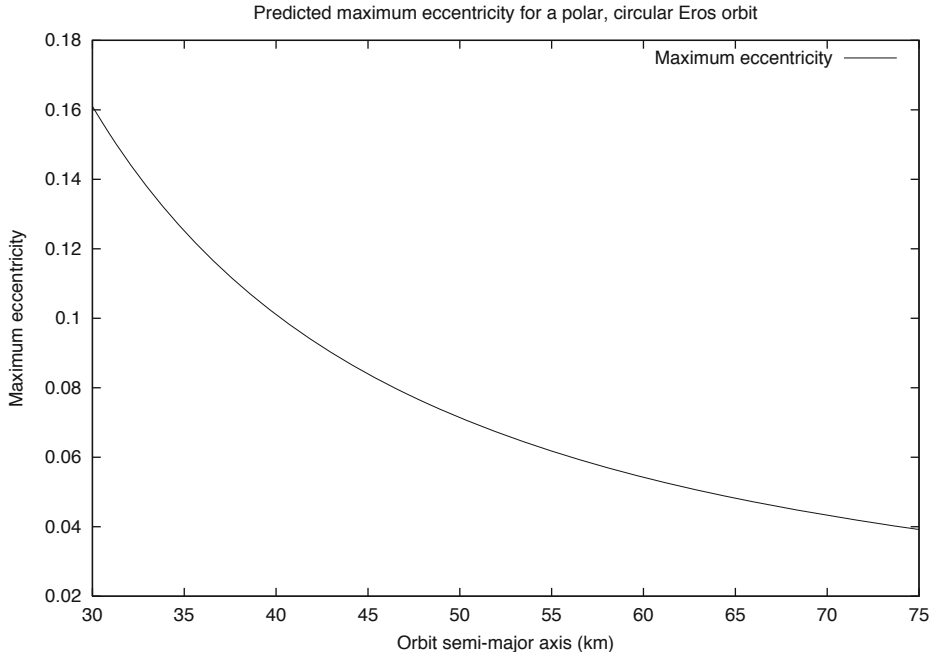


Fig. 7.1 Maximum eccentricity oscillation as a function of altitude for a polar orbiter.

When the polar orbit radius drops below 35 km, it enters a region of stronger interaction with the rotating gravity field. This also coincides with a 5:2 resonance at $a = 35$ km and a 3:1 resonance at $a = 33$ km between the orbit period and the rotating asteroid. The mechanism of the instability at this limit follows a pattern that involves the long-term eccentricity oscillation plus the (small) effect of the C_{40} gravity coefficient term.

As the strength of the Eros/orbit interactions increases with eccentricity, an initially circular orbit at a semi-major axis of 33 km will experience a mounting perturbation as the eccentricity increases (due to C_{20} and C_{30}). However, numerical integration of the spacecraft motion in an Eros gravity field taken to third degree and order only, starting at this semi-major axis value, will not experience a long-term instability. This is due to the maximum eccentricity amplitude not being large enough to bring the orbit into a region of strong interaction with the gravity field. If the effect of the C_{40} gravity term is modeled, however, it boosts the maximum eccentricity by a small amount, characterized as a function of argument of periapsis. To find this contribution consider the averaged equation for eccentricity due to this gravity term with respect to time and with respect to argument of periapsis:

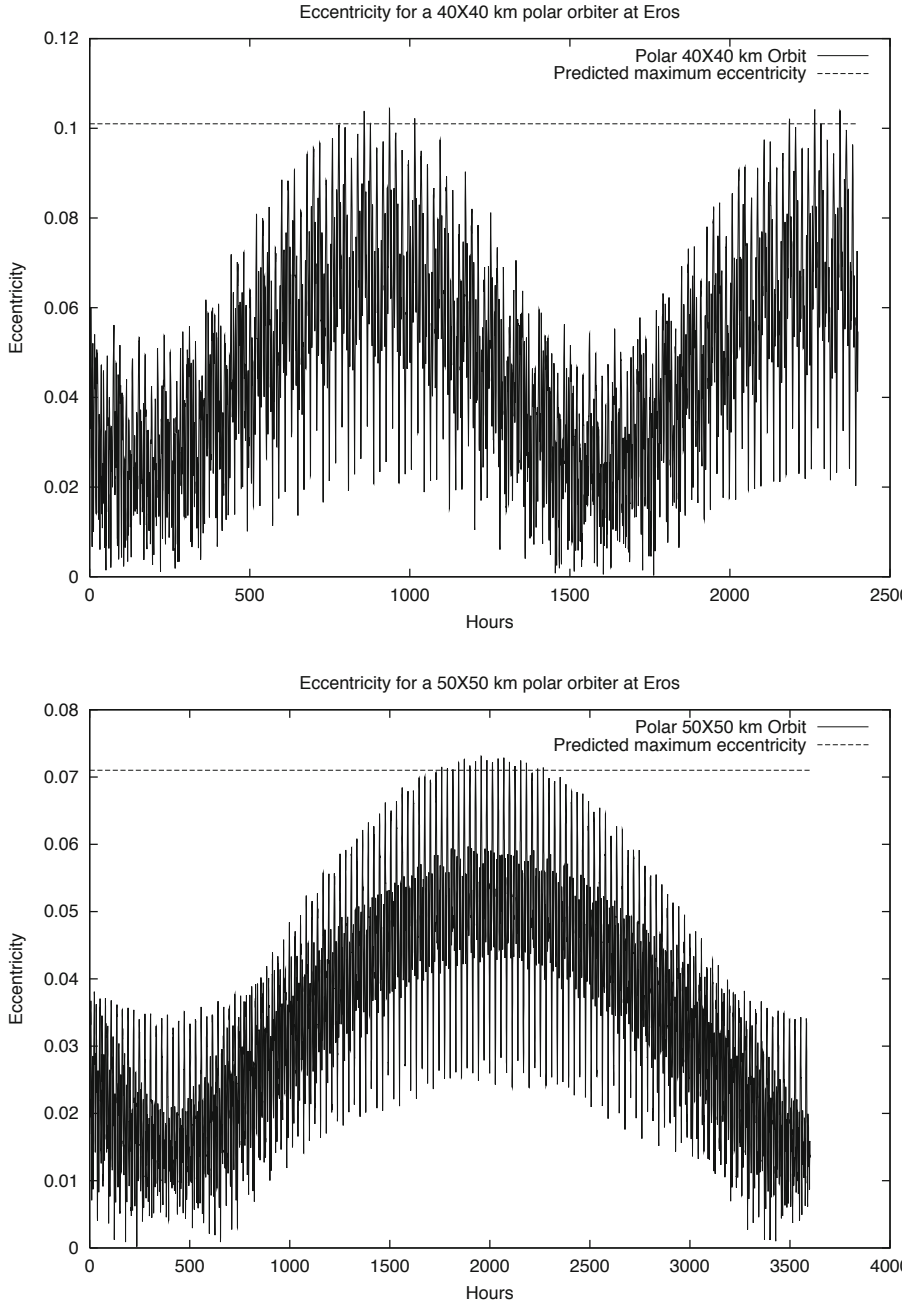


Fig. 7.2 Time integration of eccentricity for polar orbits at semi-major axes of 40 (top) and 50 (bottom) kilometers. The predicted constraints on short and long-period fluctuations are seen to agree well.

$$\dot{e}_{40} = \frac{15}{16}\frac{nr_o^4 C_{40}}{a^4(1-e^2)^3}\left(3 - \frac{7}{2}\sin^2 i\right)\sin^2 i \sin(2\omega)e \tag{7.41}$$

$$\frac{de_{40}}{d\omega} = -\frac{15}{16}\frac{r_o^4 C_{40}}{a^2(1-e^2)r_o^2 C_{20}}\frac{\left(\frac{7}{6}\sin^2 i - 1\right)}{\left(\frac{5}{4}\sin^2 i - 1\right)}\sin(2\omega)e \tag{7.42}$$

Assuming small e again, and allowing inclination to be equal to 90° gives us

$$e_{40} = e_o \exp\left[-\frac{5}{8}\frac{R_o^4 C_{40}}{a^2 R_o^2 C_{20}}\sin^2\omega\right] \tag{7.43}$$

At a semi-major axis of 35 km this perturbation can increase eccentricity by at most 4.5%. At $a = 33$ km the effect of this gravity term is sufficient to nudge the orbit into a region of unstable motion, characterized by large changes in orbit energy and angular momentum from orbit to orbit. Numerical integrations including fourth-degree and -order gravity verifies this result, as is evident from Fig. 7.3.

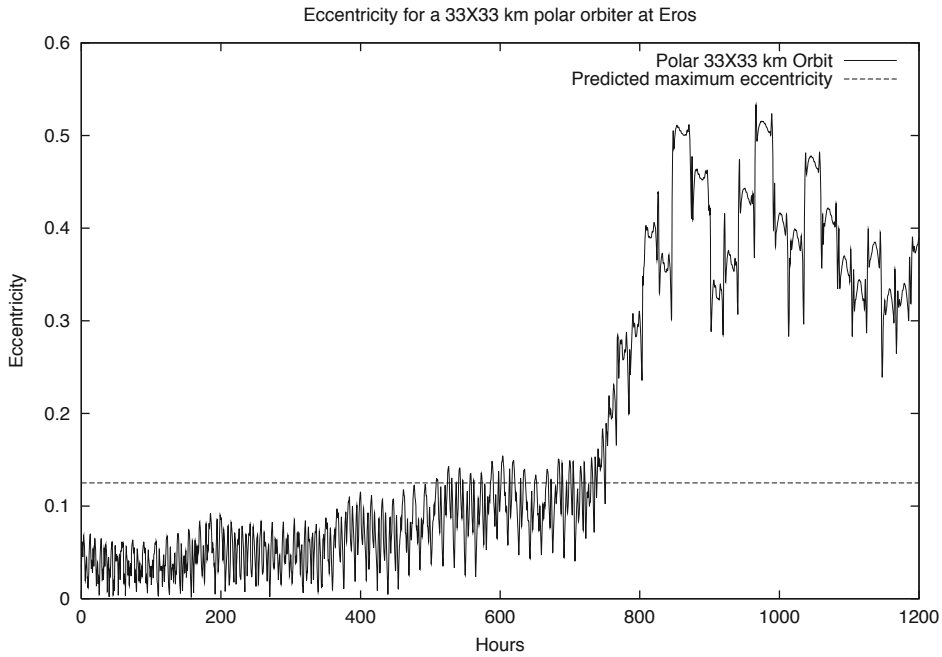


Fig. 7.3 Eccentricity of a numerically integrated polar orbit with an initial semi-major axis of 33 km, low enough to excite unstable motion.

For initial values of semi-major axis chosen lower than ∼33 km the long-term oscillation in eccentricity coupled with the increasing strength of the transient fluctuations in energy and angular momentum combine to make it infeasible to safely orbit the asteroid. It is interesting to note that this is approximately the same limit found for the linear stability of direct, near equatorial orbits and agrees with the results reported in [91].

7.3.3 Effect of Transient Perturbations

In the extremely perturbed environment close to an asteroid such as Eros the spacecraft can be subject to large changes in its orbital elements over a relatively short time period. Such situations have not arisen in classical astrodynamics, where perturbations are generally small and effects take many orbits or days to become significant. In the asteroid environment, however, large fluctuations in an orbit can be observed per orbit about the body, in many cases causing a chaotic evolution of an orbit leading to an impact or escape from the body. Even though these effects are large, it is possible to characterize them analytically and even use them in the design of close-proximity operations [3].

For analytical computations we use the canonical form of the Lagrange Planetary Equations stated earlier which express the change in orbit Keplerian energy, angular momentum, and angular momentum projected onto the z-axis as a function of the gravitational perturbations acting on them.

The strongest perturbation that the trajectory feels when in close proximity to the asteroid is mainly due to the second-degree and -order gravity field of the rotating body [175], which has the explicit form:

$$R_{20+22} = \frac{\mu R_o^2}{r^3}\left[C_{20}\left(1 - \frac{3}{2}\cos^2\delta\right) + 3C_{22}\cos^2\delta\,\cos 2\lambda\right] \tag{7.44}$$

Our analytical estimates only consider the contribution from these terms.

Although the orbital dynamics are best computed using numerical integrations, a class of estimates for the change in orbit energy and angular momentum can be derived which provide a great deal of insight into the effect of a close flyby on the resulting orbit [158]. Of particular interest is the change in orbit parameters as a spacecraft descends from a relatively high apoapsis to a periapsis close to the asteroid surface, and the change in orbit parameters as the spacecraft travels through a full orbit, from apoapsis to apoapsis. In the following the explicit formulae for the change in energy E, angular momentum H, and the projection of the angular momentum along the z-axis, G are stated. Changes in these parameters can be related to changes in the classical orbital elements (see [187, 129]). The following directly applies the theory outlined in Section 6.7.

Half-Orbit Perturbations

Over a transfer from apoapsis to periapsis variations in the orbital elements due to both the C_{20} and the C_{22} gravity coefficients occur. The projected angular momentum G is identically conserved under perturbation from C_{20}, and thus will not vary. The energy and total angular momentum, E and H will vary over a half orbit, however, yielding changes of:

$$\Delta H_{C_{20}} = 2\sqrt{\frac{\mu}{p^3}}C_{20}\cos 2\omega \sin^2 ie \tag{7.45}$$

$$\Delta E_{C_{20}} = \frac{\mu}{2p^3}C_{20}\left[-1 + 3\left(\cos^2 i + \sin^2 i\cos 2\omega\right)\right]\left(3 + e^2\right)e \tag{7.46}$$

Similarly, and leading to more detailed results, the change in orbit elements E, H, and G due to C_{22} over one-half an orbit can also be predicted approximately:

$$\begin{aligned}\Delta H_{C_{22}} = &-3E_{22}\sqrt{\frac{\mu}{p^3}} \\ &\left[\cos^4(i/2)\left\{\sin 2(\omega+\Omega)I_2^1 + \cos 2(\omega+\Omega)J_2^1\right\}\right. \\ &\left.+\sin^4(i/2)\left\{\sin 2(\omega-\Omega)I_{-2}^1 - \cos 2(\omega-\Omega)J_{-2}^1\right\}\right] \qquad (7.47)\\ \Delta G_{C_{22}} = &-3C_{22}\sqrt{\frac{\mu}{p^3}}\left[\frac{1}{2}\sin^2 i\left\{\sin 2\Omega I_0^1 + \cos 2\Omega J_0^1\right\}\right. \\ &+\cos^4(i/2)\left\{\sin 2(\omega+\Omega)I_2^1 + \cos 2(\omega+\Omega)J_2^1\right\} \\ &\left.-\sin^4(i/2)\left\{\sin 2(\omega-\Omega)I_{-2}^1 - \cos 2(\omega-\Omega)J_{-2}^1\right\}\right] \qquad (7.48)\end{aligned}$$

and the variation in E can be found from the Jacobi integral. The integrals I_m^n and J_m^n have the definitions:

$$I_m^n = 2\int_0^\pi (1+e\cos f)^n \cos(mf - 2\omega_E t)\, df \qquad (7.49)$$

$$J_m^n = 2\int_0^\pi (1+e\cos f)^n \sin(mf - 2\omega_E t)\, df \qquad (7.50)$$

These integrals cannot be expressed in closed form in general except for the particular cases:

$$I_0^{-2} = \frac{\sin\left(2\pi\sqrt{\omega_E^2 a^3/\mu}\right)}{\sqrt{\omega_E^2 p^3/\mu}} \qquad (7.51)$$

$$J_0^{-2} = \frac{\cos\left(2\pi\sqrt{\omega_E^2 a^3/\mu}\right)}{\sqrt{\omega_E^2 p^3/\mu}} \qquad (7.52)$$

The numerical quadrature of these integrals has been treated previously [158]. These integrals are intimately related to the Hansen Coefficients [159]. Figure 7.4 shows the values of these integrals over an interval of parameter values of interest. The total variation in these elements over an apoapsis to periapsis passage is then computed as:

$$\Delta H = \Delta H_{C_{20}} + \Delta H_{C_{22}} \qquad (7.53)$$

$$\Delta G = \Delta G_{C_{20}} + \Delta G_{C_{22}} \qquad (7.54)$$

$$\Delta E = \Delta E_{C_{20}} + \Delta E_{C_{22}} \qquad (7.55)$$

Using these results, it is possible to predict the change in orbital parameters between apoapsis of the transfer ellipse to periapsis of the ellipse. In general, the variation in energy and angular momentum can be quite large, implying that using Keplerian orbits to initiate numerical targeting routines may be an inefficient way in which to choose a target flyover condition. Rather, using the above relations in

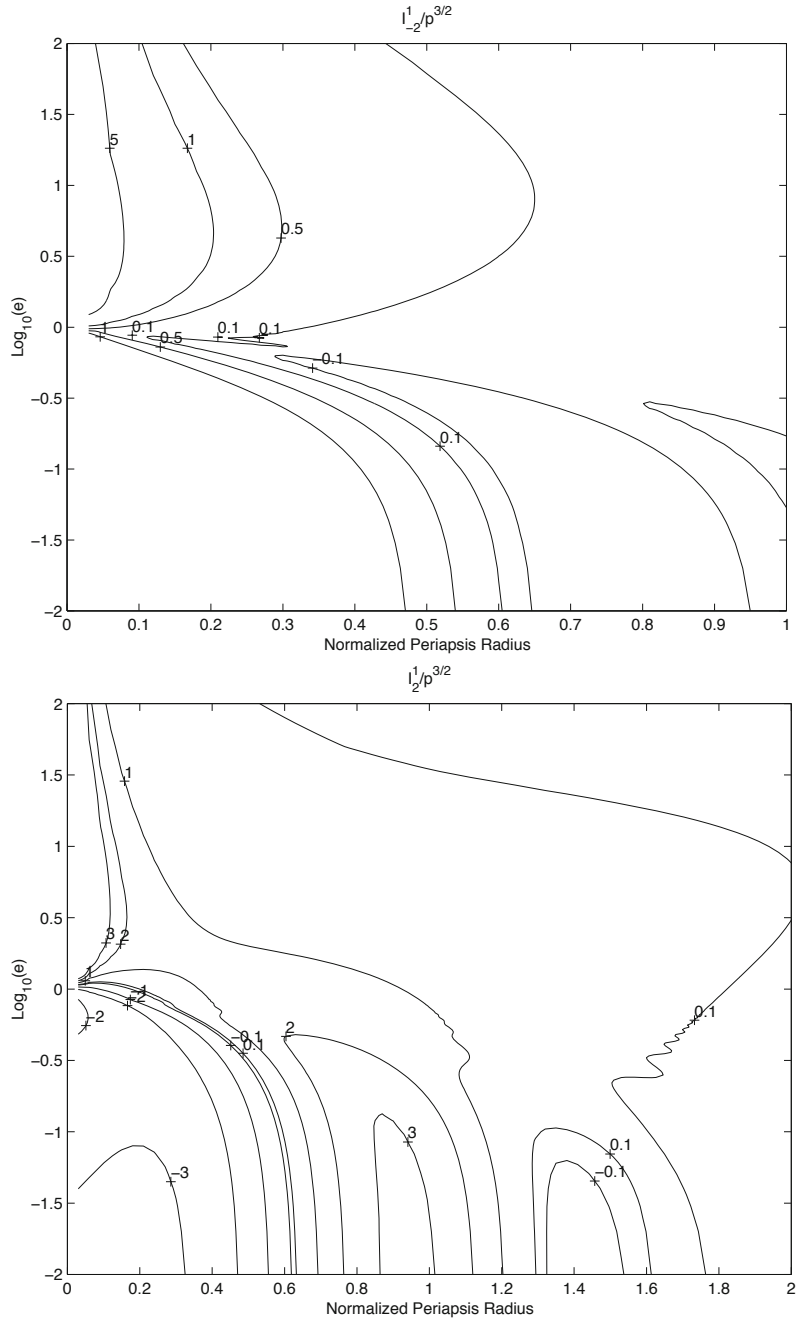


Fig. 7.4 Contours of the integrals that describe the transient effect of the C_{22} gravity coefficient on an orbiter.

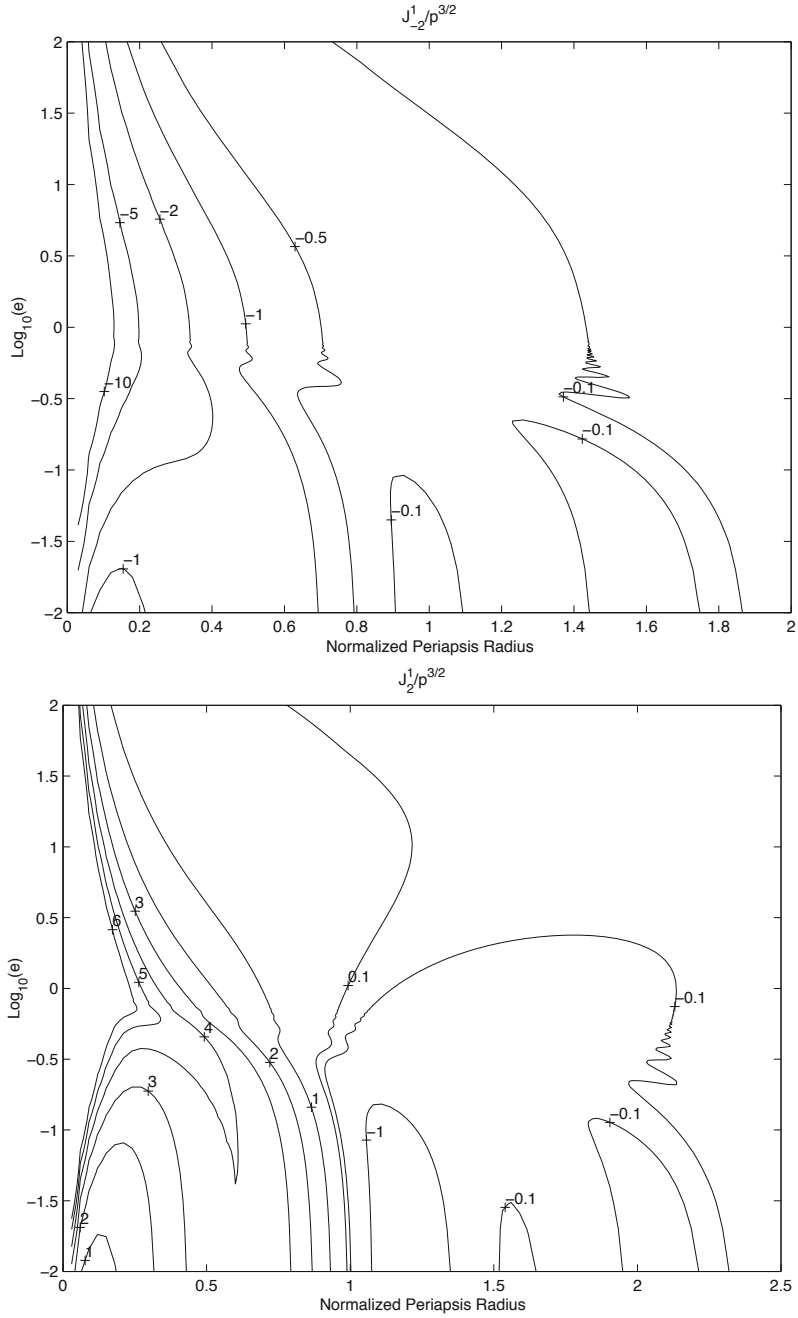


Fig. 7.4 Continued.

the initial design of a de-orbit maneuver can provide additional insight into the selection of target points for observation on the asteroid surface.

Full-Orbit Perturbations

For analysis of the low-altitude flyovers the orbit is continued through periapsis up to its next apoapsis passage, experiencing additional perturbations along the way. For the effect of the C_{20} gravity term these additional perturbations erase the fluctuations in angular momentum and energy. For the C_{22} effects, however, only a partial cancellation occurs, leaving a residual change in the orbit elements that is often rather large [158, 175]. The results are similar to the half-orbit results, except that the integrals J_m^n will cancel out over a full orbit pass, leaving only the terms containing the integrals I_m^n.

$$\Delta H = -6C_{22}\sqrt{\frac{\mu}{p^3}} \left[\cos^4(i/2)\sin 2(\omega+\Omega)I_2^1 + \sin^4(i/2)\sin 2(\omega-\Omega)I_{-2}^1\right] \tag{7.56}$$

$$\Delta G = -6C_{22}\sqrt{\frac{\mu}{p^3}}\left[\frac{1}{2}\sin^2 i \sin 2\Omega I_0^1 + \cos^4(i/2)\sin 2(\omega+\Omega)I_2^1 - \sin^4(i/2)\sin 2(\omega-\Omega)I_{-2}^1\right] \tag{7.57}$$

$$\Delta E = -6C_{22}\omega_E\sqrt{\frac{\mu}{p^3}}\left[\frac{1}{2}\sin^2 i \sin 2\Omega \left\{I_0^1 - (1-e)^3 I_0^{-2}\right\} + \cos^4(i/2)\sin 2(\omega+\Omega)\left\{I_2^1 - (1-e)^3 I_0^{-2}\right\} - \sin^4(i/2)\sin 2(\omega-\Omega)\left\{I_{-2}^1 - (1-e)^3 I_0^{-2}\right\}\right] \tag{7.58}$$

It can be seen that $I_2^1 \gg I_{-2}^1$ and $I_2^1 \gg I_0^1$ in the regions of interest (see Fig. 7.4). Thus, direct, low-inclination orbits will be subject to the terms I_2^1 while retrograde, near equatorial orbits will be primarily subject to the terms I_{-2}^1. Inspecting the contour plots, it is obvious that direct orbits will experience much larger changes in energy and angular momentum over each orbit, while the retrograde orbits will experience little, if any, change per orbit.

In Fig. 7.5 contour plots of the fractional change in orbit energy and angular momentum per orbit are presented, using the appropriate Eros constants. These contours are computed using only the I_2^1 integrals, per the above discussion. Thus, each of the contours will scale as $\cos^4(i/2)\sin[2(\omega+\Omega)]$, allowing the results to be generalized to a range of orbit inclinations and body-relative geometries (recall that the argument of periapsis and longitude of the ascending node define the location of periapsis in the body-fixed frame). Using this result it is possible to qualitatively predict the expected perturbation experienced by an orbit and use this to compute acceptable limits for proximity operations.

There are many applications of these formulae, including the setting of limits for the minimum periapsis of a spacecraft orbit and the design of very close passes

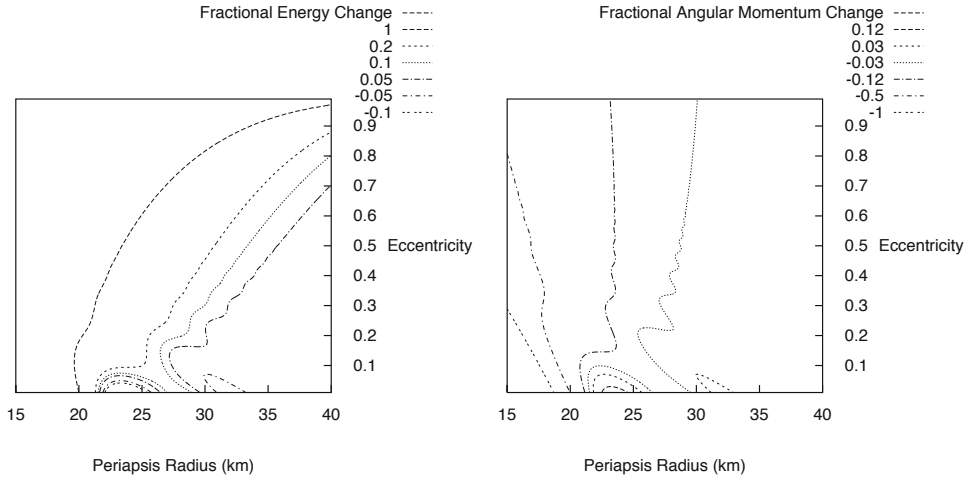


Fig. 7.5 Contour plot of fractional change in Keplerian energy (left) and angular momentum (right) per orbit, computed specifically for Eros. The contour values are scaled by $\cos^4(i/2)\sin(2(\omega+\Omega))$ to generate the predicted changes.

over the asteroid. Related formulae were used to design relatively safe close flybys over the Eros surface [3]. In that application the functions were used to predict which surface flybys would cause the spacecraft apoapsis and period to increase, ensuring ample time following a close flyby to transition the orbit into a safe orbit at a sufficient distance from the asteroid. Further application of this orbit-to-orbit mapping approach can be found in [187, 129, 143].

7.4 Phase Space Characterization

Now consider the explicit characterizations that can be made concerning the phase space about a uniformly rotating asteroid. These characterizations consist of applications of the Jacobi integral, computation of equilibrium points, and the explicit computation of families of periodic orbits and their stability about the asteroid.

7.4.1 Stability against Impact

Using the Jacobi integral it is possible to determine the conditions under which impact of a spacecraft with the asteroid surface is possible. The zero-velocity curves for the Eros system are found by computing the contours of the gravity plus centripetal potential in the body-fixed position space x, y, and z. These contour lines then define the limits of physical motion that a spacecraft can have in the body-fixed space, given that value of the Jacobi constant. In general the spacecraft dynamics

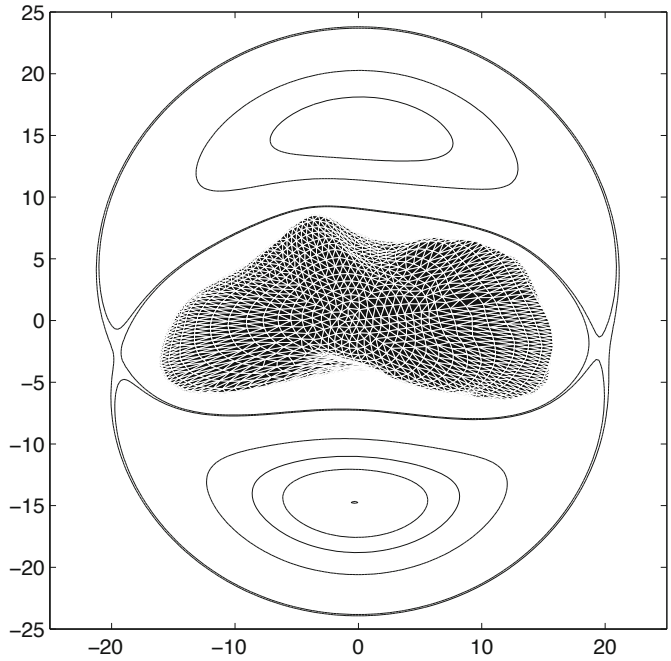


Fig. 7.6 Zero-velocity curves about Eros projected into the x-y-plane.

must satisfy the inequality constraint:

$$U + \frac{1}{2}\varpi^2\left(x^2 + y^2\right) + J_o \geq 0 \tag{7.59}$$

which constrains the spacecraft to lie on one side of the zero-velocity curves. Figure 7.6 shows the zero-velocity curves corresponding to the Eros shape model.

Of primary interest is finding the value of the Jacobi constant such that, for all values of J_o less than this, the zero-velocity curves are guaranteed to separate the trajectory space containing the asteroid and the space not containing the asteroid. This explicitly ensures that a spacecraft in orbit in the outer region with the appropriate Jacobi integral value can never, under gravitational dynamics alone, impact onto the asteroid surface. For the Eros shape model this value of J_o is found to be $J_o = -5.109 \times 10^{-5}$ km^2/s^2 and is associated with one of the equilibrium points relative to the body. To ensure stability against impact the initial spacecraft conditions must be chosen such that the spacecraft position resides in the outer portion of the zero-velocity curve and that the value of the Jacobi integral is less than or equal to J_o:

$$E - \varpi G - R \leq J_o \tag{7.60}$$

which provides a simple check in terms of osculating orbital elements for whether or not the spacecraft might impact with the surface at some point in the future. This

relation can be expressed in terms of initial osculating elements for an assumed direct, equatorial orbit specified by its periapsis radius, eccentricity and initial longitude λ in the body-fixed frame.

$$\frac{-\mu(1+e)}{2r_p} + \varpi\sqrt{\mu r_p(1+e)} + U(r=r_p,\lambda) + J_o = 0 \tag{7.61}$$

where the potential U is evaluated from the actual gravity field. Figure 7.7 shows a plot of the limiting stability-against-impact curve for the Eros system (including the full effect of the gravity field) in terms of initial periapsis radius and eccentricity for an equatorial orbit.

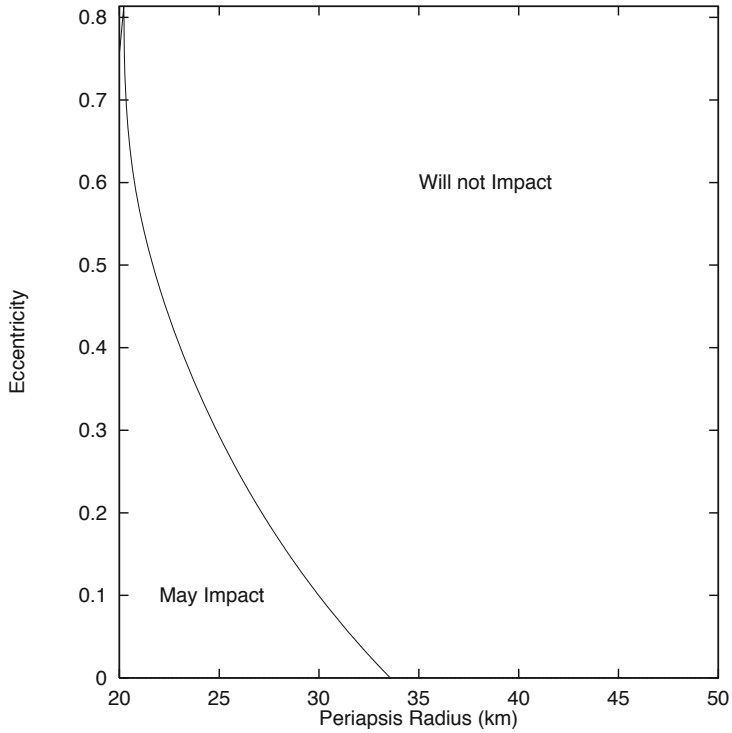


Fig. 7.7 Stability-against-impact curve for equatorial, direct orbits. Initial orbits to the left of this line may impact with Eros at some point in the future, orbits to the right of this line cannot impact with Eros unless some additional forces act on the satellites.

This analysis shows that direct, circular orbits must lie outside of 34 kilometers from Eros in order to not be able to impact on the surface. This does not mean that the spacecraft will not undergo large perturbations and changes in its orbit from the gravity field. The condition can also be computed for out-of-plane or retrograde orbits, but the sharpness of the zero-velocity surface analysis loses its power as inclination is increased, which is a known delinquency of this method.

7.4.2 Orbit Stability Characterizations

Another approach to the stability characterization of orbits about an asteroid is to study the stability of equilibrium points and periodic orbit families about the body. If stable families of periodic orbits can be found which lie near orbits of interest then these can be candidates for a mission orbit. Some basic families of periodic orbits and their stability are now discussed. In this context, the existence of a stable orbit will imply that a spacecraft placed near such an orbit will not experience large changes in its orbital elements over timespans of interest, usually days to weeks. An unstable orbit implies that, depending on the characteristic time of the instability, the spacecraft may deviate markedly from the original trajectory in a fairly short time. The characteristic instability times of orbits about asteroids can be very short, on the order of hours, and may interfere with orbital operations. This is especially true when one considers the navigation of spacecraft, as flying in an unstable orbit environment implies that the orbit uncertainty becomes stretched at a hyperbolic rate, leading to greater uncertainties about where the spacecraft is located [168].

Equilibrium Points

A uniformly rotating asteroid with an approximately ellipsoidal shape will in general have 4 synchronous orbits (equilibrium points) about it. Exceptions to this rule exist, with the asteroid Betulia being an interesting case in point with 6 equilibrium points [96]. These points will lie, approximately, along the long and short axes of the asteroid, close to the ideal synchronous orbit radius $r_s = (\mu/\varpi^2)^{1/3}$, which equals 16.581 km for Eros. Generally, the equilibrium points along the long-axis of the body will lie outside of r_s and will always be saddle points, having a hyperbolic stable and unstable manifold and two oscillatory modes (in and out of plane).

The equilibrium points along the short equatorial axis will lie inside of r_s and may either be completely oscillatory or have complex roots, leading to hyperbolic stable and unstable spirals. For many asteroid shapes considered thus far these equilibrium points are unstable. Exceptions will occur if the body has a small shape ellipticity in the equator, rotates slowly, or has a large density.

For the Eros shape model in question the four synchronous orbits were computed (all unstable) and their coordinates as well as their characteristic times and Jacobi values presented in Table 7.2. The smallest value of J for the equilibrium points, -5.109×10^{-5}, serves as the constant for the stability against impact criterion developed earlier.

Table 7.2 Positions, characteristic instability times and Jacobi constant values for the Eros synchronous orbits.

x (km)	y (km)	z (km)	τ (hours)	J (km/sec)^2
19.546	0.364	−0.839	0.69	-5.103×10^{-5}
−19.600	−0.158	−1.170	0.70	-5.109×10^{-5}
0.106	15.284	0.234	1.59	-4.279×10^{-5}
−0.110	−15.281	0.162	1.64	-4.279×10^{-5}

The osculating elements of such an equilibrium point are all constant (evident given that the period of motion is effectively zero) except for the longitude of the ascending node which has a secular rate equal to the rotation rate of the asteroid.

Resonances in Periodic Orbits

Other than equilibrium points, all other periodic orbits in the Eros-fixed frame have a non-zero period of motion and have a non-constant trajectory relative to the body. A distinction is made here between periodic orbits in the body-fixed frame and in the inertial frame. In general, a periodic orbit in the body-fixed frame will not be a periodic orbit in the inertial frame. However, at special values of the body-fixed orbit period the orbit will correspond to a periodic orbit in the inertial frame. This situation often corresponds to the intersection of two or more body-fixed periodic orbit families and usually corresponds to the onset of instability for one of these families. Thus, it is of interest for us to understand the conditions for these resonances to occur in our particular case.

Consider an orbit which repeats itself in the body-fixed space every time period T_P. Considering this same orbit, now specified in terms of osculating orbit elements, it is clear that the orbit elements (a, e, i) must repeat with period T_P, and the orbit elements of ω and M_o must shift by some value $2\pi m; m = 0, \pm 1, \pm 2, \cdots$ over a time T_P. The longitude of the ascending node, Ω, need not repeat and instead will shift by a multiple of 2π plus an added angle, all equaling $\theta = \varpi T_P$, where ϖ is the rotation rate of the asteroid. Thus, when this angle θ is commensurate with 2π the body-fixed periodic orbit is also an inertial periodic orbit. Intersections of these near-circular orbit families with eccentric orbit families can be found in the vicinity of these resonances, which result in intervals of unstable periodic orbit families.

Let us consider the two special cases of direct and retrograde periodic orbits. For each of these cases the body-fixed period is approximately:

$$\frac{1}{T_P} = \frac{1}{T_r} \pm \frac{1}{T_o} \tag{7.62}$$

$$T_o = \frac{2\pi r^{3/2}}{\sqrt{\mu}} \left[1 + \frac{3C_{20}}{4r^2}\right] \tag{7.63}$$

where the + sign is for retrograde orbits and the − sign is for direct orbits, T_r is the rotation period of the asteroid, T_o is the inertial orbital period of the spacecraft (computed with C_{20} corrections and assuming a circular orbit). The direct orbit period has a singularity when the inertial orbit period approaches the body rotation period, corresponding to the equilibrium points and the surrounding phase space. In Fig. 7.8 the ratio of the direct body-fixed orbit period over the Eros rotation period is shown, assuming the above formula. There are commensurabilities of 3:2 at a 35-km radius, 2:1 at a 27.5-km radius, and higher commensurabilities as the singularity is approached. In Fig. 7.10 the ratio of the retrograde body-fixed orbit period over the Eros rotation period is shown, again assuming the above formula. The only significant commensurability here is a 1:2 ratio at a radius of approximately 18 km, and the retrograde orbit becomes unstable in the vicinity of

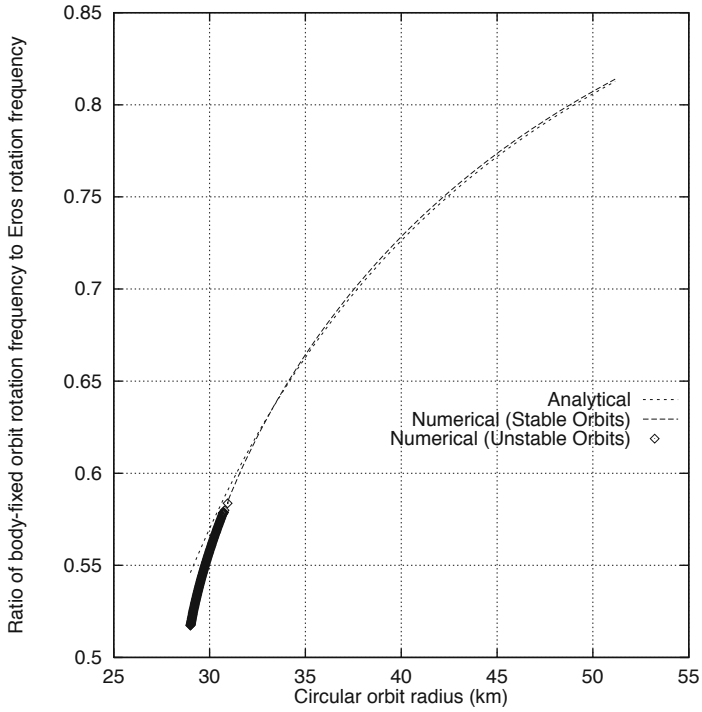


Fig. 7.8 Ratio of the periodic orbit frequency over the Eros rotation frequency for the direct family of orbits. Plotted is the predicted ratio (found using oblate planet theory) and the numerically computed ratio for the direct family of periodic orbits split into stable and unstable portions of the family.

this ratio. The direct orbit instabilities, however, are not as clearly linked to the commensurabilities, which may be due to the large perturbations in this regime. A detailed study of this phenomenon applied to periodic orbits about a general asteroid model is found in [74].

Direct Periodic Orbits

The family of direct, equatorial, near-circular, body-fixed periodic orbits were computed about this model of Eros. Figure 7.8 shows the family as a curve between mean orbit-radius and orbit-period frequency divided by Eros rotation frequency. Far from the body these orbits are stable, as expected; however, as the family comes closer to the body they become unstable – for the circular branch of the family this occurs at a radius of approximately 30.8 km and a Jacobi integral value of -4.938×10^{-5} $(\mathrm{km/s})^2$. Before this stability bifurcation occurs, the family itself splits into a circular branch and two elliptic branches – the elliptic branches lose their stability at a slightly higher value of the Jacobi constant.

The proper interpretation of this result is that the minimum stable, direct orbit radius about Eros will be ∼31 km, but this assumes some fairly special initial

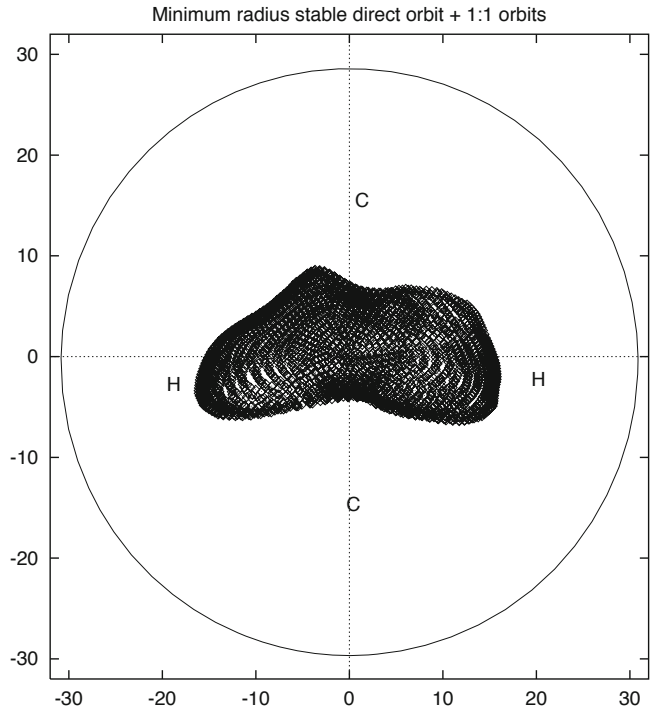


Fig. 7.9 Minimum radius stable, direct circular orbit about Eros and locations of the unstable 1:1 synchronous orbits projected into the x–y plane. Also shown is a silhouette of Eros.

conditions. It is possible, at higher radius values, to find non-periodic orbits that are characterized by having finite Lyapunov characteristic exponents (indicating that they are unstable). So, this limit should be considered to be an absolute minimum for stable, direct, equatorial orbits about Eros. Note that this minimum orbit radius violates the stability against impact limit implying that, if sufficiently perturbed from its stable orbit, it is possible for a spacecraft on that trajectory to impact on the asteroid surface.

It is significant to note from Fig. 7.8 that the onset of instability is not associated with any particular resonance, leading us to conclude that it is associated with the increasing strength of the gravity perturbations as the direct orbit draws closer to the rotating body and its C_{22} gravity term. Observing the predicted fractional change in orbit energy and angular momentum in the vicinity of where the periodic orbits first become unstable in Fig. 7.5 the fractional change in each of these quantities is on the order of 3–5% – leading to the conclusion that perturbations of this magnitude may destabilize the orbit dynamics. Shown in Fig. 7.9 is the minimum radius, direct, stable circular orbit about Eros and the 1:1 synchronous orbits, all projected into the x–y plane.

Retrograde Periodic Orbits

Periodic orbits are also computed which lie in the equatorial plane but travel in the opposite sense of the asteroid rotation. Figure 7.10 presents the orbit radius – frequency ratio curve for this orbit family. Generally speaking, such retrograde orbits are stable at almost all radii, even close to the asteroid surface. The reason for this stability is easy to understand: when orbiting in the opposite sense of the asteroid rotation the destabilizing effect of the C_{22} gravity term is diminished, as its effect becomes "averaged out" in space due to the large relative angular motion between the spacecraft and asteroid. In fact, as the spacecraft orbit becomes close to the asteroid, its rotation rate increases and hence the effective variations of the asteroid's gravity field diminish. This is clearly evident when numerically integrated retrograde orbits are compared with predictions from modeling the ellipsoid as an oblate body [153]. For retrograde orbits this comparison holds up quite well, indicating that these orbits are stable. This fact was used in the nominal design of the Eros orbital operations [154] and allowed the NEAR-Shoemaker spacecraft to come close to the Eros surface.

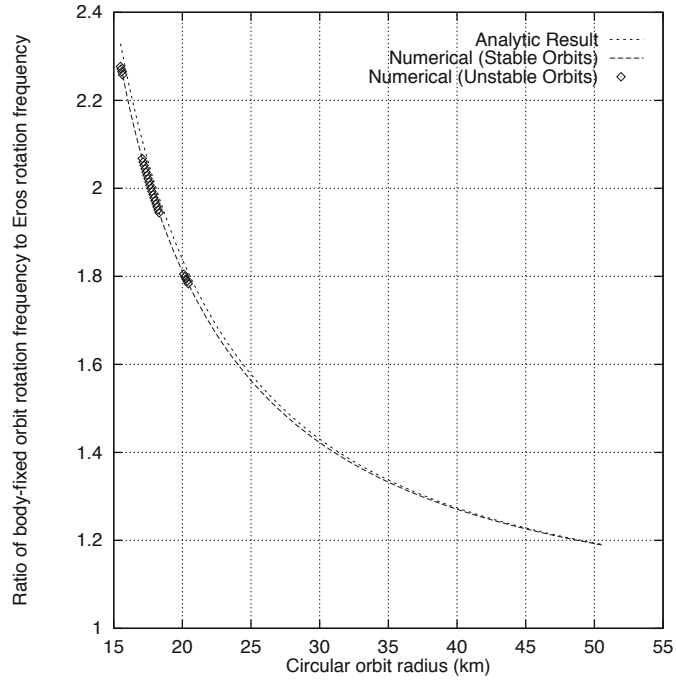


Fig. 7.10 Ratio of the periodic orbit frequency over the Eros rotation frequency for the retrograde family of orbits. Plotted is the predicted ratio (found using oblate planet theory) and the numerically computed ratio for the retrograde family of periodic orbits split into stable and unstable portions of the family.

There are some notable exceptions to this stability, however. These occur when resonances develop between the asteroid rotation rate and the secular motion of the spacecraft orbits' longitude of ascending node (as discussed above). For the case of Eros these resonances occur in the vicinity of 18 km, where a 1:2 commensurability exists between the periodic orbit period and the asteroid rotation rate. There is also an isolated instability region around 20 km, corresponding to the intersection of the retrograde orbits with an out-of-plane family of orbits with twice the period. Shown in Fig. 7.11 is the minimum radius, retrograde stable circular orbit about Eros, projected into the x–y plane, before the first instability interval is encountered.

Based on these instability limits, a reasonable constraint to place on retrograde, equatorial orbits is that they lie outside of 21 km from the center of Eros. Again, the danger of flying in or near a region of unstable orbits is due both to the possibility that the orbit may diverge from its nominal path and that the orbit uncertainty will increase significantly.

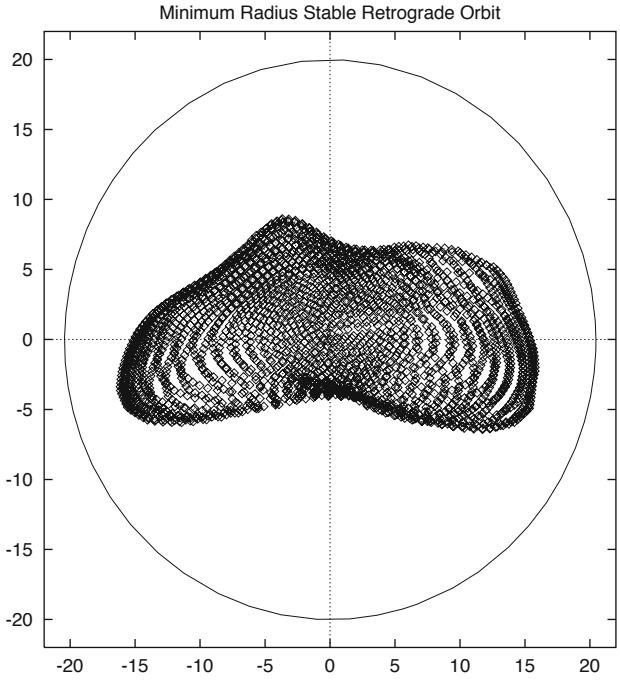


Fig. 7.11 Minimum radius stable, retrograde circular orbit about Eros projected into the x–y plane before the first instability interval is encountered. Also shown is a silhouette of Eros.

7.4.3 Out-of-Plane Orbits

The analysis thus far has focused on periodic orbits in or near to the asteroid equatorial plane. When significantly out-of-plane orbits are considered, several difficulties arise. For the periapsis and node arguments to have their required resonance, there are more stringent constraints placed on the orbit period so that, physically, the spacecraft is in the equatorial plane after one or more rotations of the central body. This restriction means that out-of-plane periodic orbits do not exist at all radii, as the equatorial families do, and thus are less useful in analyzing the stability of general motions about the asteroid. When such out-of-plane periodic orbits exist, they then correspond roughly to mean motion resonances with the rotation period of the asteroid. In [91] a detailed exploration of these periodic orbit families is made in order to characterize the stability limitations in that regime.

In order to gain additional insight into the stability of near-polar orbits results from the stability limit of direct periodic orbits (above) are combined with the analytical results found earlier for the change in energy and angular momentum. Consider, for definiteness, the stability of a polar orbit ($i = 90°$), leading to a scaling factor for the contour plots of $\cos^4(\pi/4) = 0.25$. Thus, the contour plot must be scaled by one-quarter. Due to differences in the gravitational and orbit interaction for a polar orbit (the interactions are more varied as there are more possible encounter geometries when in the nominal orbit) a reasonable criterion for stability appears to occur when the changes in orbit energy and angular momentum are on the order of 1%. This corresponds to the 5% contour lines on Fig. 7.5, and gives an indication of reasonable stability limits for polar orbits that are in agreement with the numerically determined limit of 33 km in a polar orbit.

The actual dynamics of polar (or of any) orbits about such a body are quite interesting. Due to the quasi-random movement of the argument of periapsis and longitude of the ascending node in the body-fixed space, the changes in energy and angular momentum will either add or subtract to the orbit, introducing the possibility of an orbit having random walks in terms of orbit elements. Figure 7.12 shows three numerically integrated polar orbits, started with an initial periapsis radius of 30 km and apoapsis radii of 45, 75 and 100 km respectively. Each is integrated for 600 hours (25 days), and their trajectory in the periapsis radius–eccentricity space is plotted. This plot shows how individual trajectories can either be confined to regions of orbit element space or diffuse into escaping or impacting trajectories. While there is significant correlation between Fig. 7.12 and Fig. 7.5, the controlling dynamical issues are also evidently much deeper.

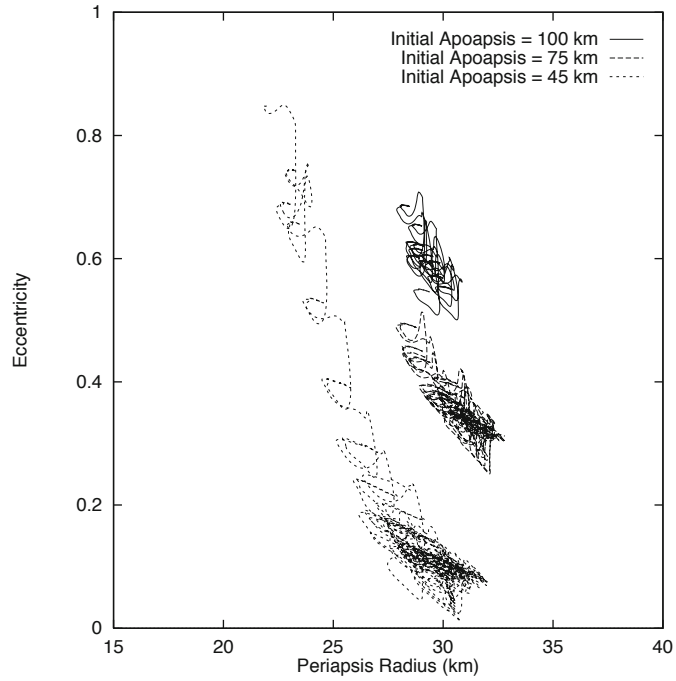


Fig. 7.12 Periapsis radius – eccentricity plots of three polar orbits over 600 hours (25 days). Each orbit starts with its periapsis radius at 30 km and its apoapsis radius at either 45, 75 or 100 km. Note the diffusion of the trajectories over time.

8. Complex Rotators: Asteroid 4179 Toutatis

In this chapter the assumption of uniform rotation is removed. Many asteroids are non-uniform rotators, potentially due to several different causes. These include spin-down due to the YORP effect, which can make an asteroid more susceptible to perturbing torques, and tumbling due to a close planetary flyby, an asteroid fission event, or an impact. Comets are also believed to often be in such a spin state due to outgassing jets on the surface that are activated around every perihelion passage. Strong evidence exists for several comets being in non-uniform rotation states including Halley and Encke [149, 12]. Once in a tumbling mode, we assume that the body will approximately follow the torque-free rigid-body dynamics described earlier. While this is not exactly true, due to solar radiation pressure torques, gravitational perturbations, and internal dissipation effects, over timespans of interest for the motion of a particle or spacecraft in the asteroid frame this is a good assumption.

In this section orbital motion close to and about asteroid 4179 Toutatis is analyzed. This is an asteroid that has been imaged to extremely high resolution using range-Doppler radar [76]. The asteroid is in a tumbling long-axis mode with body-fixed angular velocity period of 5.42 days and a period associated with its long-axis precession of 7.54 days. We again do not consider the effects of solar radiation pressure or the solar tide in order to keep our initial focus just on the gravitational problem. This mode of rotation completely changes the types of calculations and investigations one can perform for the body, and thus serves as an excellent candidate for our next case study. The results from this section are reduced from the detailed analysis of this asteroid presented in [173].

8.1 Model of 4179 Toutatis

The necessary parameters and models for this analysis are the assumed density of the asteroid (2.5 g/cm^3), the rotation state of the asteroid, and the shape of the body (as we will apply a constant density approximation to find its gravitational field) (Table 8.1). The asteroid orbit about the Sun is again not relevant as the

D.J. Scheeres, *Orbital Motion in Strongly Perturbed Environments*, Springer Praxis Books,
DOI 10.1007/978-3-642-03256-1_8, © Springer-Verlag Berlin Heidelberg 2012

solar radiation pressure and the solar tide are neglected. The density used is a pure assumption, as there are no direct measurements of the mass of Toutatis.

Table 8.1 Toutatis model parameters.

Parameter	Symbol	Value	Units
Gravitational Parameter	μ	1.279×10^{-6}	km^3/s^2
Effective Rotation Rate	ω_l	1.8548×10^{-5}	radians/s
Effective Rotation Period	$2\pi/\omega_l$	3.92	days
Dynamic Inertia	I_D/I_z	0.51983	–
Pole Declination	δ	$39.514 \rightarrow 40.450$	degrees
Right Ascension			
Period	$2\pi/\dot{\alpha}$	176.4	hours
Rate	$\dot{\alpha}$	9.894×10^{-6}	rad/s
Minimum Moment of Inertia	I_x/I_z	0.31335	–
Intermediate Moment of Inertia	I_y/I_z	0.94471	–
Oblateness Gravity Coefficient	$R_o^2 C_{20}$	0.77768	km^2
Ellipticity Gravity Coefficient	$R_o^2 C_{22}$	-0.01634	km^2

The shape model of Toutatis used in this analysis was obtained from range-Doppler imaging of the asteroid and has a surface resolution of about 84 meters and is described with a polygonal model with 1,600 vertices. It has a total estimated volume of approximately 7.67 km^3 and a mean (volumetric) radius of 1.223 km. A higher-resolution model with surface features at a precision of meters is also available [77]. For gravitational calculations outside of the circumscribing sphere of the asteroid we used a gravity field expanded up to degree and order 20, while when close to the body the polygonal gravitational field algorithm of Werner, described previously, is used. Analytical calculations are restricted to the second degree and order gravity field coefficients. Finally, the oblateness parameter has been defined relative to the body-fixed x-axis, the minimum moment of inertia, due to the rotation state of the body. As Toutatis is an a LAM rotational mode, the body makes complete rotations about its minimum moment of inertia and the dynamics become simpler to model with this change in convention. This is the reason for the positive C_{20} coefficient and the correspondingly small C_{22} coefficient, as Toutatis is nearly prolate about that axis. If we make the approximation that Toutatis is exactly prolate about this symmetry axis, then its pole declination is constant at $\sim 40^\circ$ and its precession period equals the right ascension period of 176.4 hours (7.35 days).

8.2 Equations of Motion

The equations of motion are stated in a Lagrangian framework in the body-fixed frame. The Lagrangian is again

$$L(\boldsymbol{r}, \dot{\boldsymbol{r}}, t) = \frac{1}{2}\dot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}} + \frac{1}{2}\boldsymbol{r} \cdot \tilde{\boldsymbol{\varpi}} \cdot \tilde{\boldsymbol{\varpi}} \cdot \boldsymbol{r} - U(\boldsymbol{r}) \tag{8.1}$$

however, by choosing a body-fixed frame the gravitational field is still time-invariant while now the angular velocity becomes time periodic, as it follows the torque-free rigid-body rotations (see Chapter 3). Application of Lagrange's equation leads to the second-order vector differential equation

$$\ddot{\boldsymbol{r}} + \dot{\boldsymbol{\varpi}} \times \boldsymbol{r} + 2\boldsymbol{\varpi} \times \dot{\boldsymbol{r}} + \boldsymbol{\varpi} \times \boldsymbol{\varpi} \times \boldsymbol{r} = U_{\boldsymbol{r}} \tag{8.2}$$

If switched to an inertial frame it is interesting to note that the time periodicity would be lost in general as tumbling motion is not time-periodic in inertial space. There is no longer a Jacobi integral for this problem.

A perturbation formulation using the Lagrange Planetary Equations is used with the perturbation potential specified as $R = U - \frac{\mu}{r}$. Due to the shape of Toutatis, nearly prolate with a clear pear-shaped deviation along its minimum moment of inertia, an appropriate statement of the potential can be found using the second-degree and -order gravity field perturbations.

$$R_2(\mathbf{r}) = -\frac{\mu}{2r^3}C_{20}\left[1 - 3(\hat{\mathbf{r}} \cdot \hat{\mathbf{p}})^2\right] + \frac{3\mu}{r^3}C_{22}\left[(\hat{\mathbf{r}} \cdot \hat{\mathbf{s}})^2 - (\hat{\mathbf{r}} \cdot \hat{\mathbf{q}})^2\right] \tag{8.3}$$

where we assume that the unit vectors $\hat{\mathbf{p}}$, $\hat{\mathbf{q}}$, $\hat{\mathbf{s}}$ are aligned with the body's principal axes of inertia. These unit vectors must be found from the time-varying transformation matrix for the rigid-body motion. For the asteroid Toutatis we chose to make the long-axis of that body to be the $\hat{\mathbf{p}}$-axis, giving it a positive C_{20} coefficient (i.e., a prolate body) and a very small C_{22} coefficient. In the following the specific decomposition of the body need not be specified to evaluate the relevant perturbation equations.

Many of the analysis techniques outlined in the previous chapter are still applicable, thus we will focus mostly on those which have clearly changed. These are the computation of periodic orbits and their use to characterize a dynamical system and the application of averaging.

8.3 Analytical Characterization

A large fraction of asteroids in complex rotation are characterized by having a slow overall rotation rate. The asteroid specifically discussed in this section, Toutatis, is a clear example of this having an inertial period of approximately 7.54 days and a body-fixed periodic angular velocity vector of 5.42 days. One result of this is that most orbital motion of interest about these bodies is within the resonance

region about these bodies. Orbit mechanics in these situations can become much different, with the gravity field essentially "dragging" the particle orbit about it (see [72] for a discussion of this case for a uniformly rotating asteroid). For such slow rotators, when at resonant distances one also finds that the perturbations from the gravity field are reduced due to the larger distances. Due to these physical effects, it is possible to average the perturbing potential over a particle's orbit period while keeping the mass distribution fixed in inertial space, i.e., keeping the unit vectors $\hat{\boldsymbol{p}}$, $\hat{\boldsymbol{q}}$, and $\hat{\boldsymbol{s}}$ stationary. After the averaging is performed, these can be allowed to vary in order to understand the net effect of the body's complex motion. Substituting in the averaged results (Appendix A) the singly averaged perturbation potential is found to be

$$\begin{aligned}\bar{R}_2(t) = &\frac{n^2 C_{20}}{4(1-e^2)^{3/2}}\left[1 - 3\left(\hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{h}}\right)^2\right] \\ &+\frac{3n^2 C_{22}}{2(1-e^2)^{3/2}}\left[\left(\hat{\boldsymbol{q}}\cdot\hat{\boldsymbol{h}}\right)^2 - \left(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{h}}\right)^2\right]\end{aligned} \tag{8.4}$$

where $\hat{\boldsymbol{h}}$ is the unit angular momentum vector and these are still a function of time as the body-fixed unit vectors precess and nutate in time following the torque-free rigid-body solutions.

For computation of the Lagrange Planetary Equations for this singly averaged potential there is no dependence on the mean anomaly or argument of periapsis in the perturbing potential. There is a functional dependence on eccentricity, semi-major axis, inclination and longitude of the ascending node, yielding the following partials:

$$\frac{\partial \bar{R}_2}{\partial a} = -\frac{3}{a}\bar{R}_2 \tag{8.5}$$

$$\frac{\partial \bar{R}_2}{\partial e} = \frac{3e}{1-e^2}\bar{R}_2 \tag{8.6}$$

$$\begin{aligned}\frac{\partial \bar{R}_2}{\partial i} = &-\frac{3n^2}{(1-e^2)^{3/2}} \\ &\left[\frac{1}{2}C_{20}(\hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{h}})(\hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{h}}_i) - C_{22}\left((\hat{\boldsymbol{q}}\cdot\hat{\boldsymbol{h}})(\hat{\boldsymbol{q}}\cdot\hat{\boldsymbol{h}}_i) - (\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{h}})(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{h}}_i)\right)\right]\end{aligned} \tag{8.7}$$

$$\begin{aligned}\frac{\partial \bar{R}_2}{\partial \Omega} = &-\frac{3n^2}{(1-e^2)^{3/2}} \\ &\left[\frac{1}{2}C_{20}(\hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{h}})(\hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{h}}_\Omega) - C_{22}\left((\hat{\boldsymbol{q}}\cdot\hat{\boldsymbol{h}})(\hat{\boldsymbol{q}}\cdot\hat{\boldsymbol{h}}_\Omega) - (\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{h}})(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{h}}_\Omega)\right)\right]\end{aligned} \tag{8.8}$$

$$\frac{\partial \bar{R}_2}{\partial \omega} = 0 \tag{8.9}$$

The evaluations of the partials $(\hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{h}}_i)$ and $(\hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{h}}_\Omega)$ are derived later in the analysis.

8.3.1 C_{20} Dynamics

In the following discussion consider the dynamics under the C_{20} coefficient alone, as this is the major perturbation. If we set $C_{22} = 0$, this implies that the moments of inertia are equal and the rotational dynamics of the object simplify to a constant angle between the symmetry axis and the angular momentum with a constant precession rate about the angular momentum. Let us assume that the total rotational angular momentum of the rotating body is aligned with the inertial $\hat{\boldsymbol{z}}$-axis. The axis of symmetry is then defined by a constant declination angle δ and a right ascension α, defined as

$$\hat{\boldsymbol{p}} = \cos\delta\sin\alpha\hat{\boldsymbol{x}} - \cos\delta\cos\alpha\hat{\boldsymbol{y}} + \sin\delta\hat{\boldsymbol{z}} \tag{8.10}$$

where $\alpha = \dot{\alpha}(t - t_o) + \alpha_o$. The unit angular momentum vector, the node vector and the transverse vectors are then

$$\hat{\boldsymbol{h}} = \sin i\sin\Omega\hat{\boldsymbol{x}} - \sin i\cos\Omega\hat{\boldsymbol{y}} + \cos i\hat{\boldsymbol{z}} \tag{8.11}$$

$$\hat{\boldsymbol{n}} = \cos\Omega\hat{\boldsymbol{x}} + \sin\Omega\hat{\boldsymbol{y}} \tag{8.12}$$

$$\hat{\boldsymbol{t}} = -\sin\Omega\cos i\hat{\boldsymbol{x}} + \cos\Omega\cos i\hat{\boldsymbol{y}} + \sin i\hat{\boldsymbol{z}} \tag{8.13}$$

and thus

$$(\hat{\boldsymbol{h}}\cdot\hat{\boldsymbol{p}}) = \cos i\sin\delta + \sin i\cos\delta\cos(\Omega - \alpha) \tag{8.14}$$

Define the relative node angle $\lambda = \Omega - \alpha$, which measures the orbit node relative to the rotation pole right ascension. This leads to

$$(\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{p}}) = \cos\delta\sin(\lambda) \tag{8.15}$$

$$(\hat{\boldsymbol{t}}\cdot\hat{\boldsymbol{p}}) = \sin i\sin\delta - \cos i\cos\delta\cos(\lambda) \tag{8.16}$$

Also note the following

$$\hat{\boldsymbol{h}}_\Omega = \sin i\hat{\boldsymbol{n}} \tag{8.17}$$

$$\hat{\boldsymbol{h}}_i = -\hat{\boldsymbol{t}} \tag{8.18}$$

The Lagrange equations are then (modifying $\dot{\Omega}$ to $\dot{\lambda} = \dot{\Omega} - \dot{\alpha}$)

$$\dot{a} = 0 \tag{8.19}$$

$$\dot{e} = 0 \tag{8.20}$$

$$\dot{i} = -\frac{3}{2}\frac{nC_{20}}{p^2}(\hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{h}})(\hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{n}}) \tag{8.21}$$

$$\dot{\lambda} = \frac{3}{2}\frac{nC_{20}}{p^2}\csc i(\hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{h}})(\hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{t}}) - \dot{\alpha} \tag{8.22}$$

$$\dot{\omega} = \frac{3}{4}\frac{nC_{20}}{p^2}\left[1 - 3(\hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{h}})^2 - 2\cot i(\hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{h}})(\hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{t}})\right] \tag{8.23}$$

$$\dot{\sigma} = \frac{3nC_{20}\sqrt{1-e^2}}{4p^2}\left[1 - 3\left(\hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{h}}\right)^2\right] \tag{8.24}$$

with the dot products given above.

Of specific interest for orbit design are orbits that do not change secularly in time, termed "frozen orbits" as at least some of their orbit elements are constant on average. The orbit elements a and e are both frozen under the C_{20} gravity field perturbation. By choosing the node vector $\hat{\boldsymbol{n}}$ to be orthogonal to the pole vector, $\hat{\boldsymbol{p}}$, specifically choosing $\lambda = m\pi$, then the inclination will not change on average either. Just choosing the node vector thus will not necessarily maintain the condition on λ unless $\dot{\lambda} = 0$ as well. These conditions are considered below.

First note for $\lambda = m\pi$ that $\hat{\boldsymbol{h}}\cdot\hat{\boldsymbol{p}} = \sin(\delta \pm i)$ and $\hat{\boldsymbol{t}}\cdot\hat{\boldsymbol{p}} = \mp\cos(\delta \pm i)$, where $+$ is for m even or zero and $-$ is for m odd. The condition for λ to maintain its value then requires the secular trend in the node to balance against the precession rate of the asteroid

$$\dot{\alpha} = \mp\frac{3}{4}\,\frac{nC_{20}}{p^2}\,\frac{\sin 2(\delta \pm i)}{\sin i} \tag{8.25}$$

This provides a constraint between the inclination, semi-major axis and eccentricity for a given rotation state consisting of $\dot{\alpha}$ $(= 9.894 \times 10^{-6}$ rad/s) and δ $(= 40°)$. Given that $\dot{\alpha}$ is positive yields the constraint $\mp\sin 2(\delta \pm i) > 0$, which leads to bounds on the inclinations over which frozen orbits can be found: $\frac{\pi}{2} - \delta < i < \pi - \delta$ for $\lambda = 0$ and the two intervals $i < \delta$ and $i > \delta + \frac{\pi}{2}$ for $\lambda = \pi$. Assuming these constraints and solving for the semi-major axis yields

$$a^{7/2} = \mp\frac{3\sqrt{\mu}C_{20}}{4\dot{\alpha}(1-e^2)^2}\,\frac{\sin 2(\delta \pm i)}{\sin i} \tag{8.26}$$

$$= \mp 66.668\frac{1}{(1-e^2)^2}\,\frac{\sin 2(\delta \pm i)}{\sin i} \tag{8.27}$$

The final element that can be frozen is the argument of periapsis. This condition is a function of inclination alone, requiring that the following transcendental equation be equal to zero

$$\frac{3}{2}\cos 2(\delta \pm i) - \frac{1}{2} \pm \cot i \sin 2(\delta \pm i) = 0 \tag{8.28}$$

For the Toutatis value of $\delta = 40°$ there are zeros at 25.0°, 102.3° and 151.6° for $\lambda = 0$ and at 27.7°, 79.6° and 154.1° for $\lambda = \pi$. In Fig. 8.1 the semi-major axis versus inclination for a frozen orbit to exist is plotted, and note that not all of the inclination values where the argument of periapsis is frozen are allowed. These can all be generalized to eccentric orbits, which will shift the semi-major axes to larger values in general without changing the frozen periapsis conditions.

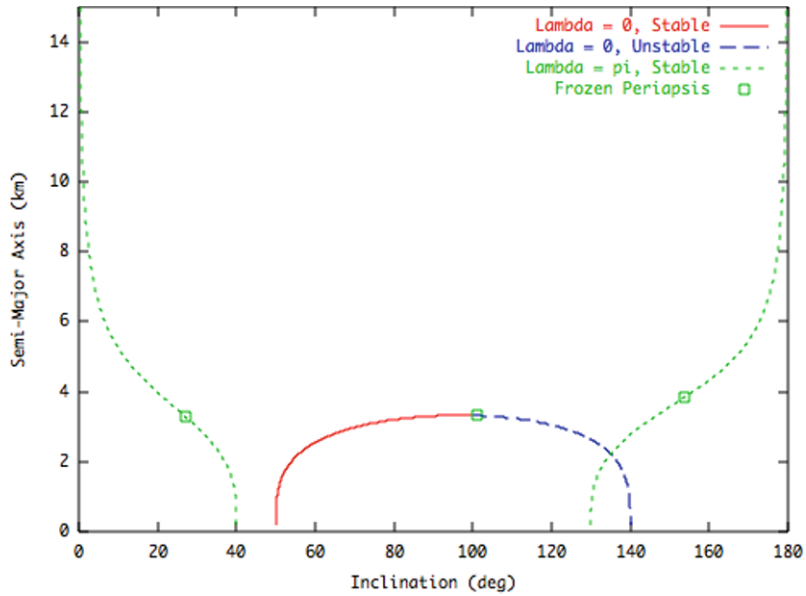


Fig. 8.1 Frozen orbit conditions for a Toutatis orbiter.

It is also instructive to carry out a stability analysis of these frozen orbits under deviations in inclination and relative node. Taking a variation in the inclination and relative node for the Lagrange equations of these elements and evaluating them at the frozen orbit conditions yields

$$\Delta\dot{i} = \mp\frac{3}{2}\frac{nC_{20}}{p^2}\sin(\delta \pm i)\cos\delta\Delta\lambda \tag{8.29}$$

$$\Delta\dot{\lambda} = \frac{3}{2}\frac{nC_{20}}{p^2}\left[\pm\frac{\cos(i)}{2\sin^2(i)}\sin 2(\delta \pm i) - \frac{\cos 2(\delta \pm i)}{\sin(i)}\right]\Delta i \tag{8.30}$$

Using the fact that

$$\dot{\alpha} = \mp\frac{3}{4}\frac{nC_{20}}{p^2}\csc(i)\sin 2(\delta \pm i)$$

these equations can be simplified and reduced to one second-order differential equation

$$\Delta\ddot{\lambda} + \dot{\alpha}^2\left[1 \mp \tan(i)\cot 2(\delta \pm i)\right]\frac{\cos\delta\cos(i)}{\cos(\delta \pm i)}\Delta\lambda = 0 \tag{8.31}$$

where the coefficient of the $\Delta\lambda$ term must be positive for the solution to be stable. Also indicated on Fig. 8.1 are the stability regions of the frozen orbits. Only the $\lambda = \pi$ orbits have fully frozen, stable orbits possible, as the only fully frozen $\lambda = 0$ orbit occurs at the transition between stability and instability. The stability of these

orbits is evaluated numerically under the full second-degree and -order gravity field in [173].

Thus there are three families of stable frozen-orbits for circular orbits, and two stable frozen orbit solutions for eccentric orbits, each with the right ascension of the rotation pole aligned with the descending node of the orbit ($\lambda = \pi$) and inclinations around 30° and 150° (we do not consider the frozen periapsis solution at the stability transition to be stable). Inclusion of higher-order dynamics can certainly affect these frozen orbits. For the three frozen angles, inclination, relative node, and argument of periapsis, higher-order terms in the gravity field can be offset by adjusting the nominal values. This is so as the gradients of the Lagrange equations for these elements are non-zero at the frozen orbits, and thus adjustments can be made in these angular values to null out a small perturbation. The same is not true for the orbit elements that are uniformly zero, namely semi-major axis and eccentricity. A higher-order perturbation in one of these cannot be nulled, and frozen orbit solutions for these higher-order perturbations must be found themselves. The semi-major axis will generally remain conserved so long as the averaging assumptions apply. The same is not true for the eccentricity, however, which is influenced by the C_{30} gravity term, a situation which is considered in more detail in [173].

8.3.2 Doubly-Averaged Solutions

For orbits far from Toutatis, beyond a semi-major axis of ∼23 km, the rotation pole of the body will precess faster than the mean motion of the orbit. In this situation the orbital dynamics can be analyzed by averaging a second time, over the time-varying rotation state of the asteroid. This averaging will be acceptable in general where there remains a non-resonance condition between the precession period of the rotation pole and the orbit period. Specifically, this implies that the relative node of the orbit is not constant ($\dot{\lambda} \neq 0$) and that there is not a mean motion resonance between the orbit and the rotational dynamics. Orbits when such resonances exist are considered in the following section. For simplicity consider the case where $C_{22} \sim 0$ and the rotation pole dynamics follow a simple rotation of the right ascension with the declination constant.

The second time average of R_2 is evaluated by varying the pole right ascension α over one precession period, specifically the time average of $(\hat{\boldsymbol{h}} \cdot \hat{\boldsymbol{p}})^2$ while keeping the angular momentum vector constant. Expanding this product yields

$$\cos^2 i \sin^2 \delta + 2 \sin i \cos i \sin \delta \cos \delta \cos(\Omega - \alpha) + \sin^2 i \cos^2 \delta \cos^2(\Omega - \alpha) \quad (8.32)$$

The averaging operator is $\frac{1}{2\pi} \int_0^{2\pi} (-)\, d\alpha$ and yields

$$\cos^2 i \sin^2 \delta + \frac{1}{2} \sin^2 i \cos^2 \delta \quad (8.33)$$

Substituting this back into the singly-averaged R_{20} potential yields

$$\overline{\overline{R}}_{20} = \frac{n^2 C_{20}}{4(1-e^2)^{3/2}} \left[1 - 3\cos^2 i \sin^2 \delta - \frac{3}{2} \sin^2 i \cos^2 \delta \right] \tag{8.34}$$

If $\delta \to \pi/2$, which is the usual case of a uniformly spinning rotationally symmetric body, the classical singly-averaged potential for the leading oblateness term is recovered.

8.4 Phase Space Characterization

Analytical results are large limited to the application of averaging approximations. To progress further in the investigation of this system it is necessary to use numerical integrations. The results shown in this section all model the full Toutatis gravity field without any of the preceding simplifications and provides a way to verify the analysis and the assumptions made above.

8.4.1 Numerical Integrations

Some general numerical integrations have been made to show the severity of the gravity field perturbations on an orbiter, with the results presented graphically. In Fig. 8.2 three different trajectories are shown. The plots on the right are shown relative to inertial space while the plots on the left are shown from a frame fixed and rotating with the asteroid. The upper diagrams correspond to a ballistic trajectory from the surface that is less than escape speed, yet which travels relatively far from the asteroid. As the body travels far from the body the non-spherical mass distribution is not as important and the inertial trajectory looks like, and is close to, an elliptic orbit (note that the inertial attitude of Toutatis at re-impact of the particle is shown on the right). On the upper-left, the complex rotation of Toutatis plays a role in the trajectory relative to the body-fixed frame, enabling the ballistic trajectory to travel to the far side of the asteroid (essentially meaning that the asteroid has rotated beneath this orbit to present a new face to the re-impacting particle). The middle frames correspond to a trajectory lofted with a relatively low speed, with the particle being captured into orbit initially due to the asteroid rotating such that the first periapsis was not under the surface. After this the trajectory is strongly influenced by the non-spherical gravity field of the asteroid and its rotation state. It is specifically interesting to note that in the inertial frame the orbit plane undergoes a large rotation from orbit to orbit. Viewing the body-fixed frame it becomes clear that the asteroid's mass distribution is actually dragging the orbit frame with itself, as it undergoes its complex rotation. This motion persists for several revolutions before impact occurs again. Finally, in the bottom images a trajectory is shown that is given a very small lofting speed. The particle remains trapped in orbit for an extremely long time, on the order of months, before it re-impacts. The fact that this occurs for the lowest energy trajectory lofted

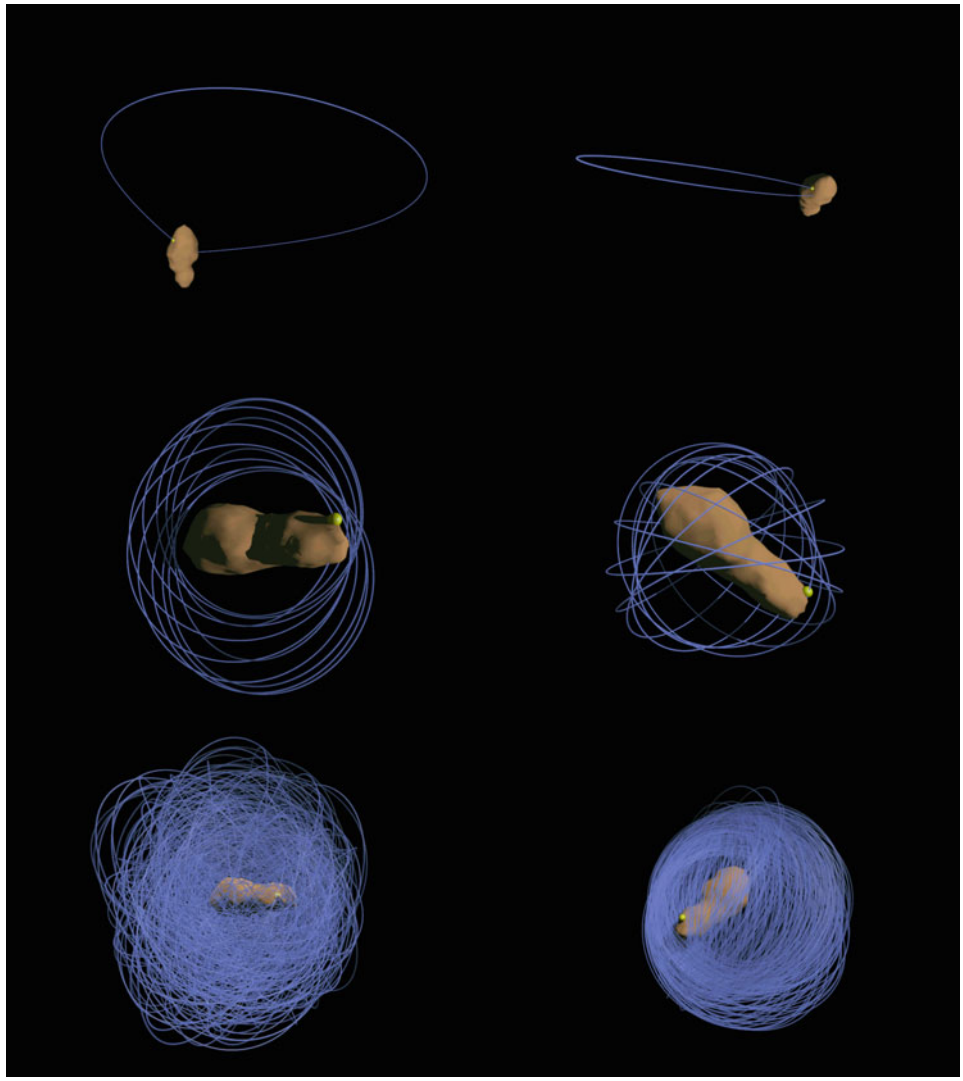


Fig. 8.2 Ballistic trajectories about Toutatis. The images on the right these are viewed from an inertial frame. The images on the left are viewed in a frame rotating with Toutatis. Credit: DIAL/JPL-Caltech.

from the surface points out a potential concern for human missions to asteroids. Should activity on the surface of such a body stir up a cloud of particles, then it may be possible to create a transient atmosphere of rocks that persist in orbit for very long timespans.

Figure 8.3 shows the time evolution of an orbit chosen to satisfy the stable frozen orbit condition derived above. This figure explicitly shows the same behavior as seen in the previous figure, with the asteroid mass distribution dragging the orbit with

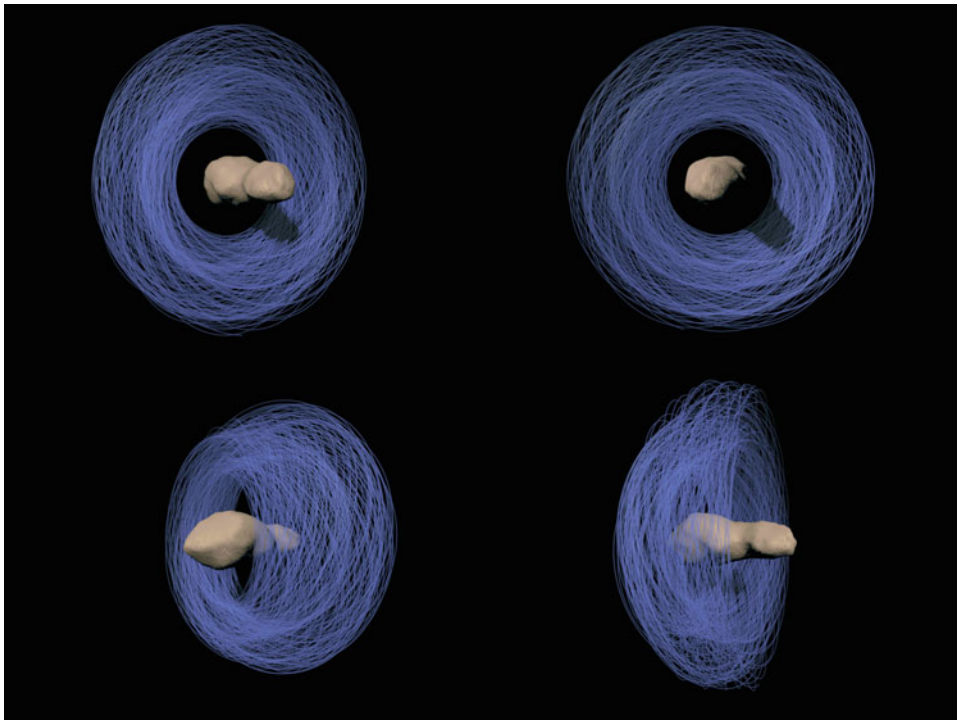


Fig. 8.3 Numerically integrated orbit about Toutatis, chosen according to the "frozen orbit" parameters, shown in a frame that rotates with the asteroid. Credit: DIAL/JPL-Caltech.

it as it rotates in inertial space. This orbit is stable in the sense that it will persist for an arbitrarily long time in the absence of other perturbations.

8.4.2 Periodic Orbits

The time-varying nature of the equations of motion destroy the existence of the Jacobi integral and of equilibrium points in these systems. Thus, two tools for understanding the dynamics about this system are not available to us. The only systematic solutions which are available to us are periodic orbits.

Periodic orbits in a time-periodic system have several properties not shared by orbits in time-invariant systems. First, the variation matrix associated with one period does not have unity eigenvalues in general. This is analogous to the monodromy matrix in a time invariant system, defined such that the energy integral has been removed from the linear variation. In those systems this implied that a periodic orbit was isolated in phase space at a fixed level of energy. For a time-periodic system the lack of an energy integral implies that the periodic orbit is isolated in the full six-dimensional phase space. Thus, these periodic orbits are similar to equilibria in time invariant systems in that they only exist at discrete points. In many ways this property makes it algorithmically easier to compute

periodic orbits as one can use the full state transition matrix without reduction. Conversely, the initial conditions of these orbits must be identified independent of a family, meaning that one cannot compute one family member and then trace out the family as the parameter varies. Choosing initial estimates for these periodic orbits then becomes an exercise in appropriately identifying candidate solutions, generally resonant orbits in the unperturbed problem.

First consider the types of periodic orbits which can be found. The periods of these orbits must be of the form $T = nT_p$, where T_p is the period of the angular velocity vector in the body-fixed frame and n is an integer. Despite this fact, it is convenient to discuss these orbits as $m : n$ orbits, where m represents the number of spin periods the rigid body has covered and n represents, approximately, the number of orbits the particle has made about the body in inertial space. This is related to the generation of initial guesses for these orbits to feed into an iteration scheme for their computation. Independent of the body-fixed rotation period, the orbit and mass distribution must repeat their configuration relative to each other in inertial space for a periodic orbit to exist. By using the unperturbed motion as the first estimate for this, periodic orbits can be located and numerically converged upon.

In Figs. 8.4 to 8.7 a series of numerically computed periodic orbits for the Toutatis system are presented. We make a special note of the intrinsic beauty that these mathematically defined objects have, especially when viewed relative to the general, non-symmetric mass distribution, as they close upon themselves.

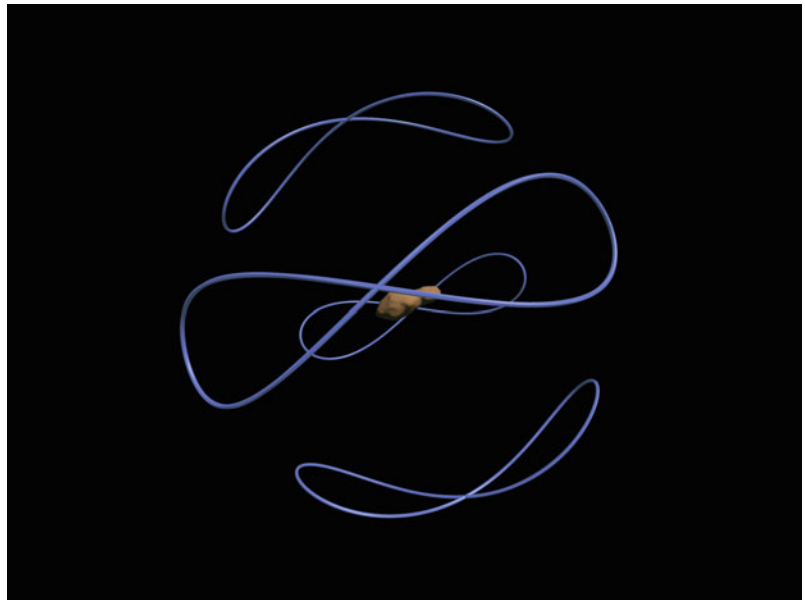


Fig. 8.4 Direct 1:1 periodic orbits about Toutatis. Two of these orbits are stable and two are unstable. Credit: DIAL/JPL-Caltech.

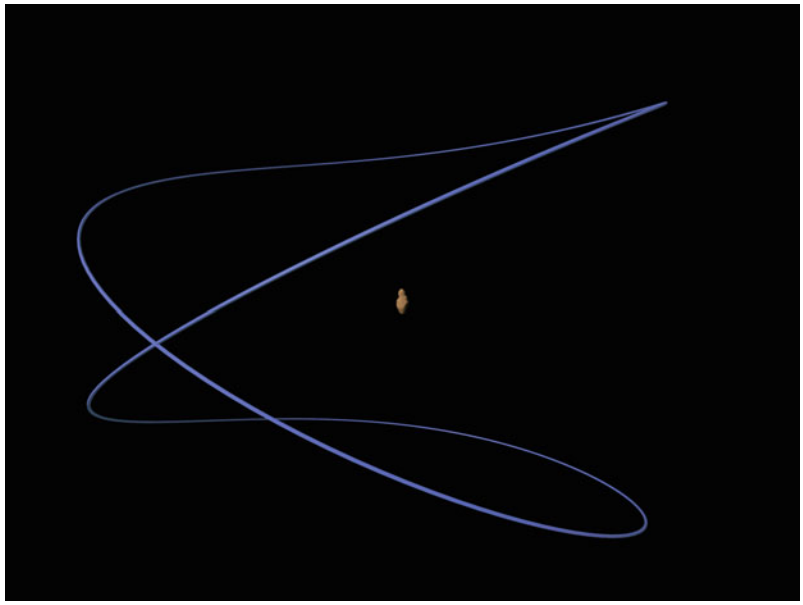


Fig. 8.5 Stable retrograde 1:1 periodic orbits about Toutatis. Credit: DIAL/JPL-Caltech.

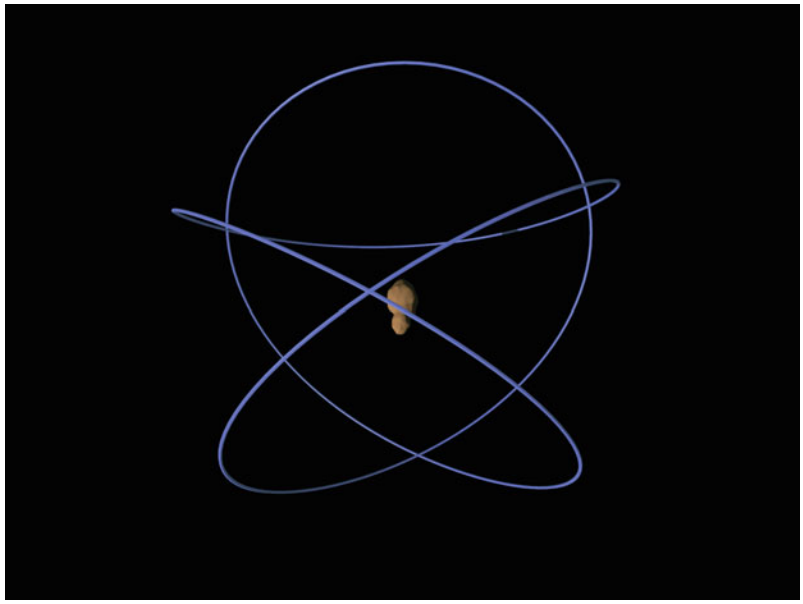


Fig. 8.6 Stable retrograde 1:2 periodic orbits about Toutatis. Credit: DIAL/JPL-Caltech.

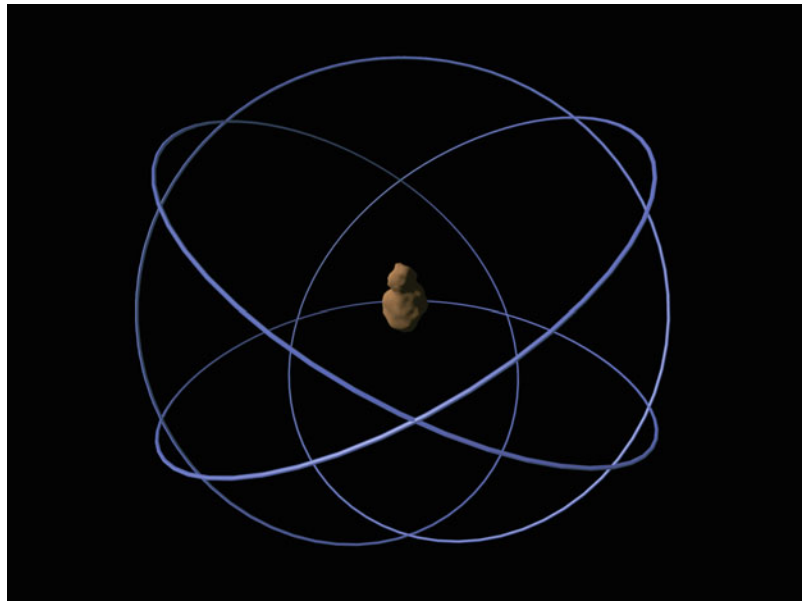


Fig. 8.7 Stable retrograde 1:3 periodic orbits about Toutatis. Credit: DIAL/JPL-Caltech.

9. Binary Asteroids: 1999 KW4

Binary asteroids are estimated to constitute over 15 percent of the Near-Earth Asteroids and may be as prevalent in the Main Belt as well. There are many important issues of science that can be investigated by studying and sending probes to investigate these systems, and their study has been one of the main foci of asteroid science over the last decade. Thus, it is highly probable that a future space mission will visit a binary asteroid and need to carry out operations in its vicinity. The problem of trajectory design and navigation in these systems is complex and challenging, as it requires modeling of the dynamics of the asteroid system in addition to the motion of the spacecraft about the system. In the most general sense the problem of binary asteroids orbiters integrates four classical problems of astrodynamics: the Hill problem, the restricted three-body problem, the non-spherical orbiter problem and the full two-body problem [162], constituting a complex simulation problem. The general problem has been considered in detail in a number of analyses [165, 36, 38, 37, 35], all involving the solution of the full two-body problem for the gravitating components of the binary asteroid, a topic beyond the current book (Fig. 9.1 shows the general geometric description of this problem). Thus in this chapter we introduce the simplest model for a binary asteroid, which is still sufficient for initial mission design and analysis for many of the observed binary systems.

In this simplest model the two component bodies of the binary are modeled as a sphere and a tri-axial ellipsoid, with the satellite assumed to have no influence on the motion of the two primaries. This particular problem will be referred to as the (sphere)-restricted full three-body problem. The general form of the equations of motion were first derived in [162]. Further, we will assume that the two bodies are in a relative equilibrium, and thus this system can be transformed into a time invariant Lagrangian through a uniform rotation transformation. As for the restricted three-body problem, five equilibrium can be found and the stability of the analogue L_4, L_5 points are investigated. Some numerical studies of periodic orbits in this system are also presented, using the binary asteroid 1999 KW4 as a motivating model.

D.J. Scheeres, *Orbital Motion in Strongly Perturbed Environments*, Springer Praxis Books,
DOI 10.1007/978-3-642-03256-1_9, © Springer-Verlag Berlin Heidelberg 2012

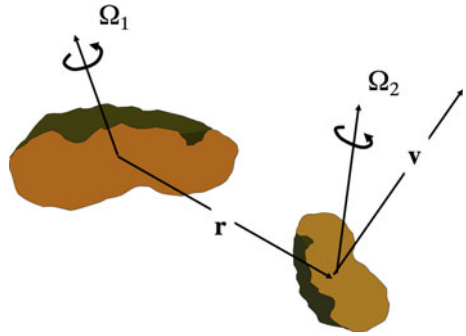


Fig. 9.1 The full two-body problem involves the coupled rotational and translational motion of two gravitating bodies.

9.1 Model of 1999 KW4

In Table 9.1 we present the dimensional model parameters for the binary asteroid 1999 KW4, while in Table 9.2 we present normalized values used later in our numerical analysis. In reference to this model, the larger component of the binary is called "Alpha" and the smaller component "Beta." Note that detailed shape models exist for each of these bodies and their motion has been analyzed extensively [121, 165]. Despite this, only a simplified form of the model is presented for definiteness in the following discussion. The largest approximation is modeling the primary as a sphere, when in actuality it is closer to an oblate spheroid. Incorporation of that component of the primary into the current analysis alone would require significant additional discussion on the modeling of the mutual potential of the binary asteroid components, a topic beyond our current scope [194].

Table 9.1 Representative dimensional values and model for 1999 KW4 [121].

Component	Parameter	Value	Units
Orbit:			
	Period	17.42	hours
	Semi-Major Axis	2.55	km
	Eccentricity	~ 0	
	Total Mass	2.50×10^{12}	kg
Alpha:			
	Mass	2.35×10^{12}	kg
	Mean Radius	0.658	km
	Rotation Period	2.76	hours
Beta:			
	Mass	0.135×10^{12}	kg
	Mean Radius	0.225	km
	Semi-Axes	$0.297 \times 0.225 \times 0.171$	km
	Rotation Period	17.42	hrs

Table 9.2 Representative non-dimensional values and model for 1999 KW4.

Component	Parameter	Value
Orbit:		
	Spin rate	0.21487
	Orbit Size	8.565
	Mass Fraction	0.9457
Alpha:		
	Mean Radius	2.215
	Distance from Center of Mass	0.465
Beta:		
	Mean Radius	0.758
	Semi-Axes	$1 \times 0.758 \times 0.576$
	Distance from Center of Mass	-8.100

9.2 The Full Two-Body Problem

In order to find the equations of motion of a spacecraft in the vicinity of a binary system, it is necessary to first consider the full two-body problem, which describes how the binary asteroid components move relative to each other. At the simplest level this can be captured by the restricted three-body problem, treating each binary asteroid component as a sphere. Our current analysis is more generalized than this, allowing for one of the asteroid components to be an ellipsoid while keeping the other as a sphere. If non-spherical bodies are to be modeled for both components the mutual potential between these bodies must be considered, which becomes more complex (reviewed in [194, 36]).

For the current problem statement M_s is the mass of the sphere and M_e is the mass of the constant density tri-axial ellipsoid. The mass fraction of the two primaries is defined as,

$$\nu = \frac{M_s}{M_e + M_s} \tag{9.1}$$

The relative position vector of the sphere with respect to the ellipsoid center of mass is $\boldsymbol{R}_r$. Then, relative to their center of mass, the positions of the asteroid components are,

$$\boldsymbol{R}_e = -\nu \boldsymbol{R}_r \tag{9.2}$$

$$\boldsymbol{R}_s = (1 - \nu)\boldsymbol{R}_r \tag{9.3}$$

where subscripts s and e refer to the sphere and the ellipsoid, respectively.

For a coordinate frame fixed to the ellipsoid, the motion of the sphere relative to the ellipsoid is modeled as [160]:

$$\ddot{\boldsymbol{R}}_r + 2\boldsymbol{\Omega} \times \dot{\boldsymbol{R}}_r + \dot{\boldsymbol{\Omega}} \times \boldsymbol{R}_r + \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \boldsymbol{R}_r) = G(M_e + M_s)\frac{\partial \mathcal{U}_e}{\partial \boldsymbol{R}_r} \tag{9.4}$$

The vector $\boldsymbol{\Omega}$ is the angular velocity of the ellipsoid, which follows Euler's equation

$$\boldsymbol{I}\cdot\dot{\boldsymbol{\Omega}}+\boldsymbol{\Omega}\times\boldsymbol{I}\cdot\boldsymbol{\Omega}=-G(M_e+M_s)\boldsymbol{R_r}\times\frac{\partial\mathcal{U}_e}{\partial\boldsymbol{R}_r} \tag{9.5}$$

The inertia dyadic of the ellipsoid is represented as $\boldsymbol{I}$ and $\mathcal{U}_e$ is the "unit gravitational parameter" mutual potential between the sphere and ellipsoid. It can be shown that this is precisely equal to the gravitational potential of the ellipsoid, which is found in Section 2.5, Eq. 2.32. In general, $\mathcal{U}_e$ is a function of the ellipsoid semi-major axes, α, β, and γ, and the position vector from the ellipsoid to the sphere, $\boldsymbol{R}_r$.

To simplify the analysis introduce the following non-dimensionalization of the problem. The maximum radius of the ellipsoid, denoted as α, and the mean motion of the system at this distance,

$$n=\sqrt{\frac{G(M_e+M_s)}{\alpha^3}}$$

are taken as length and frequency scales, respectively. Therefore, the normalized position and angular velocity are

$$\boldsymbol{R}=\frac{\boldsymbol{R}_r}{\alpha}\quad\text{and}\quad\boldsymbol{\omega}=\frac{\boldsymbol{\Omega}}{n}$$

Equations 9.4 and 9.5 now become

$$\ddot{\boldsymbol{R}}+2\boldsymbol{\omega}\times\dot{\boldsymbol{R}}+\dot{\boldsymbol{\omega}}\times\boldsymbol{R}+\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\boldsymbol{R})=\frac{\partial\mathcal{U}_e}{\partial\boldsymbol{R}} \tag{9.6}$$

$$\boldsymbol{I}\cdot\dot{\boldsymbol{\omega}}+\boldsymbol{\omega}\times\boldsymbol{I}\cdot\boldsymbol{\omega}=-\nu\boldsymbol{R}\times\frac{\partial\mathcal{U}_e}{\partial\boldsymbol{R}} \tag{9.7}$$

where $\mathcal{U}_e$ is now a function of $\bar{\alpha}=1$, $\bar{\beta}=\beta/\alpha$, $\bar{\gamma}=\gamma/\alpha$ and $\boldsymbol{R}$.

Solutions to these coupled differential equations define the relative motion of the sphere and ellipsoid and the spin of the ellipsoid. Of particular interest to the current analysis is the steady-state condition when the ellipsoid rotates synchronously with the mutual orbit. This situation is studied in detail in [160], where it is shown that the stable configuration for such a system is for the sphere to be placed along the ellipsoid's long axis and for the entire system to spin about an axis parallel to the ellipsoid's maximum moment of inertia direction at a rate that counters the gravitational attraction between the two bodies. Setting up this special initial state, and assuming the initial speed of the sphere in the ellipsoid fixed frame to be zero, the equilibrium condition reduces to

$$\boldsymbol{\omega}\times\boldsymbol{\omega}\times\boldsymbol{R}=\frac{\partial\mathcal{U}_e}{\partial\boldsymbol{R}} \tag{9.8}$$

Applying the results from Section 2.5 this condition can be reduced to

$$\omega^2 = \frac{3}{2}\int_{\lambda}^{\infty} \frac{dv}{(\bar{\alpha}^2 + v)\Delta(v)} \tag{9.9}$$

In this case, $\bar{\alpha}$ represents the ellipsoid axis along which the sphere is placed (=1 in our model) and $\lambda = R^2 - \bar{\alpha}^2$, where R is the distance between the primaries. Making the substitution $v = u + \lambda$, the spin rate equation can be simplified to,

$$\omega^2 = \frac{3}{2}\int_{0}^{\infty} \frac{du}{(R^2 + u)\Delta(u)} \tag{9.10}$$

where $\Delta(u) = \sqrt{(\bar{\alpha}^2 + u)\left(\bar{\beta}^2 + u\right)(\bar{\gamma}^2 + u)}$. This relation is precisely equal to the Carlson Elliptic Integral R_J [136].

With this spin rate, the binary asteroid primaries are now in a stable synchronous rotation about each other, forming a relative equilibrium.

9.3 The Restricted Full Three-Body Problem

9.3.1 Equations of Motion of a Particle

Having defined the motion of the binary asteroid primaries in the full two-body problem, the equations of motion for a spacecraft in this system can be defined as the restricted full three-body problem. The barycenter of the binary asteroid system is the coordinate center and the location of the spacecraft is defined by the position vector $\boldsymbol{r}$. Although the mutual potential for the relative motion of these bodies was used, for the gravitational attraction on the spacecraft a different potential must be used that separately accounts for the attraction of the two components. Define a unit gravitational parameter potential for the spacecraft as

$$U_{se}(\boldsymbol{r}, \boldsymbol{R}_r) = \frac{\nu}{|\boldsymbol{r} - (1-\nu)\boldsymbol{R}_r|} + (1-\nu)\mathcal{U}_e(\boldsymbol{r} + \nu\boldsymbol{R}_r) \tag{9.11}$$

where $\mathcal{U}_e$ is the same ellipsoid unit potential as used for the mutual potential. For $\nu \to 0$ the ellipsoid dominates the system and for $\nu \to 1$ the sphere dominates.

Skipping the details of transforming from an inertial frame to a rotating orbit-fixed frame, the equations of motion in a frame fixed with the ellipsoid are

$$\ddot{\boldsymbol{r}} + 2\boldsymbol{\Omega} \times \dot{\boldsymbol{r}} + \dot{\boldsymbol{\Omega}} \times \boldsymbol{r} + \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \boldsymbol{r}) = G(M_e + M_s)\frac{\partial U_{se}}{\partial \boldsymbol{r}} \tag{9.12}$$

where $\boldsymbol{\Omega}$ is driven by Eq. 9.5. These equations do not make any assumptions on the binary system being in a relative equilibrium, evident through the inclusion of the $\dot{\boldsymbol{\Omega}}$ term.

The same normalizations are introduced as for the full two-body problem, defining a new position vector $\boldsymbol{\rho} = \frac{\boldsymbol{r}}{\alpha}$ and the same time scaling. Then Eq. 9.12 becomes

$$\ddot{\boldsymbol{\rho}} + 2\boldsymbol{\omega} \times \dot{\boldsymbol{\rho}} + \dot{\boldsymbol{\omega}} \times \boldsymbol{\rho} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{\rho}) = \frac{\partial U_{se}}{\partial \boldsymbol{\rho}} \tag{9.13}$$

where $U_{se}(\boldsymbol{\rho}, \boldsymbol{R})$ is functionally unchanged and $\boldsymbol{\omega}$ is satisfied by Eq. 9.7. Figure 9.2 shows the geometry for the scaled problem.

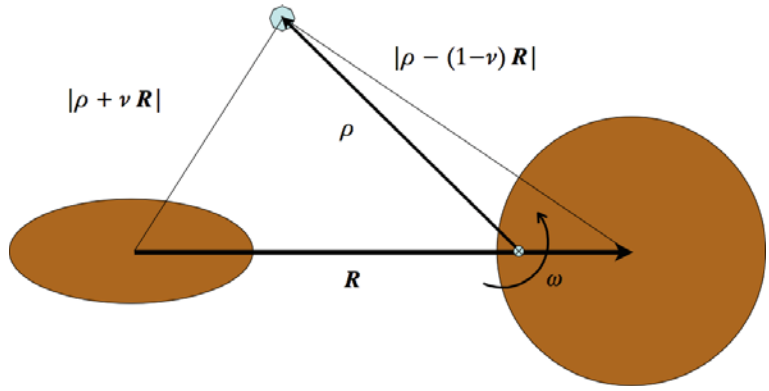


Fig. 9.2 Geometry of the restricted full three-body problem.

If the binary asteroid is in a relative equilibrium, as defined above, the only change is to set $\dot{\boldsymbol{\omega}} = 0$, treat $\boldsymbol{R}$ as constant vector and solve for ω from Eq. 9.9. In this case, the system has a Jacobi integral, specified as

$$J = \frac{1}{2}\dot{\boldsymbol{\rho}} \cdot \dot{\boldsymbol{\rho}} - \frac{1}{2}\tilde{\boldsymbol{\omega}} \cdot \tilde{\boldsymbol{\omega}} \cdot \boldsymbol{\rho} - U_{se} \tag{9.14}$$

9.3.2 Equilibrium Solutions

If the binary asteroid is in a relative equilibrium, as for the restricted three-body problem five equilibrium solutions exist for the present problem. The three collinear points, analogous to L_1, L_2 and L_3, exist along the axis joining the primaries and are always unstable. However, the two equilibrium points analogous to the $L_{4,5}$ points may be stable with their stability properties affected by the presence of the ellipsoid.

The equilibrium solutions can be obtained from Eq. 9.13 by setting all velocities and accelerations to zero and solving the resulting algebraic equations. Doing so yields

$$\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{\rho}) = \frac{\partial U_{se}}{\partial \boldsymbol{\rho}} \tag{9.15}$$

These can be reduced to three scalar equations

$$\omega^2 x + \frac{\partial U_{se}}{\partial x} = 0 \tag{9.16}$$

$$\omega^2 y + \frac{\partial U_{se}}{\partial y} = 0 \tag{9.17}$$

$$\frac{\partial U_{se}}{\partial z} = 0 \tag{9.18}$$

where expressions for the ellipsoidal portion of the partials can be found in Section 2.5, Eqs. 2.35 to 2.37.

It can be shown that $z = 0$ is required for the third equation to be satisfied, and that $y = 0$ will always satisfy the second equation. For this case the first equation has three roots along the x-axis, qualitatively similar to the L_1, L_2 and L_3 equilibrium points in the restricted three-body problem (R3BP). Finally, the first two equations also have a solution off of the x-axis, analogous to the $L_{4,5}$ solutions in the R3BP. When $\bar{\beta} = \bar{\gamma} \to 1$, the location of these equilibrium points go to $x = (0.5 - \nu)R$ and $y = \sqrt{3}/2R$ for our normalization of the problem, where R is the normalized distance between the two bodies. For non-unity values of $\bar{\beta}$ and $\bar{\gamma}$ the location of these equilibrium points can vary widely, as seen in Fig. 9.3

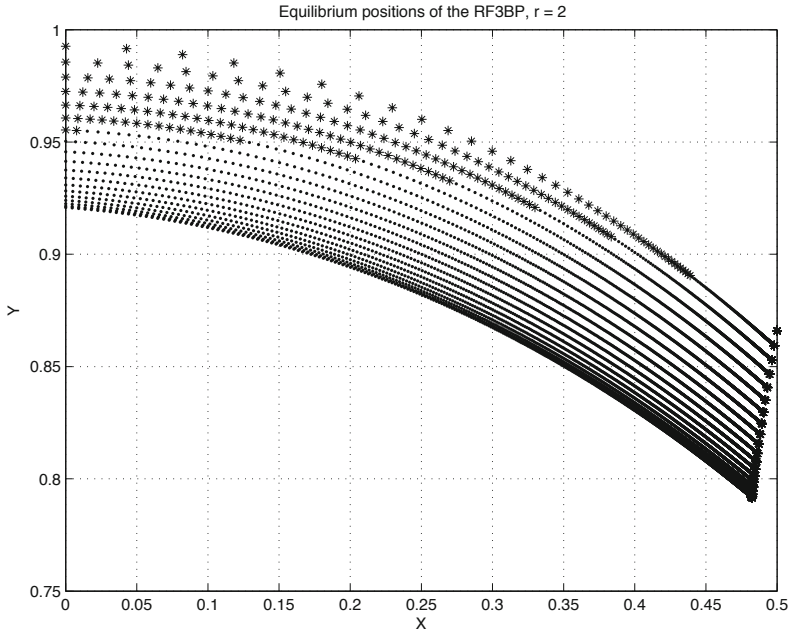


Fig. 9.3 Geometric location of the L_4 analogue libration points for $\bar{\beta} = \bar{\gamma}$ and $R = 2$. The ellipsoid is centered at the origin and the sphere at $(1, 0)$ and is off to the right. In the classical R3BP the location of L_4 is fixed at $(0.5, 0.866)$ for this choice of coordinates. The points go from left to right as ν goes from $0 \to 1$, and from bottom to top as $\bar{\beta}$ goes from $0 \to 1$. The starred points are stable locations of the L_4 analog points. Credit: [10].

For the dimensional parameters of the binary asteroid 1999 KW4 all of the equilibrium points and corresponding zero-velocity curves for that system have been computed in [10] and shown in Fig. 9.4. Qualitatively speaking, the placement of the equilibrium points and the sequence of zero-velocity curves in this system are relatively unchanged from the R3BP. The precise stability conditions for the off-axis points do change, which is discussed in the following sections.

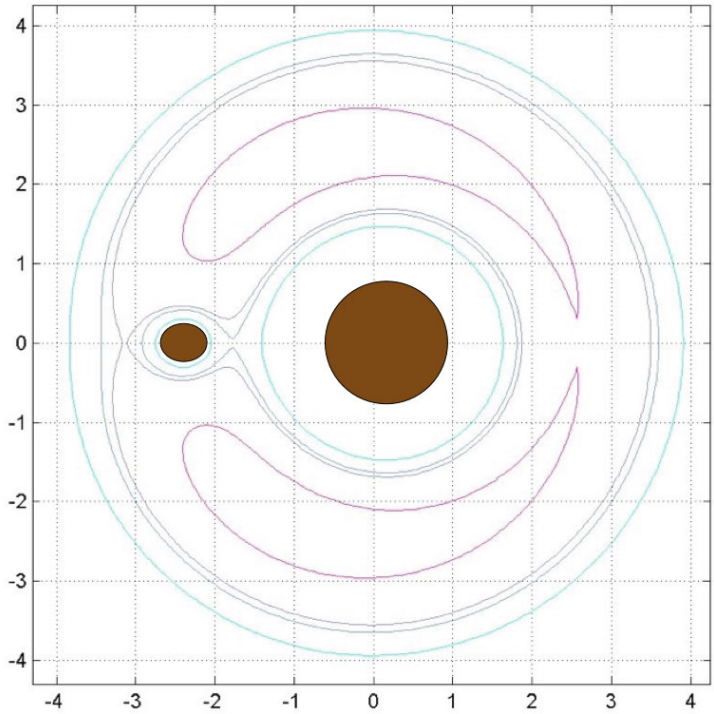


Fig. 9.4 Zero-velocity curves for a particle about the binary asteroid 1999 KW4. The axes are in units of kilometers and the equilibria are defined by the intersections of the zero-velocity curves. Credit: [10].

9.3.3 Stability of the $L_{4,5}$ points

For the R3BP, stability of the equilateral points is given by the Routh criterion, which provides a range of mass fraction values for stability. This can be precisely computed as

$$\nu < \frac{1}{2}\left[1 - \sqrt{\frac{23}{27}}\right] \sim 0.0385 \tag{9.19}$$

$$\nu > \frac{1}{2}\left[1 + \sqrt{\frac{23}{27}}\right] \sim 0.9615 \tag{9.20}$$

The presence of the ellipsoid in the present study can affect the stability of these equilibria. Specifically, they become a function of the semi-major axis ratios and the distance between the bodies in addition to the mass fraction.

It can be shown that the out-of-plane oscillations about these equilibrium points are always stable. The in-plane stability conditions for these equilibria can be reduced to finding the roots of the characteristic equation

$$\begin{Vmatrix} \lambda^2 - \omega^2 - U_{xx} & -2\omega\lambda - U_{xy} \\ 2\omega\lambda - U_{xy} & \lambda^2 - \omega^2 - U_{yy} \end{Vmatrix} = 0 \tag{9.21}$$

where the usual U_{se} subscript is suppressed and the second-order partial derivatives for the ellipsoid component of the potential are given in Section 2.5, Eqs. 2.38 to 2.43. The $L_{4,5}$ points have a non-zero U_{xy} partial.

Expanding the determinant, the characteristic polynomial is

$$\lambda^4 + A\lambda^2 + B = 0 \tag{9.22}$$

where

$$A = 2\omega^2 - U_{xx} - U_{yy} \tag{9.23}$$

and

$$B = \omega^2 + \omega^2(U_{xx} + U_{yy}) + U_{xx}U_{yy} - {U_{xy}}^2 \tag{9.24}$$

Therefore, for a system to be stable, i.e., for the roots to be purely imaginary, the conditions to satisfy are

$$A > 0 \tag{9.25}$$

$$B > 0 \tag{9.26}$$

$$A^2 - 4B > 0 \tag{9.27}$$

Figures 9.5 to 9.7 show results from this stability investigation. In each plot the horizontal axis is the $\bar{\beta}$-value of the ellipsoid, going from 0 to 1 and the vertical axis represents the mass ratio ν of the system, varying from 0 to 1 where the middle portion has been cut out. Each plot shows the Routh criterion as a horizontal line at their respective values of the limit.

Figure 9.5 shows the stability region transition as the ellipsoid goes from being spherical to a highly oblate body with a normalized distance set to 1. Figure 9.6 shows the stability region transition as the ellipsoid again goes from being spherical to highly oblate. In this case, the normalized distance has been set to 2 and the effect of the ellipsoid is seen to be significantly reduced. Figure 9.7 shows the stability region transition as the distance between the bodies is varied from 1 to 4 in normalized units. In this case, the ratio $\gamma/\beta = 1$, making the ellipsoid prolate. The presence of the ellipsoid can have a dramatic influence on the stability of these equilibrium points, but as the distance between the bodies increases the original results from the R3BP are recovered.

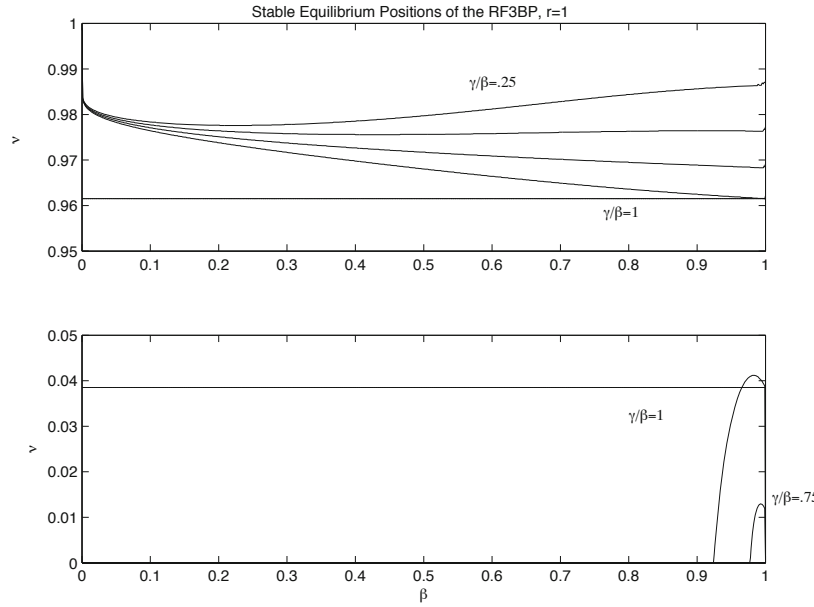


Fig. 9.5 Stability region transition for parameters: $r = 1$, $\gamma/\beta = 0.25 \to 1$. Credit: [10].

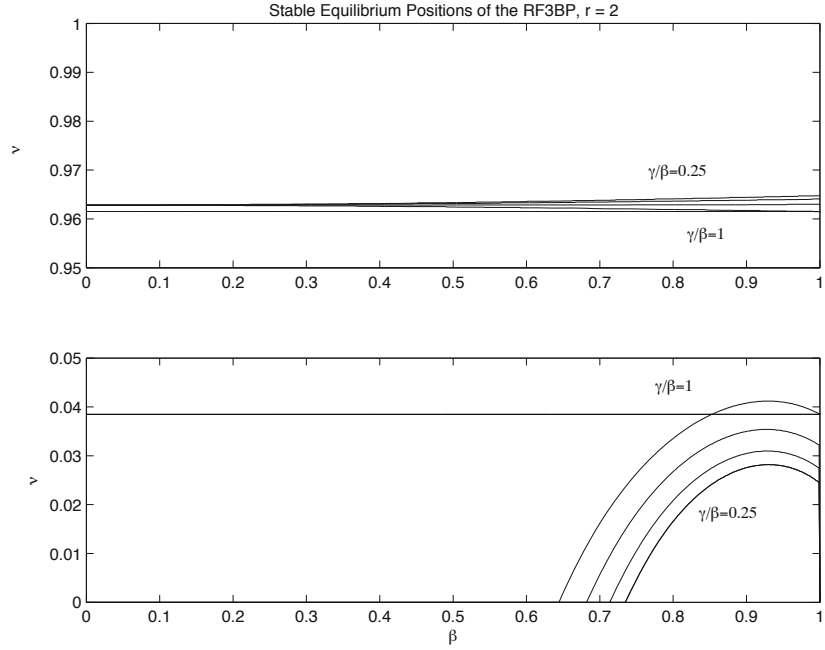


Fig. 9.6 Stability region transition for parameters: $r = 2$, $\gamma/\beta = 0.25 \to 1$. Credit: [10].

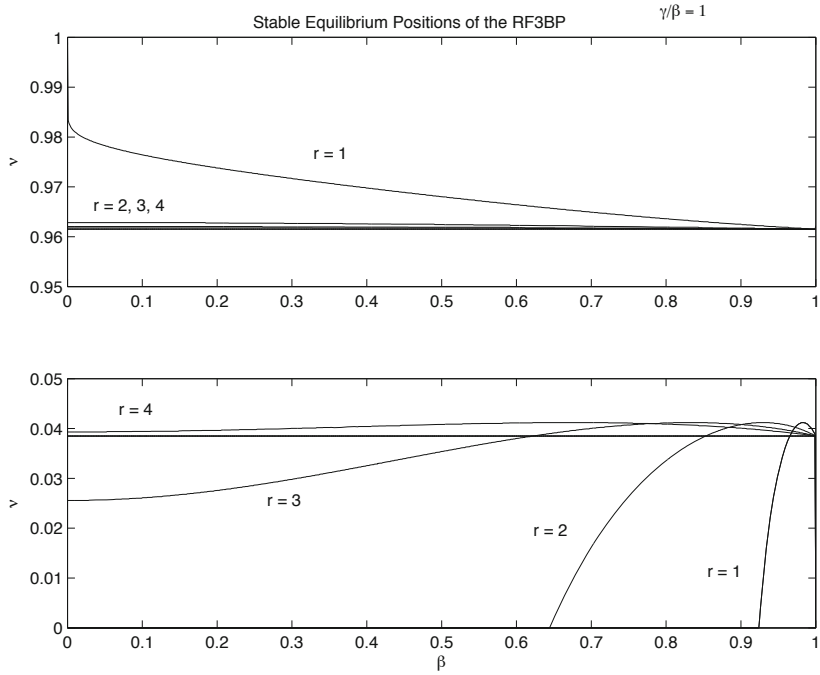


Fig. 9.7 Stability region transition for parameters: $\gamma/\beta = 1$, $r = 1 \rightarrow 4$. Credit: [10].

9.4 Periodic Orbit Computations

To finish this chapter a series of periodic orbit family computations made for the 1999 KW4 system are presented in the normalized units. Only planar orbits that are symmetric with respect to the x-axis are considered, meaning that they make an orthogonal crossing of the x-axis at least twice. For the 1999 KW4 constants the large value of $R \sim 8$ means that the ellipsoid has a relatively minor influence on the orbital dynamics, but the relatively large value of $\nu \sim 0.05$ changes the structure of the families as compared to the more familiar Sun–Jupiter value of $\nu \sim 0.000954$, which is almost two orders of magnitude smaller.

In Figs. 9.8 and 9.9 some specific periodic orbits extracted from several different families are shown, separated by whether they are direct or retrograde. By definition, direct orbits have an inclination value of 0 degrees when viewed from an inertial frame, while retrograde orbits have an inclination value of 180 degrees. These orbits are presented in a rotating frame, and thus their direction of travel in these frames are a combination of their inertial direction of travel and their angular rate as compared to the orbit rate of the binary system.

The direct periodic orbits show a variety of radial variations, ranging from near-circular to highly eccentric. The eccentric orbits can have a characteristic loop or cusp, meaning that the angular rate of the orbit matches or exceeds the rotation rate of the frame, making the orbit change from retrograde to direct in the rotating

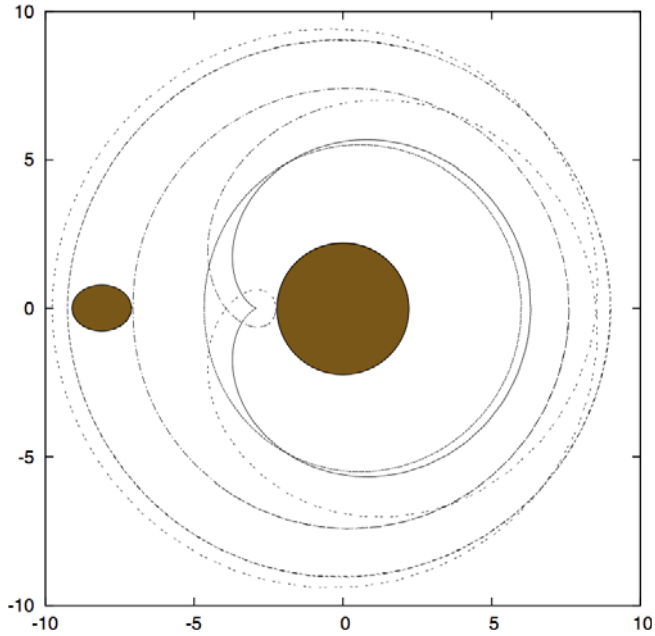


Fig. 9.8 Select direct periodic orbits about the 1999 KW4 system (shown in scaled units).

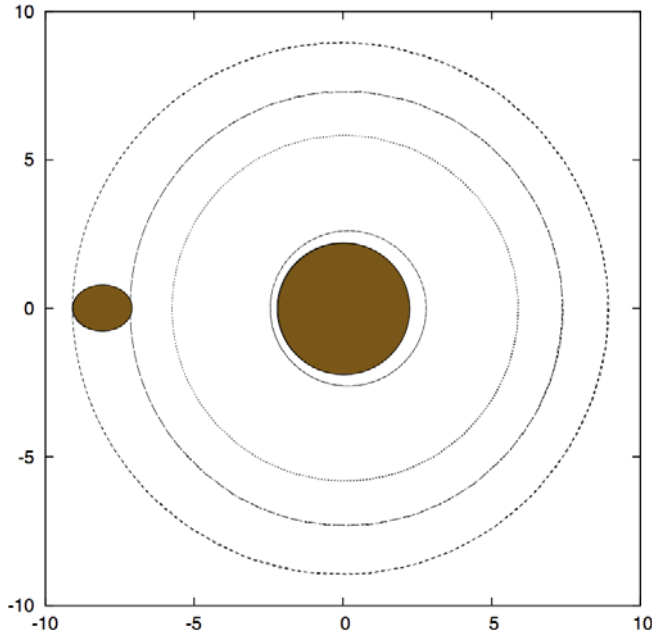


Fig. 9.9 Select retrograde periodic orbits about the 1999 KW4 system (shown in scaled units).

frame, while always being direct in the inertial frame. In Fig. 9.8 all of the periodic orbits except for the largest one are unstable, some with characteristic multipliers that exceed 10. The largest orbit, lying beyond the secondary, is stable and this family continues as a near-circular orbit to arbitrary distance from the system, maintaining stability. As this family is continued to smaller radii they become unstable at x-axis crossings of less than 9.0 units and are discussed in more detail later. Almost all of the periodic orbit families found within the secondary are unstable, implying that it would be challenging to orbit within the secondary in the same sense of motion. There is one short interval of near-circular orbits within the secondary that was found to be stable, with a positive x-axis crossing ranging between 6.5 and 7.5 units.

The retrograde periodic orbits are all uniformly near-circular, although there is an observable degree of eccentricity in the closest orbit shown in Fig. 9.9. The orbits with smallest and largest radii in the figure are both unstable, while the middle two orbits are both stable. There is a larger interval of stable orbits within the secondary for this system. Retrograde orbits with an x-axis crossing between 3.25 and 7 units are stable, with another stable interval between $\sim$ 7.3 and 7.5 units. Retrograde orbits outside of the secondary are generally stable except for an interval that ranges from an x-axis crossing of just less than 9 units to about 9.2 units. The outermost retrograde orbit in Fig. 9.9 is taken from this interval.

In Figs. 9.10 and 9.11 some additional details are presented on two families of interest as they undergo a series of stability changes. Both families are taken from just outside of the secondary, with the first set a direct family and the second a retrograde family. It can be seen that both families go through a series of extrema in the family curves showing one of the perpendicular x-axis crossings versus the Jacobi constant. From the discussion in Chapter 5 one expects to see stability transitions associated with each extremum with respect to the Jacobi constant. In the lower left images from each figure the magnitude of the eigenvalues (i.e., the stability indices) for the families are shown and in the lower right the angles that the eigenvalues make with the real axis in the complex plane are shown. These images taken together allow us to trace the sequence of stability transitions as the family evolves to smaller x-axis crossings.

Taking the direct orbits in Fig. 9.10 first, the orbit eigenvalues are stable and on the unit circle as the Jacobi constant increases (and x-axis crossings decrease), but then at the point labeled 'II' a double bifurcation occurs, with one set of eigenvalues colliding at (1,0) and going unstable (this is the set that generates the extrema) with the other colliding at (-1,0) and going unstable, signifying the intersection with a stable periodic orbit of double the period. This period doubling bifurcation corresponds to an out-of-plane bifurcation. The Jacobi constant then decreases until point 'III' is reached, where the positive real eigenvalues collide again at (1,0) and re-enter the unit circle. The curve undergoes another extremum at this point and the Jacobi constant continues to increase as the x-axis crossing is decreased. The orbit itself remains unstable, with negative real values for two of its eigenvalues, corresponding to an out-of-plane instability.

Now consider the retrograde orbits in Fig. 9.11. First, as the Jacobi constant is decreased the family has a decreasing x-axis crossing. At point 'II' the family curve reaches an extremum and one pair of the eigenvalues collides at (1,0) and departs the unit circle. The family then increases its Jacobi constant as the crossing continues to decrease, seen by the interval of zero angle eigenvalues in the plot on the lower right or by the interval of non-unity eigenvalues on the lower left. At point 'III' the family goes through another extremum and the eigenvalues return to the unit circle, with the family regaining its stability and continuing its initial trend.

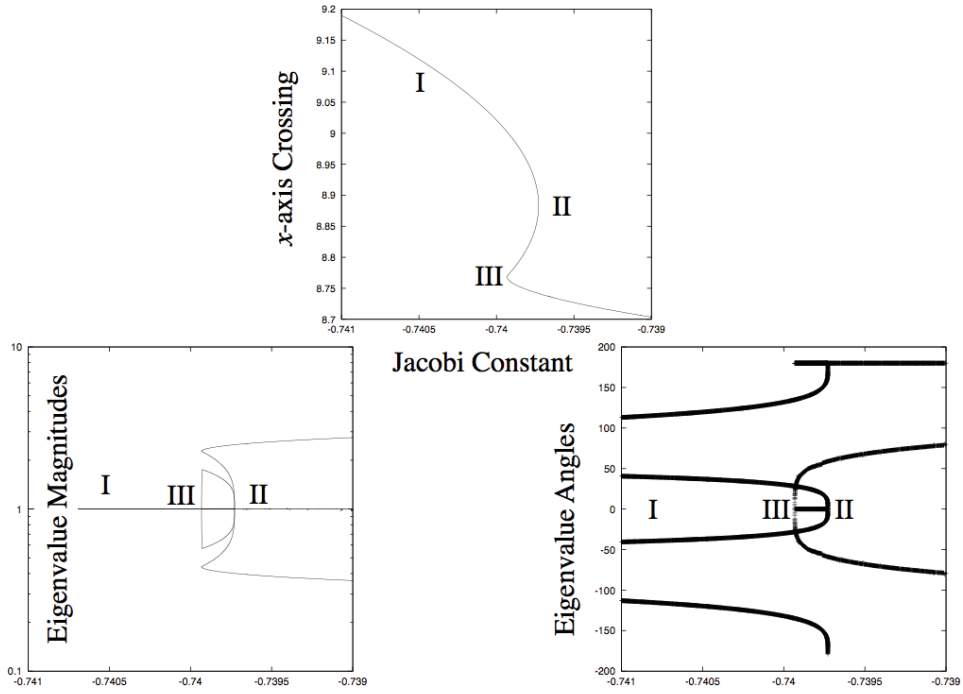


Fig. 9.10 Portion of a direct periodic orbit family about the 1999 KW4 system. The x-axis crossings as a function of Jacobi constant are shown at top, the eigenvalue magnitudes at lower left, and angles with the positive real axis in the complex plane at lower right.

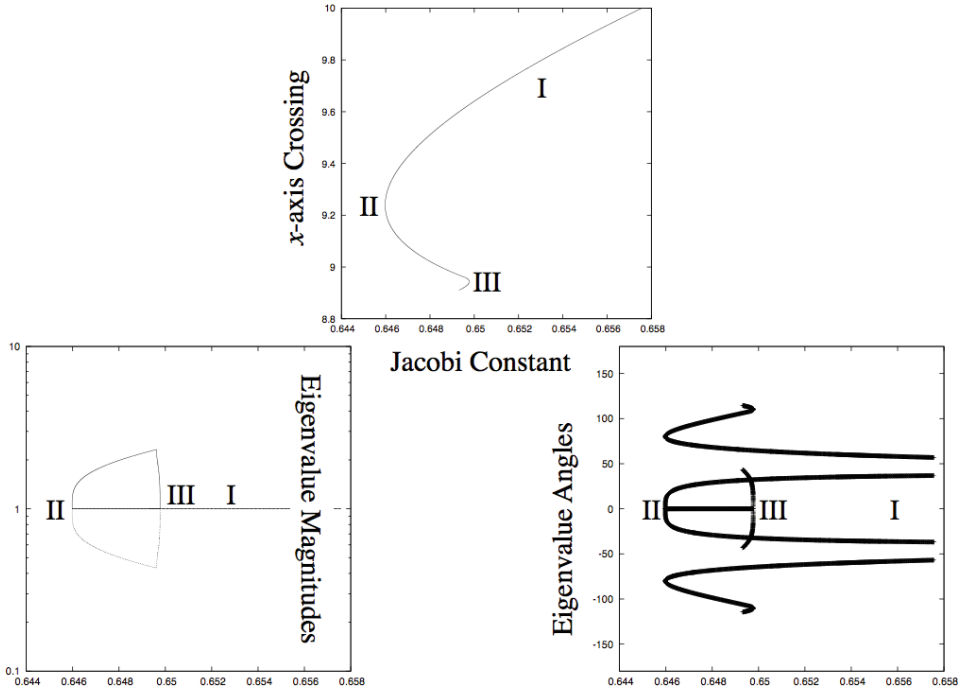


Fig. 9.11 Portion of a retrograde periodic orbit family about the 1999 KW4 system. The x-axis crossings as a function of Jacobi constant are shown at top, the eigenvalue magnitudes at lower left, and angles with the positive real axis in the complex plane at lower right.

10. The Surface Environment on Asteroids

The goal of many space missions to small bodies, such as the Hayabusa mission to Itokawa and the forthcoming OSIRIS-REx mission to asteroid 1999 RQ36, are to take samples from the surface of the body. The prospect of designing a transfer to an asteroid surface and choosing where on the surface to sample begs several questions related to the flow of material across an asteroid surface, possible regions of stable resting points, and the surface forces and slopes that a landing spacecraft will encounter. The detailed design of surface encounters rests heavily on the capabilities and design of the proposed spacecraft. A series of interesting case studies for the Hayabusa spacecraft and its successful touchdowns on the asteroid Itokawa surface were summarized in [81, 201, 64, 84]. In the following we take a more dynamics and mechanics oriented approach and describe the surface environment due to gravitational and rotational effects, describe measures of surface stability, and discuss some simple characterizations of the ballistics of bodies from the surface of an asteroid. This chapter does not probe the interesting question of how surface rovers should be designed and operated in order to achieve specific goals (see [85]).

10.1 Surface Specification

The physical surface of the asteroid to be specified can be defined by a constraint function $S(\boldsymbol{r}) = 0$, where $S > 0$ corresponds to a positive altitude above the surface and $S < 0$ would be below the surface. On the surface of the asteroid the gradient of S is defined as the surface normal, $S_{\boldsymbol{r}} = \hat{\boldsymbol{n}}$, and is itself a function of position on the asteroid surface. If the asteroid shape is specified as a polyhedron, then the constraint consists of a series of flat plates with constant normal vectors that become discontinuous at edges and vertices. The second partial of S is identically zero if on a flat facet, and is not defined as the surface point changes from facet to facet. If the surface is specified as a smooth function, such as the popular Gaussian random shape models [113], the second partial can be non-zero and would describe the local topography. An actual asteroid surface will, of course, consist of a highly discontinuous function as asteroid surfaces are often covered with boulders,

D.J. Scheeres, *Orbital Motion in Strongly Perturbed Environments*, Springer Praxis Books,
DOI 10.1007/978-3-642-03256-1_10, © Springer-Verlag Berlin Heidelberg 2012

cobbles and blocks [108]. Thus, the use of a flat or smooth surface is an obvious simplification that describes macroscopic properties of the surface while ignoring its small-scale features.

10.2 Surface Geopotential

A quantity of interest for the surface environment is the effective potential, defined by the combined gravitational potential and rotational potential terms. For a uniformly rotating body this is just the Jacobi integral discussed earlier, evaluated at rest on the asteroid surface. This provides a relative measure of the available energy that can be converted to kinetic energy (and hence easily dissipated) based on the location of a particle in the asteroid frame. The effective potential energy function of an asteroid is:

$$V(\mathbf{r}) = \frac{1}{2}\mathbf{r} \cdot \tilde{\boldsymbol{\omega}} \cdot \tilde{\boldsymbol{\omega}} \cdot \mathbf{r} + U(\mathbf{r}) \tag{10.1}$$

where $\boldsymbol{\omega}$ is the angular velocity vector of the body and can be time-varying. Uniform rotation is generally assumed in this chapter (see [173, 165] for a discussion of surface slopes on a non-uniformly rotating body). Using this, the dynamical height of the asteroid surface can be computed, a relative measure from a locally defined average gravity [181].

A more direct interpretation of the surface geopotential is to link it directly to possible dynamical motion across the asteroid surface. Recall from the Jacobi integral that the quantity $J = \frac{1}{2}v^2 - V(\boldsymbol{r})$ is conserved during ballistic flight, if the body is uniformly rotating (which is assumed for the current discussion). Here v is the speed of the particle in the body-fixed frame. Thus, given two points at rest on the surface of the asteroid, $\boldsymbol{r}_o$ and $\boldsymbol{r}$, the ideal kinetic energy required for an object to move from one location to the other can be compared. Assume that the body is at rest at point $\boldsymbol{r}_o$ and that the two points will be linked by a ballistic trajectory (independent of the existence of a ballistic trajectory between these two points). Then the two Jacobi integral values can be equated to find:

$$\frac{1}{2}v^2 - V(\boldsymbol{r}) = -V(\boldsymbol{r}_o) \tag{10.2}$$

and solved for the speed required to link the two points on the surface

$$v = \sqrt{2\left(V(\boldsymbol{r}) - V(\boldsymbol{r}_o)\right)} \tag{10.3}$$

To be defined, the quantity $V(\boldsymbol{r}) - V(\boldsymbol{r}_o) \geq 0$, and thus the point $\boldsymbol{r}_o$, must lie at a lower point in the geopotential. If this speed is to be defined everywhere on the asteroid surface relative to the rest point $\boldsymbol{r}_o$, the rest point must be taken at the lowest value of the geopotential on the surface. Then the geopotential across the asteroid surface can be defined in terms of the kinetic energy or speed required to boost a particle from the rest point to another point on the asteroid. Conversely,

this relationship can also be interpreted as the excess speed or kinetic energy that a particle would have if it falls to the lowest point on the asteroid without dissipation. This speed, or more fundamentally the value of $V(\boldsymbol{r})$, can be used to characterize the geopotential across the asteroid surface. This has advantages over the use of the "dynamical height" in that it provides direct information about the required specific kinetic energy needed to transfer from one region to the other, or the energy relieved due to downslope motion.

Figure 10.1 shows the surface geopotential mapped across the surface of asteroids Eros and Itokawa. These two asteroids are chosen as example bodies as they are the most complete models of asteroids in existence, with detailed shape models, spin states, and masses.

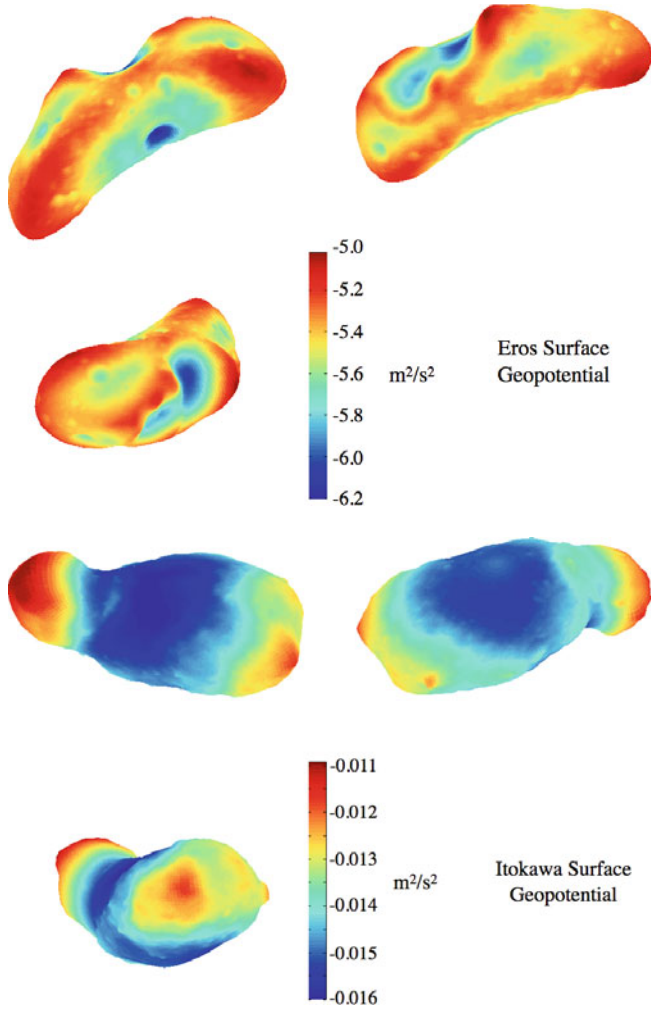


Fig. 10.1 Geopotential mapped over the surfaces of Eros and Itokawa. Note the different scales for the different bodies.

10.3 Surface Forces and Environment

The mechanical forces felt on the surface can be reduced to surface normal and transverse frictional forces acting on a particle. These forces can be reduced to surface acceleration fields and are defined by the asteroid's gravity field, surface, and rotation state. Recent interest in other forces acting on the asteroid surface have been revived by the unexpected morphology of the Eros and Itokawa surfaces. Indeed, electromagnetic forces operating on small dust particles are being considered to explain some of the dust ponding seen on Eros [93, 141]. In addition to these are occasional impulsive forces that may jolt asteroid regolith, due to impacts of other asteroids on the asteroid surface [54, 182, 7]. Finally, it has been speculated that van der Waals cohesion between grains on the surface of an asteroid may play an important role in the distribution and mechanics of surface material [169]. The current discussion only focuses on the gravitational and rotational accelerations, however.

The total acceleration that a particle feels when at rest on the surface of a rotating asteroid is:

$$\mathbf{N} = \dot{\omega} \times \mathbf{r} + \omega \times \omega \times \mathbf{r} + \frac{\partial U}{\partial \mathbf{r}} \tag{10.4}$$

If the asteroid is uniformly rotating then the surface acceleration is equal to the gradient of the geopotential, $\boldsymbol{N} = V_{\boldsymbol{r}}$. If the local surface normal is $\hat{\mathbf{n}}$, then the surface acceleration is split into a normal and tangential component:

$$N_{\perp} = \hat{\mathbf{n}} \cdot \mathbf{N} \tag{10.5}$$

$$\mathbf{N}_{\parallel} = \left[\bar{\bar{U}} - \hat{\mathbf{n}}\hat{\mathbf{n}}\right] \cdot \mathbf{N} \tag{10.6}$$

and the local slope of the system is defined as:

$$\phi = \arctan\left[\frac{|\mathbf{N}_{\parallel}|}{N_{\perp}}\right] \tag{10.7}$$

The surface slope can be related to the static coefficient of friction on the surface, μ, as $\mu \geq \tan\phi$.

The description of the surface gravitational field and surface normal vectors is simple if a constant density polyhedral model of the asteroid is used, as there is an exact concordance between the shape and the gravity field. The spherical harmonic expansions of the gravity field cannot be used for this analysis, due to their divergence at the surface of an irregularly shaped body. Thus, if a body has significant density heterogeneity, the polyhedral gravitational potential must be modified to account for these, a challenging problem in general that has been considered at an elementary level in [170].

Figures 10.2 and 10.3 show the surface normal acceleration and slope mapped across the surface of asteroids Eros and Itokawa.

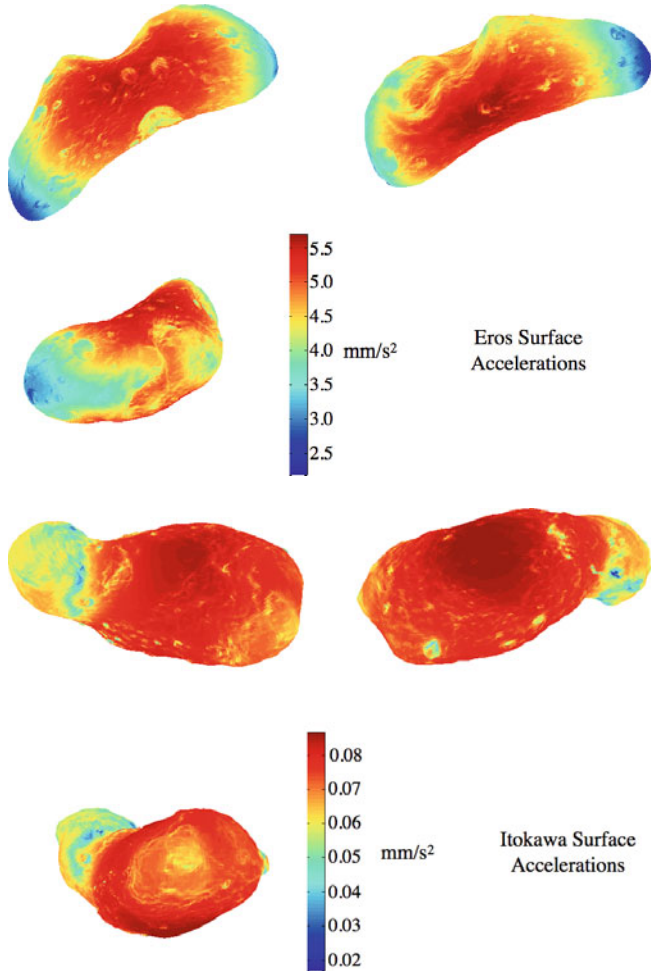


Fig. 10.2 Surface normal acceleration mapped over the surfaces of Eros and Itokawa. Note the different scales for the different bodies.

The rotational dynamics of the body can play a significant role in modifying the surface environment, and may change the stability and structure of motion on the surface. For bodies in complex rotation the slopes and surface forces are time periodic, and could potentially add sufficient "shaking" (physically realized by slowly varying slopes at each point on the surface) to cause the surface to relax, reducing the potential energy stored in local slopes. Any asteroid subject to non-uniform rotation following a large impact or planetary flyby will have these time periodic forces acting on its surface, which could play a role in smoothing a surface after an impact. This is distinguished from seismic shaking, where the asteroid feels small seismic events due to the flux of impactors striking the asteroid. While the magnitude of shaking expected from impactors should be larger than from non-

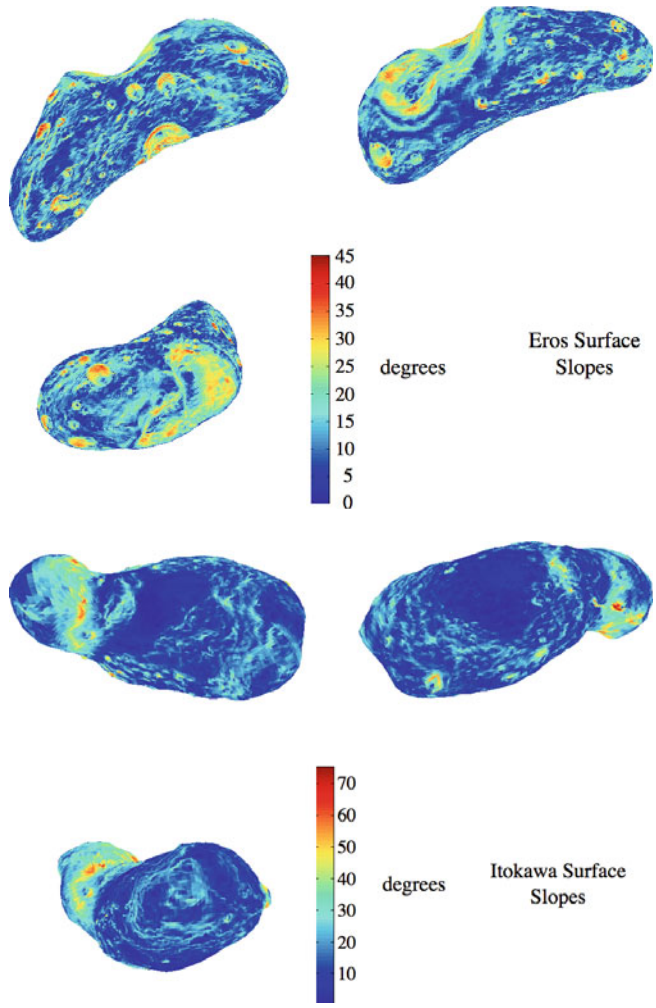


Fig. 10.3 Slopes mapped over the surfaces of Eros and Itokawa. Note the different scales for the different bodies.

uniform rotation, the non-uniform rotation will act continuously on the asteroid over the time it takes for it to relax into uniform rotation.

Finally, it should be noted that if an asteroid were to actually describe a figure of equilibrium (see [189, 39]), then the surface slope would be identically zero over the entire body. In fact, deviations of surface slope from zero indicate deviations from a figure of equilibrium. Slope distributions of asteroids have been measured from spacecraft observations and from radar imaging of asteroids. The surface slopes of most bodies measured in this way have uniformly low slopes (i.e., less than 30° on average) that, at the least, could be indicators of a relaxed surface.

10.4 Stationary and Stable Motion

Of interest is a methodology for describing the stability of a point on the surface, specifically if it is trapped at a certain region or if it is at a local "high" and is subject to downslope motion. While these concepts are trivially tied to surface topography on the surface of the Earth, or planetary bodies in general, on asteroids there can be drastic changes of the concept of where "down" is as a function of the shape, gravity field and spin state [57].

Motivated by our ability to identify a "lowest point" on the surface of an asteroid in terms of the kinetic energy required to transfer between two points, we will search for local stationary points of the geopotential across the surface of the body. To formalize this, define a geopotential cost function which can be evaluated for local stationarity subject to lying on the asteroid surface, or $S(\boldsymbol{r}) = 0$,

$$\Lambda = V(\boldsymbol{r}) + \lambda S(\boldsymbol{r}) \tag{10.8}$$

where λ is a Lagrange multiplier. The necessary conditions for a given point to be a locally stationary point on the surface is that $\delta\Lambda = 0$ or $(V_{\boldsymbol{r}} + \lambda S_{\boldsymbol{r}}) \cdot \delta\boldsymbol{r} + S\delta\lambda = 0$ for all admissible variations of $\delta\boldsymbol{r}$ and $\delta\lambda$. Noting that $S_{\boldsymbol{r}} = \hat{\boldsymbol{n}}$ and $S \equiv 0$ this condition can be reduced to

$$V_{\boldsymbol{r}} + \lambda\hat{\boldsymbol{n}} = 0 \tag{10.9}$$

For a uniformly rotating asteroid this just specifies that the total acceleration acting on the surface particle is parallel to the surface normal, that λ is the surface normal acceleration, and thus that the slope equals zero. Specifically, $\lambda = -\hat{\boldsymbol{n}} \cdot V_{\boldsymbol{r}}$. Thus, zero slope regions are candidates for local gravitational highs and lows.

To evaluate whether a candidate point with zero slope is a gravitational high or low, consider the second variation of Λ

$$\delta^2\Lambda = \delta\boldsymbol{r} \cdot [\tilde{\boldsymbol{\omega}} \cdot \tilde{\boldsymbol{\omega}} + U_{\boldsymbol{rr}} + \lambda S_{\boldsymbol{rr}}] \cdot \delta\boldsymbol{r} \tag{10.10}$$

where the other terms are identically zero given the surface constraint. If on a faceted shape model $S_{\boldsymbol{rr}} \equiv 0$ except at edges and vertices where it is undefined. Stability of a surface point then implies that $\delta^2\Lambda > 0$, or can only increase for any given admissible $\delta\boldsymbol{r}$. If $\delta^2\Lambda < 0$ then the point is at a local maximum of the geopotential, and has excess energy for moving in the downslope direction. If the second variation is semi-definite the local point is a saddle. For a smooth surface the stability condition can be reduced to the definiteness of the matrix

$$\tilde{\boldsymbol{\omega}} \cdot \tilde{\boldsymbol{\omega}} + U_{\boldsymbol{rr}} - \hat{\boldsymbol{n}} \cdot V_{\boldsymbol{r}} S_{\boldsymbol{rr}} > 0 \tag{10.11}$$

and for a faceted model the term $S_{\boldsymbol{rr}}$ disappears.

The stability of a surface point need not be constant as a function of an object's spin rate. In [57] the stable surface points of a uniformly rotating ellipsoid were computed and shown to depend on the rotation rate of the ellipsoid. Thus, if a body has a changing spin rate over time the geopotential highs and lows can

switch, which may have important consequences for the subsequent migration of material. This also implies that direct inspection of the topography of a small body may not indicate the geopotential lows on the body.

An alternate way of determining the stability of a surface region is to plot the directions of the local acceleration vectors tangent to the surface, which in general point in the direction of downslope motion. If these vectors point towards a zero slope region, then this will be a local geopotential low and material may accumulate in this region. If, instead, it points away from a local zero slope region than that is an unstable surface region with a local maximum value of the geopotential. The advantage of this approach is that it allows one figure to show both the local slopes and the local gepotential lows and highs. Figure 10.4 presents such an image for 1999 KW4 Alpha. Here it can be clearly seen that the flow of loose material will be from the polar regions towards the equatorial region. As blue signifies near-zero slopes, there exists a stable resting area around the equator, while the polar regions are unstable. This slope pattern is due to the high spin rate of the body. If the spin rate were sufficiently lower, then the slope arrows would be reversed and the polar regions would be the stable region again.

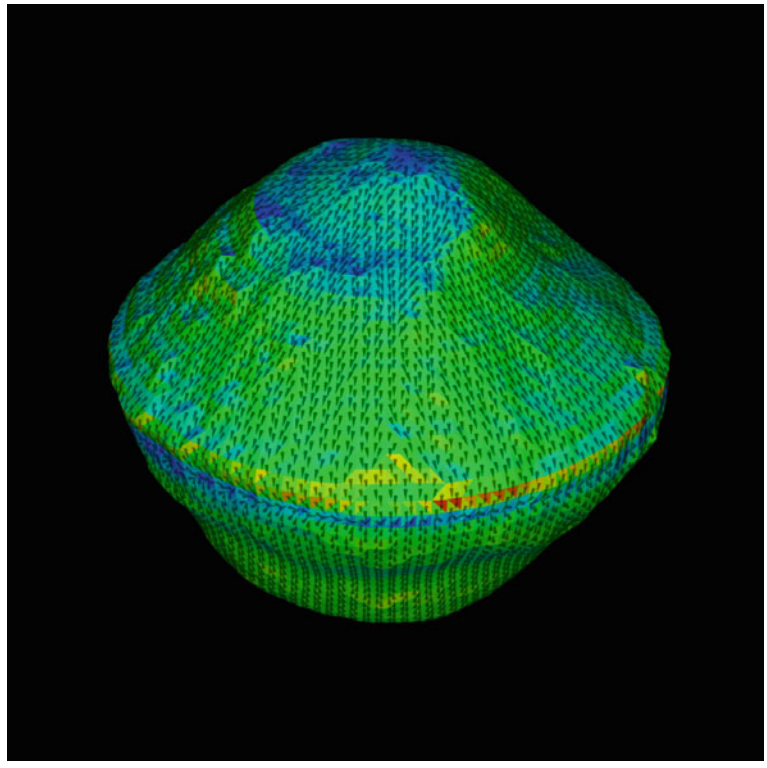


Fig. 10.4 Surface slopes and slope arrows plotted over the surface of 1999 KW4 Alpha. Blue regions are near zero-slope, green is at ∼30° slope and red is at ∼75° slope. Credit: DIAL/JPL-Caltech.

10.5 Roche Lobe and the Guaranteed Return Speed

On the surface of a uniformly rotating asteroid, the geopotential energy can also be related to the minimum amount of energy a particle requires before it can escape from the asteroid (the guaranteed re-impact speed in [174]). Specifically, the value of $V(\mathbf{r})$ at the synchronous orbits with minimum value of the Jacobi constant (C^*) defines the zero-velocity surface that surrounds and encloses the asteroid in three-dimensional space. The effective potential energy evaluated at this synchronous orbit defines the minimum energy that a particle must have before it becomes possible to escape from the surface of the asteroid, i.e., a particle with Jacobi constant greater than this value could, theoretically, escape from the asteroid following a purely ballistic trajectory. If a particle has an energy less than this, and is within the zero-velocity curve, then it is impossible for it to leave the vicinity of the asteroid. This surface has also been referred to as the Roche lobe, and was studied in the particular case of Phobos [32], and more recently has been computed for Eros [107]. Phobos was found to nearly "fill" this minimum energy surface, meaning that particles on its surface were prone to escape that body when given sufficient speeds. Conversely, Eros lies entirely within this energy surface, although 56% of that asteroid's surface lies within 1 km of this energy surface, the closest point lying only 90 meters from the energy surface. Figure 10.5 shows the computed Eros Roche lobe projected into the Eros equatorial plane. Note, it is not necessary for the asteroid surface to lie completely within the Roche Lobe.

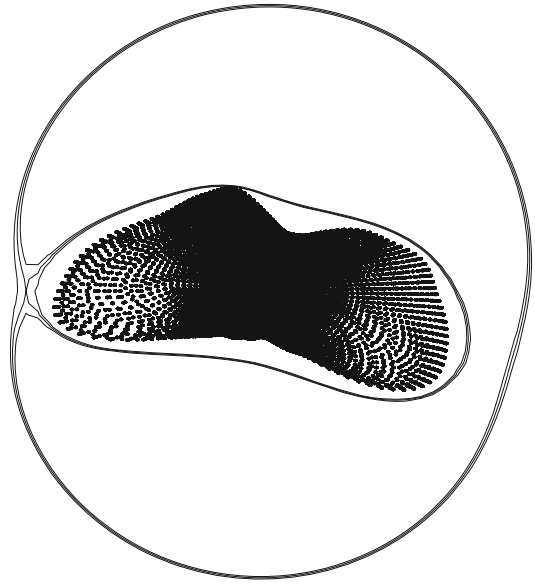


Fig. 10.5 Roche lobe of asteroid Eros, shown projected onto its equatorial plane.

The guaranteed return speed is then computed across the surface of the asteroid as

$$v_{ret} = \sqrt{2\left(V(\boldsymbol{r}) - C^*\right)} \tag{10.12}$$

and is defined to be zero when the energy of the surface exceeds C^*. This concept is important when considering the motion of surface rovers, as induced speeds relative to the surface of greater than this amount may lead to the ejection and escape of a rover from the asteroid surface while speeds less than this ensure capture. Figure 10.6 shows the surface return speed mapped across the surface of asteroids Eros and Itokawa.

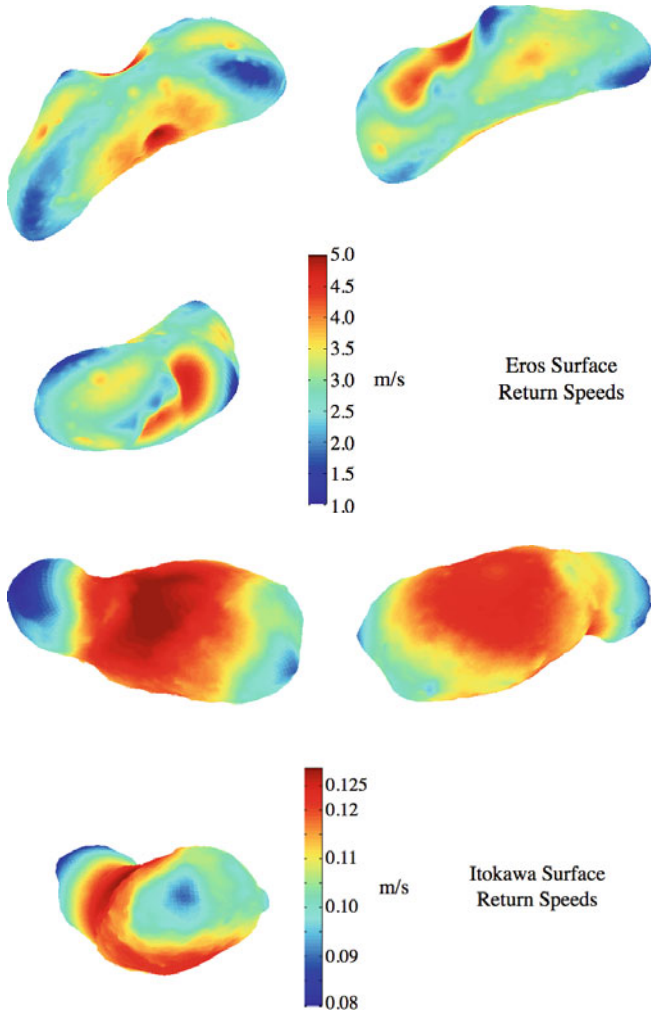


Fig. 10.6 Surface return speed mapped over the surfaces of Eros and Itokawa. Note the different scales for the different bodies.

10.5.1 Surface Escape Speeds

The guaranteed return speed can be thought of as a necessary condition for escape from the surface, i.e., a particle must have speed greater than this value if it is to escape. It is not a sufficient condition, however, as having a speed in excess of the return speed does not guarantee that the particle will escape. Such a sufficient condition for escape cannot be established as rigorously as outlined above, but conditions can be derived that yield reasonable estimates of this outcome.

The approach to this question is more constructive. First consider the inertial speed of a particle resting on an asteroid surface, which will be $\boldsymbol{v}_I = \boldsymbol{\omega} \times \boldsymbol{r}$ (see Fig. 10.7). Depending on the surface topography in relation to the spin state, this velocity can point away from the asteroid surface if the point is on the leading edge of the asteroid, or may point into the asteroid if it is on a trailing edge. Due to this, the necessary additional speed necessary to impart escape speed to a particle may vary drastically over a small body's surface.

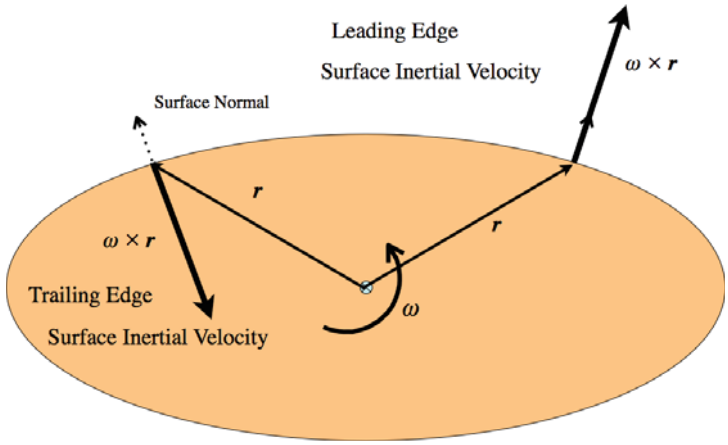


Fig. 10.7 Surface inertial speeds vary as a function of location on a rotating body.

To progress to the next step, define the escape speed from a point on the asteroid surface as the additional speed normal to the asteroid surface that would provide the particle sufficient speed to have a positive Keplerian energy when far from the asteroid. This criterion has been tested in earlier papers [174, 173] and found to be a reliable indicator, sufficient to develop surface maps to follow trends across an asteroid's surface. This definition results in a velocity vector in inertial space of $\boldsymbol{v}_I = v_{esc}\hat{\boldsymbol{n}} + \boldsymbol{\omega} \times \boldsymbol{r}$. Our definition requires the magnitude of $\boldsymbol{v}_I$ to be greater than or equal to $\sqrt{2U_{max}}$ where $U_{max} = \max\left[U(\boldsymbol{r}), \mu/r\right]$, which is a conservative result. Solving for the surface escape speed v_{esc} yields [173]

$$v_{esc}(\boldsymbol{r}) = -\hat{\boldsymbol{n}} \cdot (\boldsymbol{\omega} \times \boldsymbol{r}) + \sqrt{\left[\hat{\boldsymbol{n}} \cdot (\boldsymbol{\omega} \times \boldsymbol{r})\right]^2 + 2U_{max}(\boldsymbol{r}) - (\boldsymbol{\omega} \times \boldsymbol{r})^2} \tag{10.13}$$

This is not a well-defined quantity when the body has locally non-convex regions where a speed normal to the surface would result in a re-impact with a different region of the body. Despite this, the result will still indicate the level of speed necessary in general to generate energies consistent with escape from a body. Figure 10.8 shows the escape speed mapped across the surface of asteroids Eros and Itokawa.

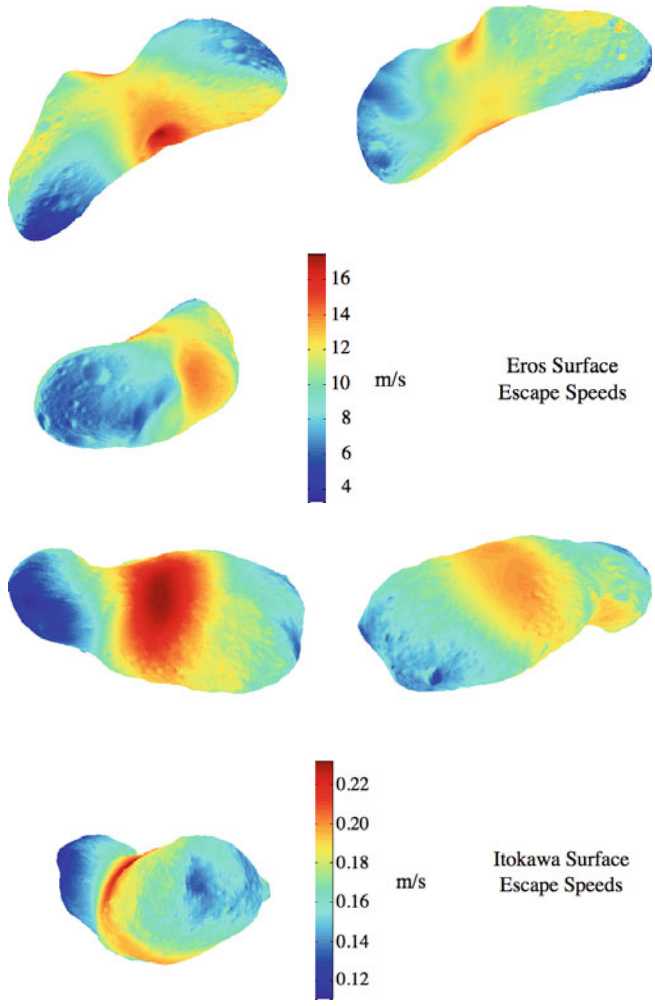


Fig. 10.8 Escape speed mapped over the surfaces of Eros and Itokawa. Note the different scales for the different bodies.

11. Controlled Hovering Motion at an Asteroid

We end our sequence of chapters on dynamics dominated by the small-body gravity in isolation by focusing on the application of control maneuvers to enact hovering about a small body. This technique was applied during the Hayabusa mission to asteroid Itokawa, where the spacecraft maintained its location on the Sun-side of the asteroid by performing occasional thrusting maneuvers. For a small body this is a feasible set of operations as the necessary propellant to maintain a given distance can be quite small. This chapter draws from a number of different analyses of such motion [157, 151, 16, 17, 22] and only presents an introduction to the problem, leaving details to these other papers and for future work.

11.1 Motivation

Performing scientific explorations of small bodies such as asteroids and comets can be simplified in many cases by abandoning an "orbital" approach in favor of a "hovering" approach in which the spacecraft thrusts continuously, near-continuously, or sporadically to null out gravitational and rotational accelerations, either fixing its position in a body-fixed frame or in a heliocentric orbit-fixed frame. Such approaches to small-body exploration make it possible to obtain high-resolution measurements, and even samples, from multiple sites over the body surface without having to make complicated transitions from orbital to body-fixed trajectories between each near-surface observation period. The Hayabusa mission to Itokawa implemented a mixture of these approaches, maintaining an orbit-fixed location when far from the asteroid and transitioning into a body-fixed frame during its descents to the asteroid surface.

A variant of such hovering is to perform a sequence of slow hyperbolic flybys of the target body, with V_∞ on the order of centimeters to tens of centimeters per second, ending each flyby with a small maneuver to turn the trajectory around for the next flyby. Such an approach can also enable the gravity field of the body to be determined [180], which may be more difficult if the spacecraft is in a fixed location and frequently thrusting (which was the case for the Hayabusa spacecraft).

D.J. Scheeres, *Orbital Motion in Strongly Perturbed Environments*, Springer Praxis Books,
DOI 10.1007/978-3-642-03256-1_11, © Springer-Verlag Berlin Heidelberg 2012

The main issue that arises with a hovering approach is that the artificial equilibrium point that the spacecraft creates in the relevant frame of reference is almost always unstable, and thus must always be implemented with the navigation of this hovering solution in mind. In the following we review some simple results that can be developed for a hovering analysis, and discuss procedures for the stabilization of such hovering trajectories.

There are two general approaches to controlled motion: near-inertial hovering and body-fixed hovering. In near-inertial hovering the spacecraft is stationed at a fixed location relative to the asteroid in the Sun–asteroid frame, the asteroid rotating beneath the spacecraft. This is the hovering mode implemented by Hayabusa during most of its mission. In body-fixed hovering the spacecraft is stationed at a fixed location relative to the rotating asteroid, implying that the spacecraft is rotating with the asteroid in inertial space. This mode is essential for sampling a small-body surface, as at some point the spacecraft must control its motion in the asteroid-fixed frame. Again, the Hayabusa mission implemented body-fixed hovering in a dynamic sense during its sampling runs close to the surface. Both of these ideas, and their generalizations, are discussed in more detail below.

For the implementation of either approach to be feasible some minimal level of sensing capability is needed on-board the spacecraft. First is the ability to directly sense altitude, either using a laser altimeter or by the efficient processing of stereoscopic optical measurements. This measurement type forms the backbone of an automatic control system to maintain altitude and position relative to an asteroid. In addition to this, it is ideal for the vehicle to have the ability to sense its location relative to the asteroid surface. This can be implemented by optical sensors or scanning lasers, both of these technologies are in different stages of development. These are not the only types of measurements available or useful, but they are the most essential. The efficient measurement of altitude allows for the implementation of automatic control algorithms that stabilize the spacecraft hovering position, while measurements of body-relative location allows for an expanded capability for the control and motion of the spacecraft. For the latter case, the spacecraft must be able to correlate measured features with a global topography map in order to locate its current location. For some specific applications it may only be necessary to measure and detect lateral motion in addition to vertical motion, but for the most general applications the ability to detemine its global location on the asteroid is necessary. This implies that a global map of the asteroid surface has been created at some earlier point, ideally using the same instruments to be used for the relative navigation. The development and implementation of such sensor systems is a technology that is currently being developed, and is available for use in the future.

In addition to the above sensing and estimation capability, the spacecraft will also require precise 6-DOF control capability. This implies a full set of thrusters for executing arbitrary control moves, perhaps augmented by momentum wheels for fine attitude control. It may be feasible to use more restrictive thruster configurations for the control of the spacecraft, although these would have to be carefully designed for specific implementation approaches. Finally, some, but not all, of these active control approaches imply that the vehicle may be out of Sun-light for con-

siderable periods of time. Thus, such power considerations should be factored into the development and design of space vehicles for these advanced approaches.

11.2 Near-Inertial Hovering

In this approach the spacecraft fixes its location relative to the body in the rotating body–Sun frame, creating an artificial libration point in this frame. A useful way to think about this approach is to first consider the Sun–asteroid libration point. A spacecraft placed in this location will, ideally, remain fixed in its position. If, however, the spacecraft adds a constant thrust acceleration away from the asteroid, it would have to move its location closer to the asteroid in order for the forces to balance again. If a sufficiently large acceleration is added, it could conceivably hold its position relatively close to the asteroid. If close to the asteroid, it would have to supply an acceleration of $\sim\mu/r^2$ to "hover" at a radius of r from the attracting asteroid.

From this relationship an approximate measure of the propellant cost to hover can be developed. Assuming a spherical asteroid of density ρ and radius R the control acceleration needed to null out the gravity at a given distance r is

$$u \sim \frac{4\pi}{3}\mathcal{G}\rho R\left(\frac{R}{r}\right)^2 \tag{11.1}$$

$$\sim 24\rho R\left(\frac{R}{r}\right)^2 \text{ m / s / day} \tag{11.2}$$

where the units have been converted to meters per second of ΔV per day and R is in km. Thus, assuming a density of 2 g/cm^3, hovering at the surface of a 0.1-km body costs 4.8 m/s/day while hovering at the surface of a 10-km body costs 480 m/s/day – the former value being reasonable and the latter being unreasonable for extended hovering.

Considering the more general case, it is possible to specify the necessary control acceleration to maintain position at an arbitrary location in the asteroid–Sun rotating frame. Due to the relatively slow motion of the body about the Sun (on the order of degrees per day at fastest), this position can be considered to be nearly inertial over relatively short periods of time. Thus this discussion will assume that the spacecraft wishes to fix itself in an inertially oriented frame. Assuming that the spacecraft is attracted to the rotating asteroid and also has some thrusting capability, represented as a control acceleration $\boldsymbol{u}$, the full equations of motion are

$$\ddot{\boldsymbol{r}} = \frac{\partial U}{\partial \boldsymbol{r}} + \boldsymbol{u} \tag{11.3}$$

where $\boldsymbol{r}$ is the position of the spacecraft relative to the asteroid and is specified relative to an inertially fixed frame, $U(\boldsymbol{C}^T\cdot\boldsymbol{r}) = \frac{\mu}{r} + R(\boldsymbol{C}^T\cdot\boldsymbol{r})$ is the gravitational field, $\boldsymbol{C}$ is the transformation dyad that takes the asteroid-fixed frame into the inertial frame, and R is the perturbing, non-spherical component of the gravity field. Sim-

ply put, open-loop inertial hovering is enacted by choosing the control acceleration to balance against the local attraction at the desired hovering location $\boldsymbol{r}^*$.

$$\boldsymbol{u} = -\frac{\partial U}{\partial \boldsymbol{r}^*} \tag{11.4}$$

For hovering above a uniformly rotating non-spherical asteroid the control law is time-varying and periodic with the rotation of the asteroid. If the central body is tumbling then the control acceleration will not repeat in general, as a tumbling body will generally not present the same orientation in inertial space twice.

Although this hovering specification is simplistic, as it is not necessary to continuously throttle the propulsion system to hover, it is a useful point from which to consider the stability of this approach. Specifically, consider the effect of a small error in positioning of the spacecraft – this could either be due to navigation uncertainties in the precise placement of the spacecraft or could be due to an error in the propulsion system, meaning that the current control law is actually the hovering law for a neighboring spacecraft location. Linearizing about $\boldsymbol{r} = \boldsymbol{r}^* + \delta\boldsymbol{r}$ and assuming that the control is exactly canceling the gravitational attraction at $\boldsymbol{r}^*$ yields

$$\delta\ddot{\boldsymbol{r}} = \left.\frac{\partial^2 U}{\partial \boldsymbol{r}^2}\right|_* \cdot \delta\boldsymbol{r} \tag{11.5}$$

$$\left.\frac{\partial^2 U}{\partial \boldsymbol{r}^2}\right|_* = -\frac{\mu}{r^{*3}}\left[\boldsymbol{U} - 3\hat{\boldsymbol{r}^*}\hat{\boldsymbol{r}^*}\right] + \left.\frac{\partial^2 R}{\partial \boldsymbol{r}^2}\right|_* \tag{11.6}$$

The stability of this system for a uniformly rotating body can be evaluated using Floquet's theorem, and for complex rotation would require the computation of Lyapunov characteristic exponents. These more detailed issues are considered in [17], while for current purposes It is more instructive to ignore the gravitational perturbation component R and instead just focus on the point-mass attraction term. This is often justified as inertial hovering will generally occur relatively far from the central body where the gravitational perturbations are small and could be ignored.

When considering the point-mass term the stability is tractable as the dynamics matrix $U_{\boldsymbol{rr}}$ is time-invariant, and its eigenvalues and eigenvectors can be easily identified. There exists one eigenvector along the $\hat{\boldsymbol{r}}^*$ vector, and dotting this unit vector with the dynamics matrix yields

$$\frac{2\mu}{r^3}\hat{\boldsymbol{r}^*} = -\frac{\mu}{r^{*3}}\left[\boldsymbol{U} - 3\hat{\boldsymbol{r}^*}\hat{\boldsymbol{r}^*}\right] \cdot \hat{\boldsymbol{r}^*} \tag{11.7}$$

Note that the eigenvalue is $\pm\sqrt{(2\mu/r^3)}$, and thus has an exponentially stable and unstable component. The other eigenvector is found by dotting the dynamics matrix with any vector perpendicular to $\hat{\boldsymbol{r}^*}$, denoted as $\hat{\boldsymbol{r}^*_\perp}$, yielding

$$-\frac{\mu}{r^3}\hat{\boldsymbol{r}^*_\perp} = -\frac{\mu}{r^{*3}}\left[\boldsymbol{U} - 3\hat{\boldsymbol{r}^*}\hat{\boldsymbol{r}^*}\right] \cdot \hat{\boldsymbol{r}^*_\perp} \tag{11.8}$$

In this case the eigenvalues along this direction are imaginary and equal to $\pm i\sqrt{(\mu/r^3)}$. Thus the 3-DOF system can be reduced to a 1-DOF system along $\hat{\boldsymbol{r}}^*$ which has a stable and unstable component and a 2-DOF system consisting of two uncoupled harmonic oscillators perpendicular to the hovering position vector. The oscillation period of the motion in the perpendicular direction has period equal to the orbit period of a circular orbit with semi-major axis equal to the hovering radius, $2\pi {r^*}^{2/3}/\sqrt{\mu}$. The characteristic time of the hyperbolic manifolds equals ${r^*}^{2/3}/\sqrt{2\mu}$. In the radial direction this leads to motion of the form

$$\delta r = \alpha_+ e^{\sqrt{\frac{2\mu}{r^{*3}}}t} + \alpha_- e^{-\sqrt{\frac{2\mu}{r^{*3}}}t} \tag{11.9}$$

Thus any small disturbance in this direction will grow exponentially with a characteristic time that is only a function of body density and r^*/R. As this is the dominant term in the stability analysis it implies that hovering must always be carried out with control along the radial, hovering direction. Thus, practical implementation requires the addition of a closed-loop feedback control (potentially through the ground) that senses the altitude or distance deviation of the spacecraft from its ideal hovering point. The necessary control loop to stabilize this motion is actually quite simple, and can be implemented in an automatic way using minimal spacecraft resources [16]. There are limits to this approach, however. A spacecraft cannot inertially hover within the maximum radius of the asteroid at its hovering latitude, due to obvious physical constraints. Additionally, as the radius of hovering becomes closer to the body, the automatic control approach described here can become unstable due to the neglected higher-order terms, potentially leading to difficulties in implementation.

It is not necessary, however, to force the spacecraft to be fixed precisely at one location. A generalization of this idea places the spacecraft in an elliptic or hyperbolic orbit relative to the asteroid, and has its velocity vector "reflected" whenever it gets within a certain distance to the asteroid (for an elliptic orbit) or gets a certain distance away from the asteroid (for a hyperbolic orbit), forcing the spacecraft to travel back on, or close to, its original path but in the opposite direction. Figure 11.1 shows this graphically, while Fig. 11.2 presents the trajectory of a spacecraft hovering relative to an asteroid in an asteroid–Sun rotating frame (this hovering point is away from the sub-solar point and below the equatorial plane, as seen in the graphics, and includes the attraction of the asteroid, the attraction of the Sun, and the solar radiation pressure). Despite the simplicity of this control law – providing a fixed ΔV every time a distance to the asteroid is crossed – it can maintain the spacecraft within a fixed control volume over an arbitrarily long timespan. This approach can be thought of as hovering with a relatively large dead-band control about the nominal hovering point, and requires essentially the same control and sensing capability on-board the spacecraft. This is essentially the approach used by the Hayabusa spacecraft during its mission at Itokawa [82]. In this approach the time between control maneuvers can be made arbitrarily long by increasing the size of the dead-band box away from the asteroid. The Lyapunov stability of this approach to hovering was investigated in detail in [17] and found to be robust. The use of hyperbolic-flyby only interactions with a small body has

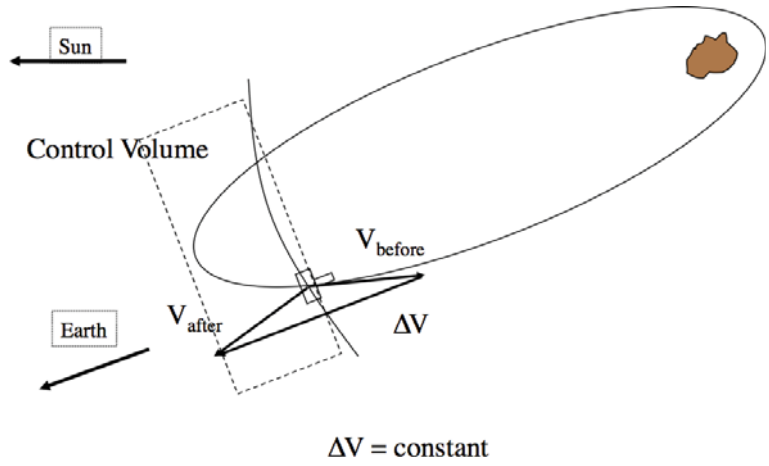


Fig. 11.1 Simple inertial hovering scheme, where a constant ΔV vector is applied (relative to the Sun–asteroid frame) whenever a threshold distance is crossed.

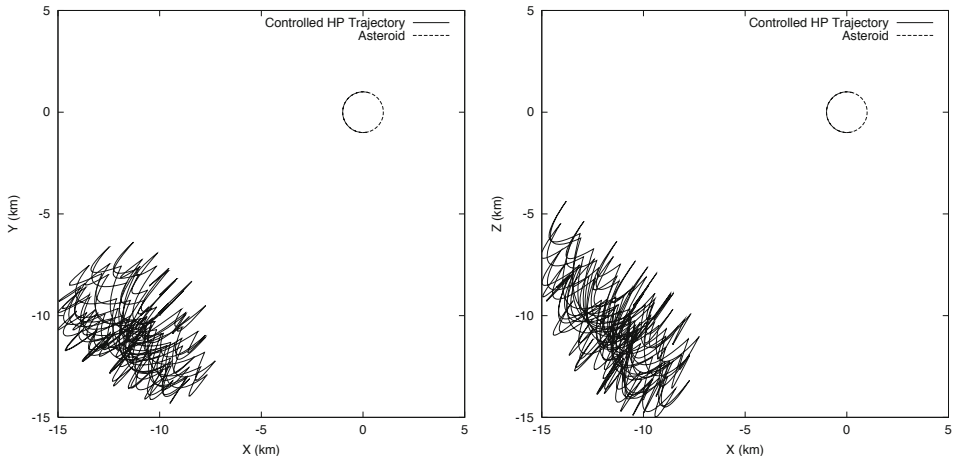


Fig. 11.2 Dynamics of a fixed ΔV hovering scheme for a spacecraft situated off the sub-solar point and below the asteroid orbital plane. Projection of the trajectory into the orbital X–Y plane (left) and into the X–Z plane (right). The trajectory remains bounded and nonlinearly stable for arbitrary periods of time.

been proposed for a number of small-body missions, although it has yet to be fully implemented [180].

Inertial hovering, or its above variations, has several attractive attributes which may make it a mainstay approach of future exploration. There are also a number of drawbacks and limitations, however. On the positive side, this approach can be applied to any small body, and the cost of inertial hovering can theoretically always be driven to zero by hovering at a high enough altitude (not accounting for the statistical control to stabilize the hovering point). However, the position where

hovering is feasible may be far from the body, and may not afford the optimal viewing geometry. For example, if the NEAR-Shoemaker spacecraft had taken a hovering approach to its mission to Eros and implemented inertial hovering at a distance of 50 km from the asteroid (which was the nominal orbit radius for most of the mission), it would have required over 15 m/s per day to maintain this position, or for its prime 9-month mission would have required a total ΔV on the order of 4 km/s. Contrasted with the actual fuel usage (on the order of a few tens of m/s), hovering was clearly not a reasonable approach for that body. Thus, to gain high resolution scientific measurements this approach is largely limited to smaller bodies with their associated smaller hovering cost.

A related drawback pertains to the ability of a spacecraft to accurately measure the mass and gravity field of the asteroid, and hence to compile an accurate understanding of the body's internal structure. When using such a controlled hovering mode, errors in the spacecraft thrusters and solar radiation pressure parameters will compete with the signature of the asteroid gravity field acting on the spacecraft. While it might be possible to extract some averaged results on the total mass of the asteroid, these results would be corrupted by many different uncertainties that will not be uniquely separated from the gravity signature. Additionally, higher-order gravity fields will be nearly impossible to extract. Again, this was the situation for the Hayabusa mission as it was only able to estimate a relatively imprecise estimate of the asteroid's mass and was not able to determine any of its higher-order gravity field coefficients [1, 200, 166]. This is a serious limitation, as it deprives the mission of essential scientific data and may make it difficult to subsequently transition to a body-fixed hovering exploration of the asteroid. Recent research has investigated the ability of a spacecraft to obtain mass and higher-order gravity field coefficients for a small body by carrying out several slow hyperbolic flybys of the body [180]. Using this approach it is feasible to gain precise information on the gravity field up to second degree and order (assuming that the shape has also been estimated) and to detect up to fourth degree and order gravity coefficients for an asteroid the size of Itokawa ($\sim$100 meters across).

11.3 Body-fixed Hovering

The counterpart to inertial or near-inertial hovering, with its range of implementation options, is hovering in the small-body-fixed frame. In this approach the spacecraft fixes its position relative to the rotating body. A natural way to visualize this approach is to imagine using a "jet-pack" to levitate off of the surface of a rotating body, such as the Earth or an asteroid. Since the gravitational attraction is relatively weak at asteroids, it is possible to implement such hovering trajectories for extended periods of time (hours) with total costs that can be relatively modest, on the order of meters per second. This approach to controlling motion in close proximity to small bodies has been analyzed in detail [157, 151] and a detailed simulation of this approach has been developed for analysis of hovering over arbi-

trary models of asteroids [16, 17]. During its final descent, the Hayabusa spacecraft also enacted a form of body-fixed hovering in its sampling mission [81].

The implementation of this approach has similarities to inertial hovering, but now everything must be done relative to the asteroid-fixed frame, which generally has a rotation period on the order of hours to days at most. Thus, the spacecraft acceleration must accommodate both the gravitational and centrifugal accelerations, although there are locations where the hovering cost is zero (at the synchronous orbits). Additionally, the spacecraft must reorient its attitude in inertial space in order to maintain the same orientation relative to the asteroid surface. This body-fixed hovering approach also suffers from the same basic instability noted for the inertial hovering case, although there are regions where this approach yields completely stabilized motion [157]. A similar control strategy, using altimetry to maintain a fixed altitude, can stabilize a hovering point so long as it is located within the synchronous radius of the body. This result holds approximately true over the entire body and places an altitude "ceiling" on hovering for a simple control law to be able to stabilize its location.

There are a number of drawbacks related to body-fixed hovering as well. First, it is essential that a fairly accurate model of the asteroid spin, topography and gravity field be available. The gravity field must be defined down to the surface of the body as well, something which is not always easy to do (as described previously). Thus, body-fixed hovering should be preceded by a period of characterization at a relatively high level of accuracy. In the future, it may be possible to dispense with this requirement, but that would only be after the basic technology and approach has been proven. Most important, however, is that a body-fixed hovering vehicle could experience periods of solar occultation, making the presence of batteries or non-solar power generation essential for long-term operations at the surface of a body. Additional operations issues also exist, such as communications, attitude determination, and the mechanical interface of its control thrust plumes with the surface.

In the following the stability of hovering in a uniformly rotating asteroid-fixed frame is considered. An analytical consideration of hovering is only investigated above a point mass in the following, although some examples of the more complex cases of hovering over non-spherical mass distributions are presented. The equations of motion for a point mass in a uniformly rotating frame can be restated from Chapter 7 with the addition of a control acceleration

$$\ddot{\boldsymbol{q}} + 2\tilde{\boldsymbol{\omega}} \cdot \dot{\boldsymbol{q}} + \tilde{\boldsymbol{\omega}} \cdot \tilde{\boldsymbol{\omega}} \cdot \boldsymbol{q} = -\frac{\mu}{q^3}\boldsymbol{q} + \boldsymbol{u} \tag{11.10}$$

To generalize this to a uniformly rotating arbitrary mass distribution replace the gravitational acceleration $-\frac{\mu}{q^3}\boldsymbol{q}$ with $U_{\boldsymbol{q}}$ and have the angular velocity vector be aligned with the maximum moment of inertia and equal in magnitude to its rotation rate. Hovering at a specified point in the asteroid-fixed frame, $\boldsymbol{q}^*$, requires a control acceleration of $\boldsymbol{u}^* = \frac{\mu}{q^{*3}}\boldsymbol{q}^* + \tilde{\boldsymbol{\omega}} \cdot \tilde{\boldsymbol{\omega}} \cdot \boldsymbol{q}^*$. For simplicity, assuming that hovering occurs

in the equatorial plane, this simplifies to

$$\boldsymbol{u}^* = \left(\frac{\mu}{q^{*3}} - \omega^2\right)\boldsymbol{q}^*$$

Thus the nominal hovering cost can go to zero if the hovering radius equals the synchronous orbit radius. For a non-spherical mass distribution this is equivalent to placing the spacecraft at one of the libration points. These are often unstable, and have a more complex instability structure than found for the inertially hovering equilibrium points, and thus the stability properties of body-fixed hovering orbits are expected to be different than the simple structure found for inertial hovering points.

For this point-mass hovering analysis denote the hovering location vector as $\boldsymbol{q}^* = q^*\left[\cos\theta\hat{\boldsymbol{x}} + \sin\theta\hat{\boldsymbol{z}}\right]$, where θ is the latitude of hovering, the asteroid rotates about the $\hat{\boldsymbol{z}}$ axis, and q^* is the distance from the body center of mass. Assuming $\boldsymbol{q} = \boldsymbol{q}^* + \delta\boldsymbol{q}$ one can linearize the equations of motion about this point to find

$$\delta\ddot{\boldsymbol{q}} + 2\omega\tilde{\hat{\boldsymbol{z}}}\cdot\delta\dot{\boldsymbol{q}} = \left\{-\omega^2\tilde{\hat{\boldsymbol{z}}}\cdot\tilde{\hat{\boldsymbol{z}}} - \frac{\mu}{q^{*3}}\left[\boldsymbol{U} - 3\hat{\boldsymbol{q}}^*\hat{\boldsymbol{q}}^*\right]\right\}\cdot\delta\boldsymbol{q} \tag{11.11}$$

Writing out the characteristic matrix for this case, with λ as the test eigenvalue, yields

$$\begin{bmatrix} \lambda^2 - \omega^2 + \omega_q^2\left(1 - 3\cos^2\theta\right) & -2\omega\lambda & -3\omega_q^2\sin\theta\cos\theta \\ 2\omega\lambda & \lambda^2 - \omega^2 + \omega_q^2 & 0 \\ -3\omega_q^2\sin\theta\cos\theta & 0 & \lambda^2 + \omega_q^2\left(1 - 3\sin^2\theta\right) \end{bmatrix} \tag{11.12}$$

where $\omega_q^2 = \mu/q^3$ and the explicit hovering position has been substituted. The characteristic equation for the system is a fully coupled polynomial of degree 3 in λ^2. In Fig. 11.3 the different stability results are given as a function of scaled radius and θ, with different colors indicating different stability types. There are four types of stable and unstable motion found for hovering in the body-fixed frame. For all latitudes of hovering with $\omega_q > \omega$ (i.e., hovering within the resonance radius of the asteroid) there is unstable motion with characteristics similar to inertial hovering: one stable and unstable manifold and two center manifolds. For this case the stable and unstable manifolds no longer line up with the radius vector in general. Transitioning to hovering locations above the resonance radius there is a larger set of hovering stability types. The red regions correspond to complex stable and unstable manifolds and one center manifold, the white regions are fully stable, and the dark regions have two pairs of stable and unstable manifolds and one center manifold.

The inclusion of non-spherical mass distributions can change the distribution of stability types markedly. This is analyzed in [17] where the hovering stability type is mapped over the surface of a number of different asteroids. Some specific examples of this are shown in Figs. 11.4 to 11.6, which plot the hovering fuel cost, the stability type, and the characteristic time for a spacecraft hovering above the Eros surface at an altitude of 100 meters. A further generalization of body-fixed

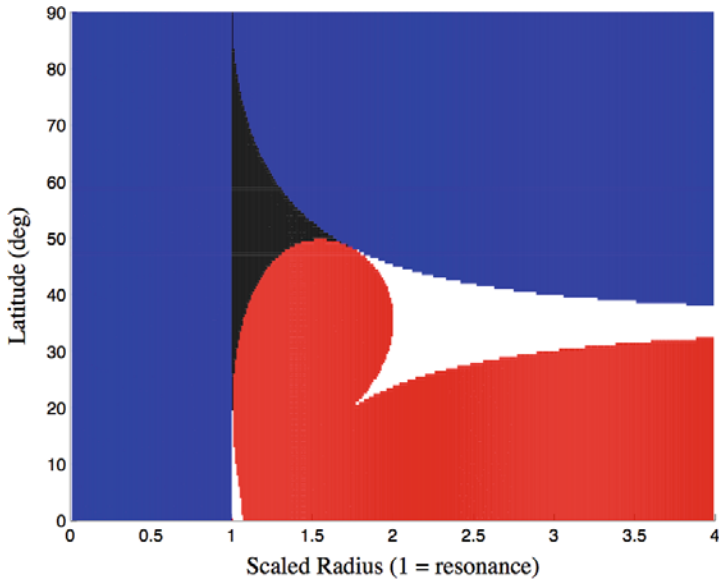


Fig. 11.3 Body-fixed hovering stability as a function of radius and latitude above an ideal spinning sphere. White is stable (oscillatory), blue regions have a single stable and unstable manifold, and two center manifolds red regions have complex stable and unstable manifolds and one center manifold, and the dark regions have two pairs of stable and unstable manifolds and one center manifold.

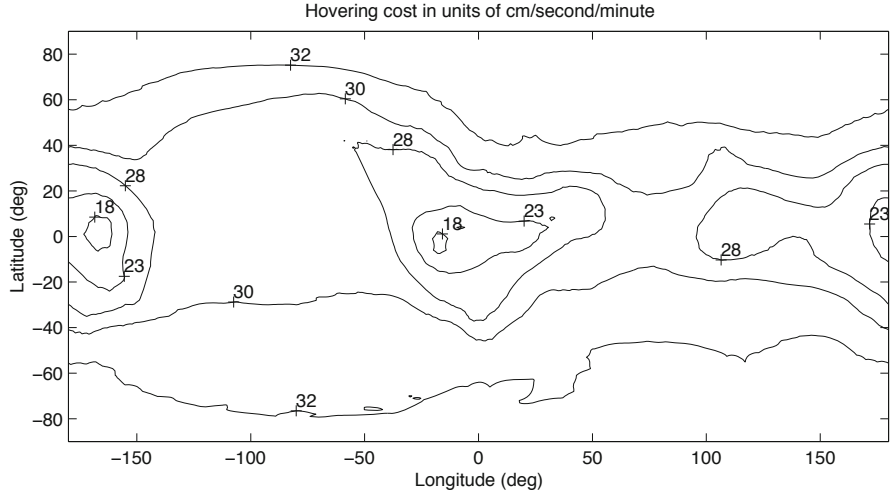


Fig. 11.4 The average cost of hovering at a 100-meter altitude above asteroid Eros, as a function of latitude and longitude of the hovering point.

hovering is to model the translational motion relative to the asteroid surface. This is analyzed in [151] where it is shown that motion against the rotation of the asteroid is in general more stable than motion in the same sense over the asteroid surface.

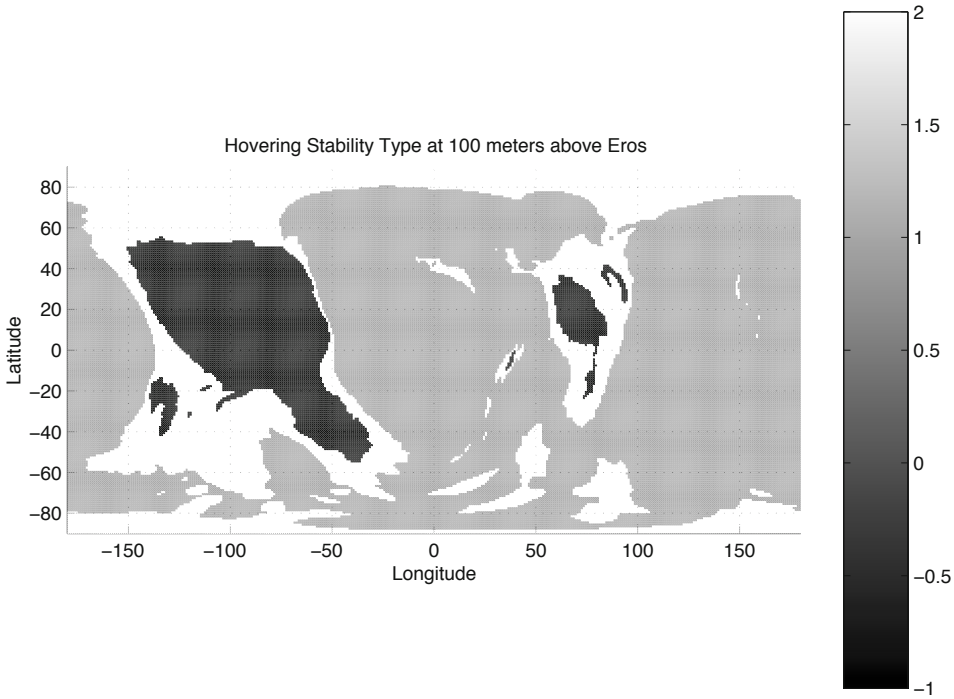


Fig. 11.5 The stability of hovering points at a 100 meter altitude above Eros. An index of 0 is stable (oscillatory), 1 has a single unstable / stable manifold, 2 has two pairs of unstable / stable manifolds, and -1 has a set of complex unstable / stable manifolds. Note that there are no stable hovering locations, which is expected as the 100 meter altitude is uniformly within the resonance radius of the body.

The implementation of body-fixed hovering can also be accomplished using a simple dead-band controller acting on the altitude of the spacecraft, in conjunction with a single thrust direction properly aligned relative to the asteroid gravity field, technologies that are clearly feasible to implement now. If, in addition to maintaining a single altitude at a specific location, it is desired for the spacecraft to translate in the asteroid-fixed frame, moving from one site to another, and to descend and ascend from the surface, additional sensing and control technology will have to be used. First, the spacecraft must maintain its attitude in the body-fixed frame – it should be noted that the spacecraft will not do so naturally, as its attitude will remain fixed in inertial space and will want to spin in the asteroid-fixed space. Second, for it to perform translational motions will require that the spacecraft have the capability to locate itself relative to the surface and perform some higher-level control to null out transverse oscillations about the hovering point.

The development of this surface-relative motion capability is perhaps the most advanced concept tendered here. This idea also solves the problem of rover locomotion over an asteroid surface, as instead of relying on natural trajectories induced by mechanical "jumpers" it provides controlled motion from one location to an-

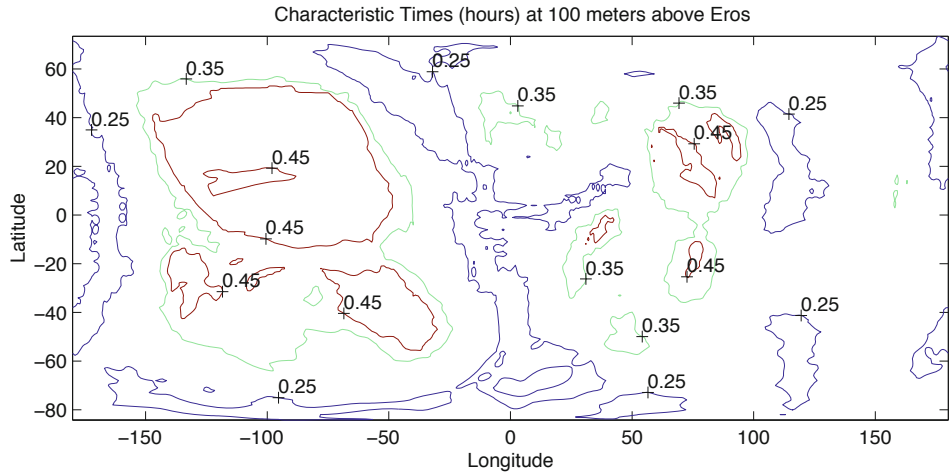


Fig. 11.6 The minimum characteristic time of the unstable 100 meter altitude hovering locations mapped in contours above the Eros surface.

other. There are a number of interesting peculiarities associated with such surface relative motion, such as the fact that there is a preferred direction of motion about an asteroid in this mode [151]. Motion in the same direction as asteroid rotation can actually destabilize the dynamics of the spacecraft control, while translational motion in the opposite direction will tend to stabilize the control system. Other than this observation, which can be easily proven, there is little known about the stability and control of surface-relative motion at small bodies, making it an essential topic for future research.

12. Solar Radiation Pressure: Exact Analysis

Moving beyond gravity-only dynamics about small bodies, we first consider the combined effect of solar radiation and solar tide perturbations on a spacecraft orbiting about an asteroid or comet. In this section we assume the central body can be modeled as a sphere, and neglect gravitational perturbations. This situation models orbital dynamics when far from asteroids or comets where the dominant perturbation will be from solar effects. We shall also see that for large enough bodies, such as Eros, solar radiation pressure only plays a minimal role. For missions to bodies whose sizes are on the order of a few kilometers or less, however, solar radiation pressure is the principal concern for orbital stability. It is interesting to note that the solar tide is generally negligible when compared to solar radiation pressure, yet it is included in this discussion for completeness. From a direct analysis of the equations of motion and their equilibrium points specific limits can be derived on orbit semi-major axis for when solar radiation can strip a spacecraft out of orbit. The analysis given here is based on some earlier work by Dankowicz [27] and the analysis given in [171].

12.1 Models

In the current discussion a focus on comets is introduced in addition to asteroids, although not on any one particular body. Listed in Table 12.1 are a number of different spacecraft and their target bodies. We identify Rosetta with the comet Wirtanen, which was its original target, in addition to listing its properties relative to its current target Comet Churyumov-Gerasimenko. The ST-4 spacecraft to comet Tempel-1 was a proposed mission to that body and is included as it had a very low mass-to-area ratio leading to it being a highly perturbed body, and thus is of interest. The asteroid and comet mass properties are all approximate. This analysis fully accounts for the potentially elliptic motion of the small body about the Sun, which is appropriate given that the solar radiation pressure will vary strongly as a function of distance from the Sun. Estimates of spacecraft values of mass-to-area ratio are highly speculative, and are estimated based on very simple

D.J. Scheeres, *Orbital Motion in Strongly Perturbed Environments*, Springer Praxis Books,
DOI 10.1007/978-3-642-03256-1_12, © Springer-Verlag Berlin Heidelberg 2012

descriptions of the spacecraft total mass and total surface area that would potentially be facing the Sun. As these values were obtained from the gray literature in many cases, they are likely to be off by a factor of a few.

Table 12.1 Approximate values of the spacecraft mass-to-area ratio (B), the small-body gravitational parameter (μ), the small body's orbital eccentricity (e_S) and orbit parameter (p_S), the non-dimensional parameter $\epsilon = (\mu/\mu_{Sun})^{1/3}$, and the non-dimensional parameter β, defined later in this chapter.

Spacecraft Body Name	B (kg/m^2)	μ (km^3/s^2)	e_S (—)	p_S (AU)	ϵ (—)	β (—)
Rosetta Wirtanen	20	3×10^{-7}	0.658	1.752	1.31×10^{-6}	28.5
Rosetta Churyumov-Gerasimenko	20	7×10^{-7}	0.64	2.047	1.74×10^{-6}	21.5
ST-4 Tempel 1	7.5	5×10^{-6}	0.5175	2.285	3.34×10^{-6}	29.8
NEAR-Shoemaker Eros	40	5×10^{-4}	0.223	1.385	1.55×10^{-5}	1.20
Hayabusa Itokawa	33	5×10^{-8}	0.136	1.25	7.2×10^{-7}	31.41

12.2 Equations of Motion

The solar derived forces acting on the spacecraft are the solar radiation pressure (SRP) force that acts on the craft and the tidal force from the sun, necessitating a model for the asteroid's motion about the sun. The small body's orbit about the sun is modeled using 2-body theory, which is adequate for describing the body dynamics over time periods of relevance for a spacecraft mission. All the relevant models and equations were previously derived for the general case in Chapters 2 and 4. This chapter focuses on a very simple model for motion about a body dominated by SRP, as it provides useful insight into the fundamental aspects of this problem and also enables the derivation of a simple condition for when a spacecraft will be bound to the small body.

12.2.1 Point Mass with No Rotation

The simplest system to analyze is that of a non-rotating, constant acceleration acting on a spacecraft orbiting about a point mass. This problem is integrable if motion is constrained to a plane containing the acceleration vector, as it is a limiting case of the fixed two-center problem with one of the centers being moved to infinity [13]. This problem was also analyzed extensively by Dankowicz [27, 28, 29, 30] using advanced methods of dynamical systems theory. Dankowicz's analysis considered

both the unperturbed problem and various perturbations of it. From the initial analysis of Dankowicz some very useful results can be extracted for spacecraft orbit design, which are based on the ideal conservation of the angular momentum about the constant force direction (i.e., the line of action of the solar radiation pressure). These results are considered first.

The relevant potential in this case is $U = \mu/r + a_{SRP}\hat{\mathbf{d}} \cdot \mathbf{r}$ where $\hat{\mathbf{d}}$ is assumed to be stationary with respect to inertial space. The resulting equations of motion are:

$$\ddot{\mathbf{r}} = -\frac{\mu}{r^3}\mathbf{r} + a_{SRP}\hat{\mathbf{d}} \tag{12.1}$$

It can be easily shown that the total angular momentum projected onto the $\hat{\mathbf{d}}$ direction is conserved, or $H_d = \hat{\mathbf{d}} \cdot (\mathbf{r} \times \dot{\mathbf{r}})$ is a constant. The form of the equations are simplified if one shifts to a cylindrical frame with the axis of the cylinder along $\hat{\mathbf{d}}$ and the radius, ρ, and polar angle, θ, measured perpendicular to this direction. Assign the $\hat{\mathbf{x}}$-axis to the direction of the acceleration and the $\hat{\mathbf{y}}$- and $\hat{\mathbf{z}}$-axis perpendicular to this to find the simplified set of equations

$$\ddot{x} = -\frac{\mu x}{r^3} + a_{SRP} \tag{12.2}$$

$$\ddot{\rho} - \rho\dot{\theta}^2 = -\frac{\mu\rho}{r^3} \tag{12.3}$$

$$\rho\ddot{\theta} + 2\dot{\rho}\dot{\theta} = 0 \tag{12.4}$$

where $r^2 = x^2 + \rho^2$. Figure 12.1 shows the geometry of the orbit, with the radius ρ measured from the x-axis and the displacement away from the asteroid center measured along the x axis.

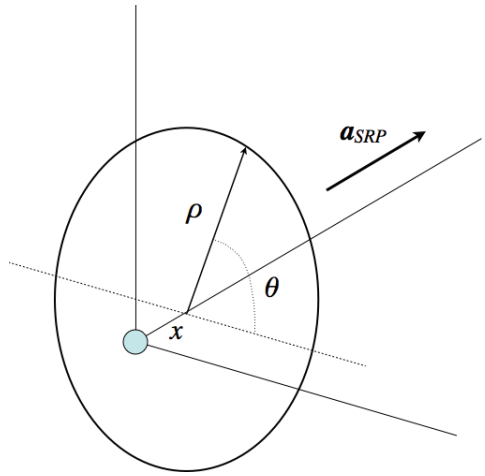


Fig. 12.1 Geometry of the orbit relative to the asteroid and the anti-Sun line (along $\boldsymbol{a}_{SRP}$).

The conserved angular momentum is now $H_d = \rho^2\dot{\theta}$ and can be immediately inferred from integrating Eq. 12.4. Eliminating $\dot{\theta}$ through this parameter yields the simplified set of equations and related potential

$$\ddot{x} = \frac{\partial U_d}{\partial x} \tag{12.5}$$

$$\ddot{\rho} = \frac{\partial U_d}{\partial \rho} \tag{12.6}$$

$$U_d = \frac{\mu}{r} - \frac{1}{2}\frac{H_d^2}{\rho^2} + a_{SRP}x \tag{12.7}$$

It is important to note that the reduced set of equations still have a Jacobi integral:

$$J = \frac{1}{2}\left(\dot{x}^2 + \dot{\rho}^2\right) - U_d \tag{12.8}$$

This is directly related to the energy of the system, as if $H_d = \rho^2\dot{\theta}$ is substituted the following results

$$J = \frac{1}{2}\left(\dot{x}^2 + \dot{\rho}^2 + \rho^2\dot{\theta}^2\right) - U \tag{12.9}$$

where $U = \mu/r + a_{SRP}x$ is the original perturbation and the velocity term $\dot{x}^2 + \dot{\rho}^2 + \rho^2\dot{\theta}^2 = v_I^2$ where v_I is the magnitude of the total inertial velocity.

12.2.2 Point Mass with Rotation

Once the motion of the asteroid about the Sun is modeled, by allowing the unit vector $\hat{\mathbf{d}}$ to rotate as a function of time, the conservation of H_d is destroyed and the dynamics of the system become much more complex. Additionally, the solar gravitational effect must be taken into account as it balances the centripetal accelerations that arise from the frame rotation. Finally, the elliptic orbit of the body about the Sun must be accounted for, as this causes large fluctuations in the magnitude of the solar radiation pressure. Incorporating our simple SRP model into a sun-fixed rotating frame and expanding the solar gravitational attraction to the lowest order yields

$$\ddot{\mathbf{r}} + 2\dot{f}\tilde{\mathbf{Z}}\cdot\dot{\mathbf{r}} + \ddot{f}\tilde{\mathbf{Z}}\cdot\mathbf{r} + \dot{f}^2\tilde{\mathbf{Z}}\cdot\tilde{\mathbf{Z}}\cdot\mathbf{r} = -\frac{\mu}{r^3}\mathbf{r} + a_{SRP}\hat{\mathbf{d}} - \frac{\mu_S}{2d^3}\left[\mathbf{r} - 3(\hat{\mathbf{d}}\cdot\mathbf{r})\mathbf{r}\right] \tag{12.10}$$

where f is the small-body heliocentric orbit true anomaly and μ_S is the Sun's gravitational parameter ($\sim 1.327 \times 10^{11}$ km^3/s^2). The equations are difficult to analyze in this form, due to the time-varying coefficients. A standard approach for such a problem is to introduce the true anomaly as an independent parameter and to scale the position vector by the time-varying distance d, defining a new position vector $\mathbf{R} = \mathbf{r}/d$. These transformations, along with the Hill approximation, were derived in detail in Chapter 4 and initially derived for the asteroid/comet problem

in [171]. Making these transformations gives us the new equations of motion

$$\mathbf{r}'' + 2\hat{\mathbf{z}} \times \mathbf{r}' + (\hat{\mathbf{z}} \cdot \mathbf{r})\,\hat{\mathbf{z}} = \frac{1}{1 + e_S \cos f} \frac{\partial U}{\partial \mathbf{r}} \tag{12.11}$$

$$U = \frac{1}{|\mathbf{r}|} + \beta \hat{\mathbf{d}} \cdot \mathbf{r} + \frac{3}{2}\left(\hat{\mathbf{d}} \cdot \mathbf{r}\right)^2 \tag{12.12}$$

The parameter β is constant and describes the relative acceleration of the SRP on the spacecraft. In terms of fundamental quantities it equals

$$\beta = \frac{(1+\rho)P_0}{B\mu_S \epsilon} \tag{12.13}$$

where $\epsilon = (\mu/\mu_S)^{1/3}$. Note that β is not necessarily a small quantity. For example, the Rosetta spacecraft at comet Wirtanen, the DS-4 spacecraft at Tempel 1, and the Hayabusa spacecraft at Itokawa will all have values of $\beta \sim 30$. In contrast the NEAR-Shoemaker spacecraft at the asteroid Eros had $\beta \sim 1$ and planetary orbiters will have $\beta \ll 1$.

It is significant to note that Eq. 12.11 only contains two parameters, the eccentricity of the orbit e_S and the normalized effect of the SRP force, β, and that the equations are time-periodic in the true anomaly f. These equations have a close affinity with the elliptic restricted three-body problem, with the change of moving the origin to the smaller primary, the addition of Hill's approximation for the effect of the larger primary, and the effect of the solar radiation pressure. Rewriting the equations in scalar form (assuming that $\hat{\mathbf{d}} = \hat{\mathbf{x}}$) yields:

$$x'' - 2y' = \frac{1}{1 + e_S \cos f}\left[-\frac{x}{r^3} + \beta + 3x\right] \tag{12.14}$$

$$y'' + 2x' = \frac{1}{1 + e_S \cos f}\left[-\frac{y}{r^3}\right] \tag{12.15}$$

$$z'' + z = \frac{1}{1 + e_S \cos f}\left[-\frac{z}{r^3}\right] \tag{12.16}$$

12.3 Analysis of the No-Rotation Case

First consider the equations of motion with no rotation, given in Eqs. 12.5 to 12.7. These equations have many interesting properties, but the one focused on here is the existence of a relative equilibrium point that corresponds to a circular orbit, offset from the center of the point mass along the direction $\hat{\mathbf{x}}$ and perpendicular to this same direction. A modified form of this orbit will also play a special role in the more general case accounting for motion of the body about the Sun, described in the next chapter. The unique aspect of this solution is that it can lose its stability at a certain value of energy, and that this agrees well with observed limits for the escape of a spacecraft due to the SRP perturbation.

The equilibrium point is simple to find, solving for $\partial U_d/\partial x = 0$ and $\partial U_d/\partial \rho = 0$:

$$x_{eq} = \frac{a_{SRP}}{\mu} r_{eq}^3 \tag{12.17}$$

$$\rho_{eq}^4 = \frac{h_d^2}{\mu} r_{eq}^3 \tag{12.18}$$

where these two equations are coupled through the relation $r^2 = x^2 + \rho^2$. The crucial parameter is the orbit radius, r, and the ratio a_{SRP}/μ. Once these are specified the value of x and ρ are fixed. By definition $x \leq r$ and thus leads to a fundamental limit on the orbit radius, $r \leq \sqrt{\mu/a_{SRP}}$. Conversely, as $\rho \leq r$ as well, a lower limit can be found as $H_d^2/\mu \leq r$, which implies a limit $H_d \leq 1/\sqrt{\mu a_{SRP}}$.

The equilibrium conditions can be simplified. Since ρ_{eq} is stationary then $\dot{\theta}_{eq}$ must also be stationary and can be used as a parameter of the equilibrium. Rewriting in terms of this yields the simplified form of the equilibrium

$$x_{eq} = \frac{a_{SRP}}{\dot{\theta}_{eq}^2} \tag{12.19}$$

$$\rho_{eq}^2 = \frac{H_d}{\dot{\theta}_{eq}} \tag{12.20}$$

$$\dot{\theta}_{eq}^2 = \frac{\mu}{r_{eq}^3} \tag{12.21}$$

The Jacobi integral at the equilibrium point can also be determined. Substituting and simplifying reduces the dependence of C on only the value of x_{eq} as:

$$J = -\frac{\mu}{2}\left(\frac{a_{SRP}}{\mu x_{eq}}\right)^{1/3} - \frac{3}{2} a_{SRP} x_{eq} \tag{12.22}$$

It is instructive to find the maximum value of J as a function of x, found by solving $J_x = 0$ for x^* and substituting back into the expression for J. Solving for these values yields

$$x^* = \frac{1}{3\sqrt{3}}\sqrt{\frac{\mu}{a_{SRP}}} \tag{12.23}$$

$$J_{max} = -\frac{2}{\sqrt{3}}\sqrt{\mu a_{SRP}} \tag{12.24}$$

These become important when considering the stability of the equilibrium points.

To study the stability of this equilibrium point form the linearized equations of motion and compute the characteristic equation. For these equations this is simple, and yields the characteristic equation

$$\lambda^4 - \lambda^2 \left(U_{\rho\rho} + U_{xx}\right) + U_{\rho\rho} U_{xx} - U_{\rho x}^2 = 0 \tag{12.25}$$

$$U_{\rho\rho} = -4\dot{\theta}_{eq}^2 \left[1 - \frac{3}{4}\left(\frac{\rho_{eq}}{r_{eq}}\right)^2\right] \tag{12.26}$$

$$U_{xx} = -\dot{\theta}_{eq}^2 \left[1 - 3\left(\frac{x_{eq}}{r_{eq}}\right)^2\right] \tag{12.27}$$

$$U_{\rho x} = 3\dot{\theta}_{eq}^2 \frac{\rho_{eq} x_{eq}}{r_{eq}^2} \tag{12.28}$$

Inserting these into the characteristic equation yields:

$$\lambda^4 + 2\dot{\theta}_{eq}^2 \lambda^2 + \dot{\theta}_{eq}^4 \left[1 - \left(\frac{3x_{eq}}{r_{eq}}\right)^2\right] = 0 \tag{12.29}$$

with the simple factorization:

$$\left(\frac{\lambda}{\dot{\theta}_{eq}}\right)^2 = \frac{3x_{eq}}{r_{eq}} - 1, -\left(\frac{3x_{eq}}{r_{eq}} + 1\right) \tag{12.30}$$

For stability the roots must be pure imaginary, i.e., λ^2 must be real and negative. One root is always of the proper sign for stability, while the other root only has the proper sign for stability when the condition

$$\frac{3x_{eq}}{r_{eq}} < 1 \tag{12.31}$$

holds, which then becomes the condition for stability of this special orbit. Note that r_{eq} can be solved for in terms of x_{eq}, yielding

$$r_{eq} = \left(\frac{x_{eq}\mu}{a_{SRP}}\right)^{1/3}$$

Substituting this into the relationship defines a limit on x_{eq} for stability of the system:

$$x_{eq} < \frac{1}{3\sqrt{3}}\sqrt{\frac{\mu}{a_{SRP}}} \tag{12.32}$$

The upper bound on x_{eq} is the value at which the Jacobi integral value takes on its maximum value. Figure 12.2 shows the zero-velocity curves of this system in the ρ–x space, with the equilibrium points plotted on them.

In [27] the orbit dynamics of the system were studied when the relative equilibria were unstable, and escape was found to be the common occurrence. From the above analysis it is clear that when $J < J_{max}$ there will be two relative equilibrium orbits, one stable with $x < x^*$ and one unstable with $x > x^*$. When $J = J_{max}$ these two equilibrium points coincide at the maximum value of J. For larger values of J the equilibrium point does not exist. This relationship allows us to derive a necessary

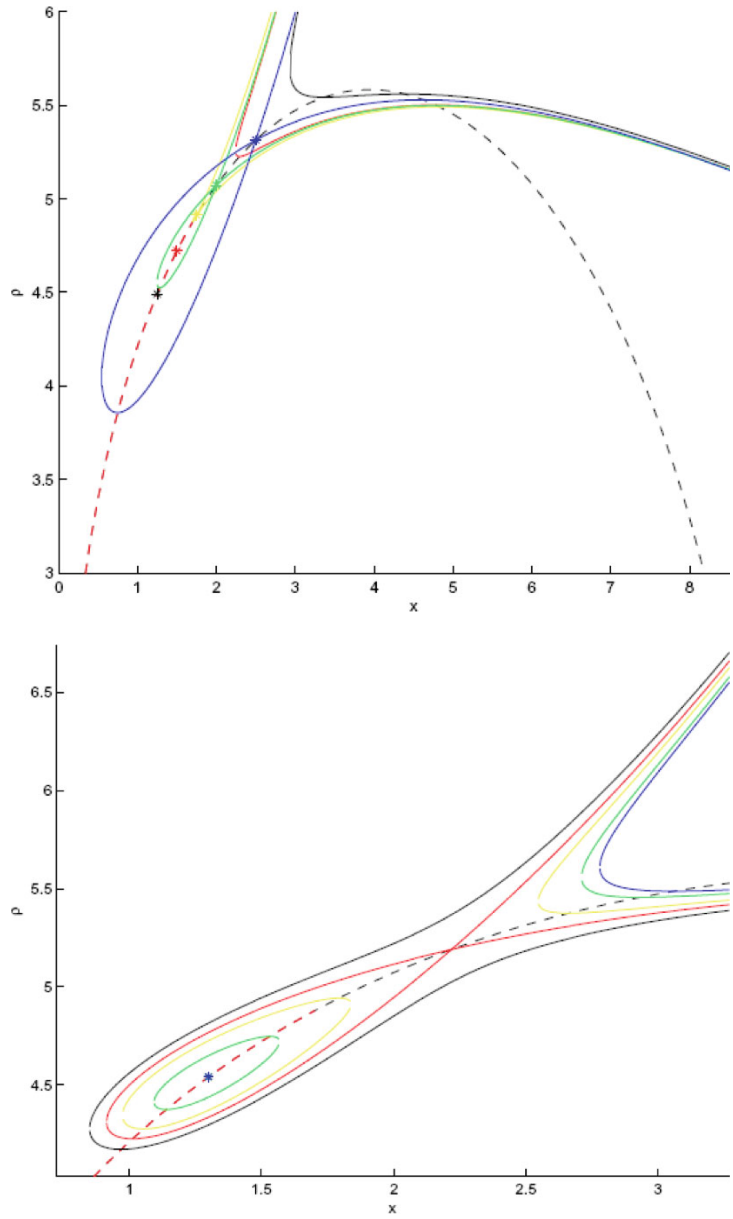


Fig. 12.2 Zero-velocity curves in the ρ–x space with circular orbits plotted as stars. The full range of the zero-velocity curves are shown in the top, with a detail shown in the bottom. Credit: [21].

condition for escape of a spacecraft from the asteroid in this simple model, or conversely a sufficient condition for stability. Simply put, if the Jacobi energy of a spacecraft is less than J_{max} and the spacecraft is in the interior region of the zero-

velocity surface, then it cannot escape. If its Jacobi energy is equal to or greater than J_{max}, then it is possible for it to escape, and in practice escape is the usual situation. Thus, given an initial value of H_d, x and ρ a limit for the reduced speed of the system can be found, $(\dot{x}^2 + \dot{\rho}^2)/2 \leq J_{max} + U_d$. This limit can be made simpler and more useful by taking advantage of the special structure of the Jacobi integral to find a limit on the inertial speed of the spacecraft

$$\frac{1}{2}v_I^2 \leq \frac{\mu}{r} + a_{SRP}x - \frac{2}{\sqrt{3}}\sqrt{\mu a_{SRP}} \tag{12.33}$$

This limit can be related to the osculating semi-major axis of an initial orbit evaluated at $x = 0$, defined by the relationship $v_I^2/2 - \mu/r = -\mu/(2a)$, to find:

$$\frac{\mu}{2a} \geq \frac{2\sqrt{\mu a_{SRP}}}{\sqrt{3}} \tag{12.34}$$

$$a \leq \frac{\sqrt{3}}{4}\sqrt{\frac{\mu}{a_{SRP}}} \tag{12.35}$$

This serves as a useful design parameter in constraining the maximum orbit size for mission design purposes.

12.4 Analysis of the Rotation Case

Now the dynamics associated with Eqs. 12.14 to 12.16 are addressed. The main focus of this discussion is on the equilibrium points of these systems and their use in deriving limits on semi-major axis for an orbiter to be bound to the small body.

12.4.1 Equilibrium Points of the System

Computing the Location of the Equilibrium Points

Equations 12.14 to 12.16 will admit an equilibrium solution that is independent of the true anomaly if $y = z = 0$ and x is chosen such that $-x/|x|^3 + \beta + 3x = 0$. This is a generalization of the well-known Lagrange equilibrium points in the Hill problem, the current result having been generalized to include the solar radiation pressure and the elliptic motion of the body. The condition for the equilibrium points becomes:

$$3x^3 + \beta x^2 - \frac{x}{|x|} = 0 \tag{12.36}$$

where the equation takes on a different form depending on whether x is greater than or less than 0. For any value of β an exact solution can be computed using Cardano's formula (for the solution of a cubic equation), or by applying a numerical procedure. It is more practical, however, to generate approximate expansions for the cases of β small and large. Parameter values of $\beta \ll 1$ apply to spacecraft that

are heavier and smaller about bodies that are larger and denser, values of $\beta \gg 1$ apply to spacecraft that are lighter and larger about bodies that are smaller and less dense.

For $\beta \ll 1$ the roots of the equation can be found by expanding the solution about the classical Hill's solution, $x^* = \pm\,(1/3)^{1/3}$, in powers of β. To the second order this procedure yields:

$$x^* \sim \pm \left(\frac{1}{3}\right)^{1/3} - \frac{\beta}{9} \pm \frac{3^{1/3}}{81}\beta^2 + \ldots \tag{12.37}$$

For $\beta \gg 1$ the roots of the equation can be found by expanding the solution in powers of $1/\sqrt{\beta}$ for $x > 0$ and in powers of $1/\beta$ for $x < 0$, with a leading term on the order β. Again, applying this procedure yields:

$$x^* \sim \begin{cases} -\frac{1}{3}\beta - \frac{9}{\beta^2} + \ldots \; x^* < 0 \\ \frac{1}{\sqrt{\beta}} - \frac{3}{2\beta^2} + \ldots \; x^* > 0 \end{cases} \tag{12.38}$$

In Fig. 12.3 comparisons of the leading order of these expansions are compared to numerical solutions of Eq. 12.36.

In the coordinate frame rotating with the body–Sun line these equilibrium points have position and velocity values relative to the body of:

$$\mathbf{R}^* = d\epsilon x^* \hat{\mathbf{d}} \tag{12.39}$$

$$\dot{\mathbf{R}}^* = \epsilon \sqrt{\frac{\mu_S}{p_S}} e_S \sin f x^* \hat{\mathbf{d}} \tag{12.40}$$

Expressing the position and velocity of the equilibrium points in an inertial frame relative to the Sun provides an even more compact result:

$$\mathbf{D}^* = [1 + \epsilon x^*]\,\mathbf{d} \tag{12.41}$$

$$\dot{\mathbf{D}}^* = [1 + \epsilon x^*]\,\dot{\mathbf{d}} \tag{12.42}$$

where $\mathbf{d}$ and $\dot{\mathbf{d}}$ denote the inertial body position and velocity vector with respect to the Sun. In all cases x^* is a solution of the equilibrium condition, Eq. 12.36. It is important to note that these equilibrium points correspond to elliptic orbits with the same eccentricity and period but with a constant fractional offset in distance from the body location.

In Fig. 12.4 the location of the Sun-side equilibrium point is shown (in metric units) as a function of true anomaly both with and without the solar radiation pressure, using parameter values from the Rosetta spacecraft and comet Wirtanen. From this figure it becomes obvious that the solar radiation pressure plays a major role in changing the location of these equilibrium points. A spacecraft with large β, such as Rosetta, that is controlled close to the libration point would have a much "safer" environment with which to view a comet than a spacecraft that was not strongly affected by the radiation pressure. Due to this geometric property we will study the stability of the Sun-side equilibrium point in greater detail, to ascertain

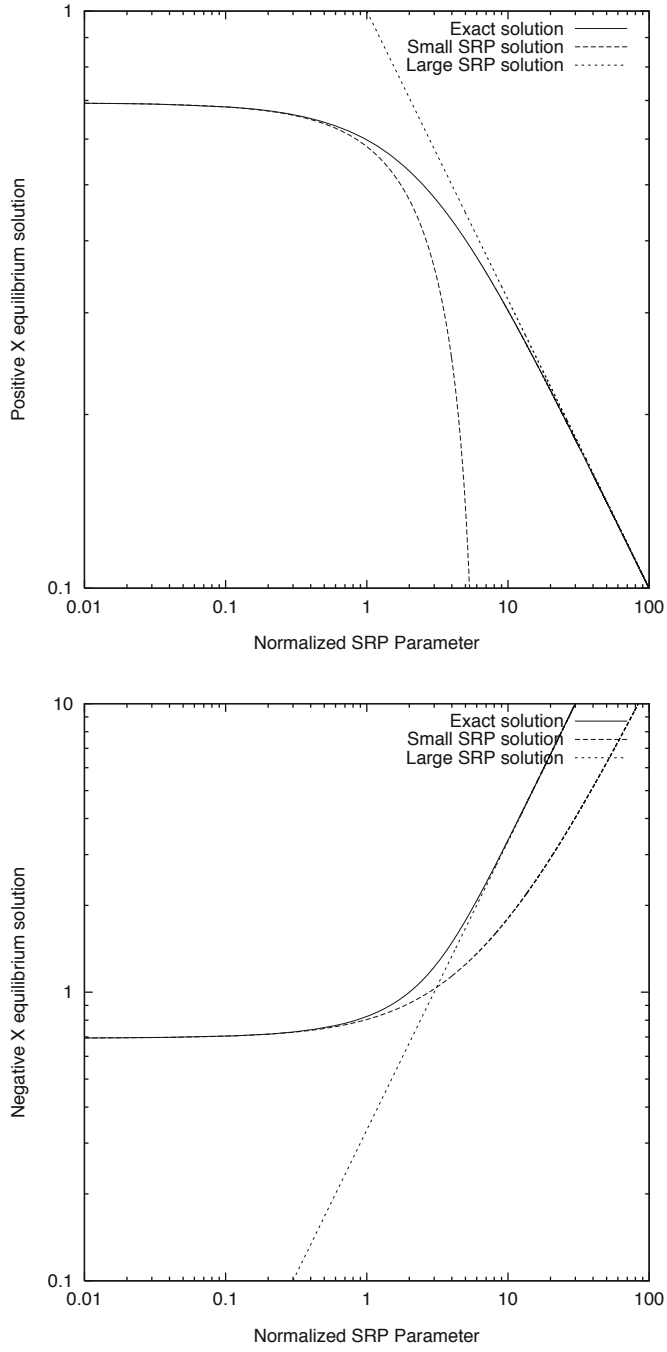


Fig. 12.3 Analytical and numerically computed equilibrium points for $x > 0$ (top) and $x < 0$ (bottom).

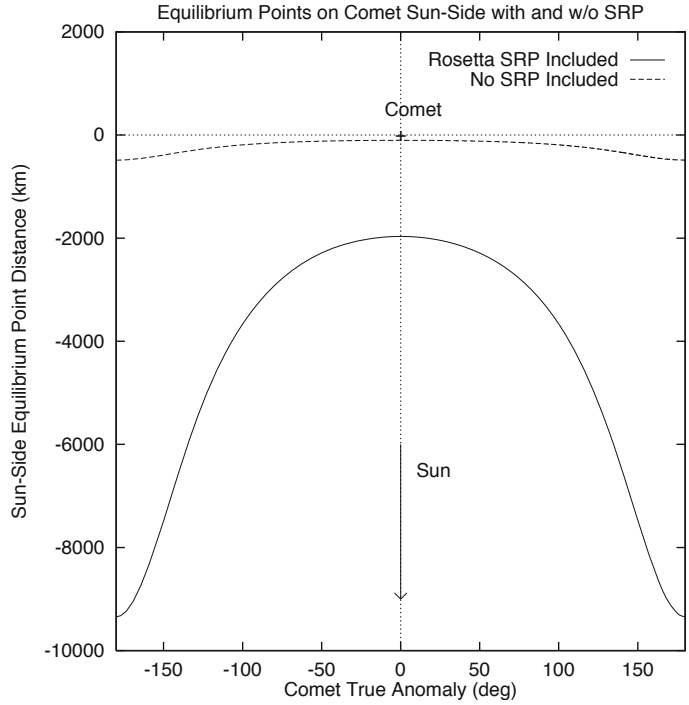


Fig. 12.4 Geometry of the equilibrium points relative to the central body as a function of true anomaly.

whether or not it is a suitable location for a spacecraft as a comet travels through perihelion.

Stability of the Equilibrium Point

In the classical Hill problem, the equilibrium points equivalent to those discussed above are all hyperbolic unstable, meaning that a particle placed near their location will depart at an exponentially increasing rate along a one-dimensional manifold, and could be attracted to the point at an exponentially decreasing rate along a one-dimensional manifold (see Chapter 16 for an explicit analysis of the Hill problem). The strength of this instability can be measured by the characteristic time for the equilibrium point. The characteristic time is a measure of how swiftly a particle will depart from the equilibrium point if slightly perturbed from it. In the limit for small perturbations, the initial displacement will increase by an order of magnitude in less than 3 characteristic times. The characteristic time can also be used as a measure of the control effort needed to maintain the spacecraft in the vicinity of the libration point. A smaller value indicates that more frequent control is needed and that the spacecraft must detect its position deviation over a shorter timespan [137, 59].

In the current application, these equilibrium points are more properly viewed as periodic orbits, and hence the computation of their stability must numerically integrate the state transition matrix about the Sun-side equilibrium point over a full period and then study its eigenvalues. We perform this computation as the true anomaly travels from -180° to 180° and the characteristic exponents of the system are found from the eigenvalues of the resulting state transition matrix. The general form of the variational equation is:

$$\Phi(f, f_0)' = A(f)\Phi(f, f_0) \tag{12.43}$$

$$\Phi(f_0, f_0) = I \tag{12.44}$$

$$A(f) = \begin{bmatrix} O & I \\ G & J \end{bmatrix} \tag{12.45}$$

where O is a 3 by 3 zero matrix, I a 3 by 3 identity matrix, and

$$J = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{12.46}$$

$$G = \frac{1}{1 + e_S \cos f} \begin{bmatrix} 3 + \frac{2}{x^{*3}} & 0 & 0 \\ 0 & -\frac{1}{x^{*3}} & 0 \\ 0 & 0 & -\frac{1}{x^{*3}} \end{bmatrix} \tag{12.47}$$

where the matrix G has been evaluated at the equilibrium point. For a given equilibrium point Φ is parameterized by the eccentricity, e_S, and the spacecraft non-dimensional SRP parameter β.

Figure 12.5 shows the characteristic time of the Sun-side equilibrium point as a function of the parameters β and e_S. For weak SRP effects (β small) the characteristic time is short, indicating that a perturbed particle will rapidly depart from the vicinity of the equilibrium point. For strong SRP effects (β large) the characteristic time is long, indicating a relatively slow departure. As the eccentricity of the body grows there is a decrease in the characteristic time.

This result is interesting for spacecraft with large values of β, as it implies that it becomes easier to control a spacecraft at the Sun-side equilibrium, a point from which it can monitor a comet through perihelion passage. Due to the larger values of β, not only is the characteristic time longer, meaning that the spacecraft has more time between applications of its control, but the geometric distance of the spacecraft from a comet is greatly increased over the distance of the tide-only equilibrium points. This is an important consideration, as the spacecraft will be less subject to comet outgassing and ejected particles.

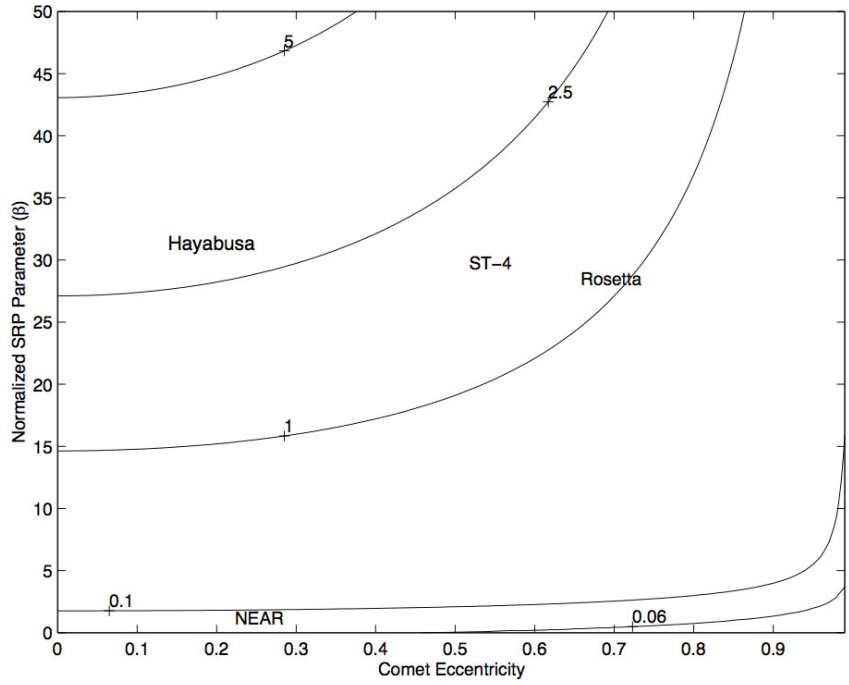


Fig. 12.5 Characteristic time of the unstable equilibrium points ($x < 0$) as a function of comet eccentricity and spacecraft β.

12.4.2 Spacecraft Capture

Using the same formalism of the Hill equations of motion, combined with results on the equilibrium point locations, it is also possible to derive a sufficient condition for a spacecraft's trajectory to be bound to a small body. This is an important consideration as it is desirable to place the spacecraft into an orbit that is definitely bound to the nucleus – accounting for both the solar radiation pressure and the solar tide. While the boundedness of any single trajectory can usually be definitively decided by numerical integration, useful analytical results can also be obtained which provide sufficient conditions for a spacecraft to be bound to the body. This condition is found by forming and studying an algebraic relation derived from the equations of motion, and using the results from the previous section on equilibrium points.

Algebraic Relation

To form the algebraic relation proceed as if forming the total Jacobi integral for Eqs. 12.14 to 12.16, multiplying each equation by x', y' and z' respectively and summing them. Treating the term $(1 + e_S \cos f)$ instantaneously as a constant the expression can be integrated to find:

$$J' = \frac{1}{2}\left(v^2 + z^2\right) - \frac{U(\mathbf{r})}{1 + e_S \cos f} \tag{12.48}$$

$$U(\mathbf{r}) = \frac{1}{r} + \beta x + \frac{3}{2}x^2 \tag{12.49}$$

where $v^2 = x'^2 + y'^2 + z'^2$. This expression is only a constant of integration in the special case when the body is in a circular orbit about the sun ($e_S = 0$). It serves a useful purpose, however, in establishing the sufficient criterion for spacecraft capture at the body. Following Marchal [97], Eq. 12.48 is re-expressed as

$$\Gamma = 2U(\mathbf{r}) - (1 + e_S \cos f)\left[v^2 + z^2\right] \tag{12.50}$$

The complete differential of Γ with respect to true anomaly yields

$$\Gamma' = e_S \sin f \left[v^2 + z^2\right] \tag{12.51}$$

showing that this quantity is only conserved for a circular orbit about the Sun, or for an equilibrium point in the orbital plane.

Whether or not Γ is conserved, if $\Gamma < \Gamma^*$, then the interior region surrounding the body is definitely separated from the outer region. Thus, for the duration when this inequality holds a spacecraft cannot escape from the body if in this interior region.

Zero-Velocity Curves in the Plane

The utility of this relation becomes obvious when considering the planar ($z = 0$) zero-velocity curves, as then the relationship becomes invariant with respect to true anomaly:

$$\Gamma = 2U(x, y, z = 0) \tag{12.52}$$

Thus, even though Γ itself is not constant, boundaries for bounded motion that are constant can still be derived – analogously to the circular Hill problem. Shown in Fig. 12.6 are zero-velocity curves for Eq. 12.52. It is significant to note that the equilibrium points serve the same function in this curve as in the standard circular problem, as they denote the points at which the region containing the nucleus connect the regions that are separated from the nucleus. The value of Γ at the equilibrium points is a constant and provides a useful criterion for whether the spacecraft is guaranteed to be bound to the nucleus or not, valid independently of the true anomaly.

For $\beta \ll 1$ the value of Γ at the equilibrium points is:

$$\Gamma^*_{\pm} \sim 3^{4/3} \pm 2\left(\frac{1}{3}\right)^{1/3} \tilde{\beta} - \frac{1}{9}\beta^2 + \ldots \tag{12.53}$$

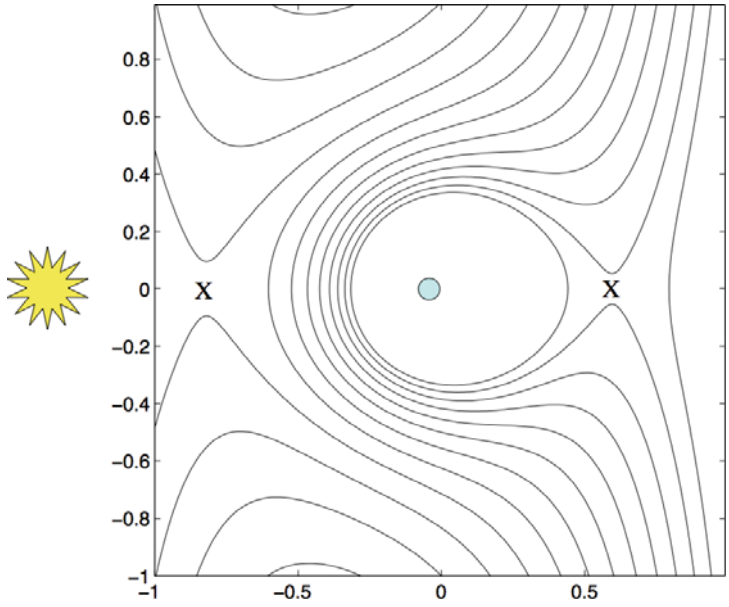


Fig. 12.6 Zero-velocity curves for the exact SRP case.

while for $\beta \gg 1$ the value is computed to be:

$$\Gamma_+^* \sim 4\sqrt{\beta} + \ldots \tag{12.54}$$

for $x^* > 0$ and

$$\Gamma_-^* \sim -\frac{1}{3}\beta^2 + \ldots \tag{12.55}$$

for $x^* < 0$. From the values given above, it is clear that for $\Gamma > \Gamma_+^*$ the region surrounding the nucleus is separated from the region containing infinity.

Sufficient Condition for Capture

Combining the previous results shows that there is a simple sufficient criterion for the capture of a spacecraft at a body: if the value of Γ for a spacecraft is always greater than Γ_+^*, and the spacecraft is inside of the zero-velocity surface, then the spacecraft is definitely bound to the body. Since the value of Γ changes over time, the satisfaction of the criterion at one time does not guarantee that it will be satisfied at some point in the future, so its use is somewhat limited. This does fit well, however, with using the criterion for a spacecraft mission as operations will generally focus on shorter time spans over which this criterion provides useful results.

To generate such results it is useful to express Γ in terms of osculating elements. To properly evaluate the orbital elements, and to make them useful for practical

computation, it is necessary to transform from the non-dimensional coordinates to a metric set of coordinates. In a coordinate frame with an inertially fixed orientation the position and velocity vector are evaluated as:

$$\mathbf{R}_I = \mathbf{R} \tag{12.56}$$

$$\dot{\mathbf{R}}_I = \dot{\mathbf{R}} + \boldsymbol{\Omega} \times \mathbf{R} \tag{12.57}$$

and the resulting expression for Γ becomes:

$$\Gamma = \frac{2d}{\epsilon^2 \mu_S}\left[-E + \Omega H \cos i + \frac{\epsilon \mu_S \beta}{d^2}\left(\hat{\mathbf{d}} \cdot \mathbf{R}_I\right) + \frac{3}{2}\frac{\Omega^2}{(1 + e_S \cos f)}\left(\hat{\mathbf{d}} \cdot \mathbf{R}_I\right)^2 - \frac{1}{2}\Omega^2 R^2 - \frac{1}{2}\Omega' \dot{\mathbf{R}}_I \cdot \mathbf{R}_I - \frac{1}{8}\Omega'^2 R^2 \right] \tag{12.58}$$

where H is the total angular momentum of the spacecraft and $\hat{\mathbf{d}}$ varies in time.

It is important to note that Γ is multiplied by the body–Sun distance, meaning that even if a spacecraft were in an orbit with constant orbital elements, the value of Γ would decrease as the body neared the Sun, and increase as the body moved away from the Sun. This is the main contributor to the mechanism that can cause a previously stable trajectory about a body to become unstable near perihelion – allowing the spacecraft to be ejected from its orbit about the body.

To apply this sufficient condition to the stability of an orbit about a body the value of Γ should be computed along with the trajectory of the spacecraft. If the trajectory has sustained periods where the condition $\Gamma > \Gamma_+^*$ is violated, then this trajectory is a candidate for ejection from the body. On the other hand, should the condition be satisfied throughout the time period of interest, the trajectory is guaranteed to be bound to the body.

Evaluating the"Average" Sufficiency Condition

For the general design of an orbit, using the above osculating criterion is limited as the value of Γ that a trajectory has may vary considerably within one spacecraft orbit around the body. Thus it is of interest to examine what the average value of Γ is over a spacecraft orbit, so a more general criterion can be developed. Due to the strong perturbations acting on the spacecraft orbit the averaged values of the orbit elements will change in time. However, these changes will in general move more slowly than the variation in Γ over a single orbit, giving some justification to this approach. Standard averaging can be applied to Γ over one particle orbit

$$\bar{\Gamma} = \frac{1}{2\pi}\int_0^{2\pi} \Gamma \, dM \tag{12.59}$$

where M is the mean anomaly of the orbit. The body position about the Sun remains constant over one spacecraft orbit about the body. Note that after application of this averaging, the condition $\bar{\Gamma} > \Gamma_+^*$ is no longer a true sufficiency condition, but rather an indicator of when escape is likely. Also, it should be noted

that this averaging may not be valid for spacecraft in long-period orbits, when the body moves through a considerable range of true anomaly over one spacecraft period.

Evaluating this average yields:

$$\begin{aligned}\bar{\Gamma} = \frac{2d}{\epsilon^2 \mu_S} \Bigg[& -E + \Omega H \cos i - \frac{3ae}{2} \frac{\epsilon \mu_S \beta}{d^2} \left[\cos\omega \cos\lambda - \sin\omega \sin\lambda \cos i \right] \\ & + \frac{3a^2}{4} \frac{\Omega^2}{(1 + e_S \cos f)} \Bigg[\left(1 - \sin^2\lambda \sin^2 i\right) \left(1 + \frac{3e^2}{2}\right) + \\ & \frac{5e^2}{2} \left\{ \left(\cos^2\lambda - \sin^2\lambda \cos^2 i\right) \cos 2\omega - \sin 2\lambda \sin 2\omega \cos i \right\} \Bigg] \\ & - \frac{a^2}{8} \left(4\Omega^2 + \Omega'^2\right) \left(1 + \frac{3e^2}{2}\right) \Bigg] \end{aligned} \tag{12.60}$$

where λ is the longitude of the ascending node measured in inertial space.

If the mean orbit elements of a spacecraft trajectory are computed, using an analytical approach (see the next chapter), then the average value of Γ can be computed as the orbital elements change.

In Fig. 12.7 the stability bound on the Hayabusa spacecraft at Itokawa, the Rosetta spacecraft at Wirtanen (its original target), the ST-4 spacecraft at Tempel 1, and the NEAR-Shoemaker spacecraft at Eros is shown as a function of body–Sun distance. The bounding semi-major axis is found by numerically solving for the orbit semi-major axis from the equation $\bar{\Gamma} = \Gamma^*_+$, assuming the spacecraft is initially in a circular, zero inclination orbit about the body starting on the Sun line between the body and the Sun. Included in this figure are checks on the condition found by starting a spacecraft at the same initial conditions and integrating forward to see whether the spacecraft escapes or is in a bound orbit. The numerical integration included the effect of solar gravitation and solar radiation pressure without approximation. Each spacecraft/body pair was checked at body periapsis and apoapsis, with the resulting agreement being remarkably good. That the averaged condition is no longer a true sufficient condition can be seen in that the numerically determined limiting semi-major axis for the Hayabusa spacecraft was less than the "averaged" sufficiency condition. All the other cases, however, give excellent agreement. Were the numerically integrated spacecraft started in a different initial orbit – such as a retrograde or polar orbit – the agreement might not be so good, which is the usual case for such a sufficiency criterion [61].

The above formulae are too involved analytically to develop any sort of simple "back of the envelope" criterion. However, under some reasonable spacecraft assumptions such a criterion can be developed. If the nominal spacecraft orbit is relatively close to the body, then the semi-major axis will become relatively small as compared to the orbit energy, E. Using this as a justification to ignore all the terms except the orbit energy in the above averaged results yields the simpler, and approximate, result:

$$\Gamma \sim \frac{\epsilon d}{a} \tag{12.61}$$

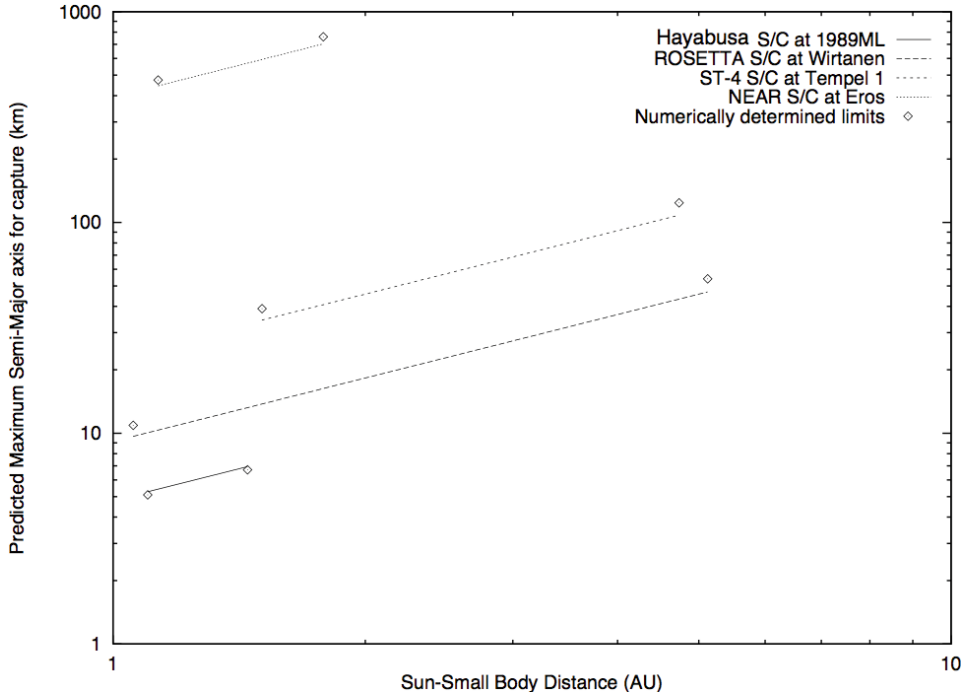


Fig. 12.7 Numerically and analytically determined stability limits.

Using this to evaluate the stability condition for both small and large values of β yields:

$$a < \begin{cases} \frac{1}{3^{4/3}} \left(\frac{\mu}{\mu_S}\right)^{1/3} d & \beta \ll 1 \\ \left(\frac{\mu}{\mu_S}\right)^{1/3} \frac{d}{4\sqrt{\beta}} & \beta \gg 1 \end{cases} \tag{12.62}$$

Expressing the $\beta \gg 1$ bound into natural units yields

$$a < \frac{d}{4}\sqrt{\frac{\mu}{(1+\rho)P_0/B}} \tag{12.63}$$

$$= \frac{1}{4}\sqrt{\frac{\mu}{a_{SRP}}} \tag{12.64}$$

which differs from the bound found in Eq. 12.35 by a factor of $\sqrt{3}$. It is important to note that these bounds were computed using very different approaches and that they represent different criteria. The current bound represents a necessary condition for escape and is derived under approximation. In contrast, in the context of the no-rotation model the earlier bound represents a necessary and sufficient condition

for escape for the class of terminator orbits. It is remarkable that they only differ by a factor of $\sqrt{3}$.

Using these bounds it is possible to place limits on the maximum orbit size attainable about small bodies. In Fig. 12.8 we show the limiting semi-major axis to small-body radius ratio as a function of small-body radius for different nominal densities. In this plot we assume a small body–Sun distance of 1 AU, a spacecraft mass-to-area ratio of 30 kg/m^2 and a spacecraft reflectance of 0. With these values we find a limit for the orbit radius over small body radius as

$$\frac{a}{R} < \frac{\sqrt{3}d}{4}\sqrt{\frac{4\pi\mathcal{G}/3}{(1+\rho)P_o/B_{SC}}}\sqrt{\sigma R} \tag{12.65}$$

$$\sim 10.9\ldots\sqrt{\sigma R} \tag{12.66}$$

where σ is the bulk density in g/cm^3 and R is the small-body radius in km. Note that the larger bound is plotted, the smaller bound for capture will equal 2/3 of the plotted value. For this relationship, we note that orbital operations are possible even around small, ∼10-meter-sized bodies (assuming asteroid-type densities).

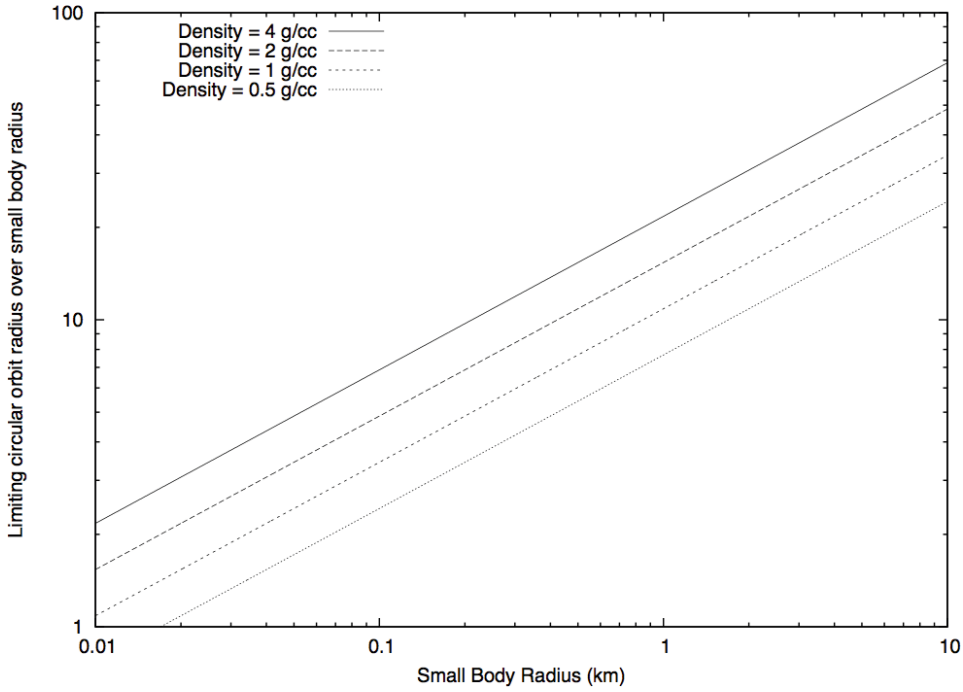


Fig. 12.8 Maximum a/R for being bound to a small body, using the less conservative (i.e., larger) limit.

Finally, it is important to note that these limits are all a function of the spacecraft mass-to-area ratio. This is a design parameter, however, and can be increased should larger orbits be desired about very small bodies. To scale this plot to other values of mass-to-area ratio multiply the given values by $\sqrt{B_{SC}/30}$, where B_{SC} is in kg/m^2, and to scale it to other values of spacecraft reflectance divide the given values by $1/\sqrt{1+\rho}$.

13. Solar Radiation Pressure: Averaged Analysis

Stating the same problem analyzed in the previous chapter as a perturbation problem allows us to introduce averaging to the dynamics of the system. In the following we show that the averaged dynamics of an orbiter subject to solar radiation pressure (SRP) and orbiting about a point mass can be solved in closed form with a very simple solution that still exhibits significantly complex behavior. Given this solution we are able to identify a set of stable "frozen orbits" that are suitable for spacecraft mission design and which enable a spacecraft to orbit about very small asteroids without the need for active control.

13.1 SRP Perturbation Formulation and Averaging

Given that the perturbation from solar radiation pressure is usually small, it is useful to cast the problem into a perturbation form. The problem is first studied in the absence of rotation, and then with rotation later. The perturbing potential associated with the simple SRP model used can be stated as

$$R_{SRP} = a_{SRP}\hat{\boldsymbol{d}} \cdot \mathbf{r} \tag{13.1}$$

where $\hat{\boldsymbol{d}}$ points from the Sun to the small body. The concept of averaging is introduced as this allows for the evaluation of the secular effect of the perturbation on this system. Using the Lagrange planetary equations, the potential can be averaged prior to application in this system. Thus the first step is to define the averaged potential

$$\bar{R}_{SRP} = \frac{1}{2\pi}\int_0^{2\pi} R_{SRP}\, dM \tag{13.2}$$

$$= a_{SRP}\hat{\boldsymbol{d}} \cdot \bar{\mathbf{r}} \tag{13.3}$$

where M is the mean anomaly. The average is taken over the unperturbed two-body motion of the spacecraft about the asteroid with the direction $\hat{\boldsymbol{d}}$ fixed initially. Thus one only needs to compute the average of the position vector, a classically

D.J. Scheeres, *Orbital Motion in Strongly Perturbed Environments*, Springer Praxis Books,
DOI 10.1007/978-3-642-03256-1_13, © Springer-Verlag Berlin Heidelberg 2012

known result:

$$\bar{\mathbf{r}} = -\frac{3a}{2}\boldsymbol{e} \tag{13.4}$$

where $\boldsymbol{e}$ is the eccentricity vector and has magnitude equal to the eccentricity and points towards the orbit periapsis. This leads to:

$$\bar{R}_{SRP} = -\frac{3aa_{SRP}}{2}\boldsymbol{e}\cdot\hat{\boldsymbol{d}} \tag{13.5}$$

The same form of the equations holds if $\hat{\boldsymbol{d}}$ is assumed to be fixed or assumed to be rotating about the sun at an angular rate $\dot{f}$.

Stated in this form, the rates of change of the orbit elements can be computed by substituting the averaged potential $\bar{R}_{SRP}$ into the Lagrange planetary equations. The simplest observation to make is that the semi-major axis is conserved on average, and hence the energy of the orbit is conserved on average. The Lagrange planetary equations for this case can be integrated in closed form, a fact originally realized by Mignard and Hénon [106], and worked out in detail in [156]. An alternate statement of the perturbation equations is given by Richter and Keller in [139] and uses the angular momentum and eccentricity vectors as the nominal orbit elements. These are not an immediately obvious choice, as this system is overdetermined in that the 6 components of these two vectors are functions of only 4 orbit elements with secular rates, e, i, ω and Ω. Still, this formulation has a significant advantage in being able to be solved in closed form, a fact originally realized by Richter and Keller for the non-rotating case. Our current analysis generalizes this to the rotation case, and also points out several important features of the solutions.

First the Gauss equations for the angular momentum and eccentricity vector are derived for the solar radiation pressure force and then these equations are averaged. The angular momentum equation is simple, and thus it is derived first. Recall the definition of angular momentum, $\boldsymbol{H} = \boldsymbol{r}\times\boldsymbol{v}$. From the fundamental form of the Gauss equations (Eqs. 4.92) this is

$$\dot{\boldsymbol{H}} = \frac{\partial \boldsymbol{H}}{\partial \boldsymbol{v}}\cdot\boldsymbol{a}_{SRP} \tag{13.6}$$

$$= \tilde{\boldsymbol{r}}\cdot\boldsymbol{a}_{SRP} \tag{13.7}$$

$$= -\tilde{\boldsymbol{a}}_{SRP}\cdot\boldsymbol{r} \tag{13.8}$$

where $\boldsymbol{a}_{SRP}$ is the constant SRP acceleration vector pointing away from the Sun. From above the average value of the position vector to find the average value of $\dot{\boldsymbol{H}}$.

$$\overline{\dot{\boldsymbol{H}}} = \frac{3a}{2}\tilde{\boldsymbol{a}}_{SRP}\cdot\boldsymbol{e} \tag{13.9}$$

The derivation of the averaged equation for $\dot{\boldsymbol{e}}$ is more complicated. First recall the definition of the eccentricity vector $\boldsymbol{e} = \frac{1}{\mu}\boldsymbol{v}\times\boldsymbol{H} - \hat{\boldsymbol{r}}$. Applying the Gauss equation

rule to this one must take partials of the angular momentum as well. Carrying this out yields

$$\dot{\boldsymbol{e}} = \frac{1}{\mu}\left[2\boldsymbol{r}\boldsymbol{v} - \boldsymbol{v}\boldsymbol{r} - (\boldsymbol{r}\cdot\boldsymbol{v})\boldsymbol{U}\right]\cdot\boldsymbol{a}_{SRP} \tag{13.10}$$

$$\dot{\boldsymbol{e}} = \frac{1}{\mu}\left[\tilde{\boldsymbol{a}}_{SRP}\cdot\boldsymbol{H} + (\boldsymbol{r}\boldsymbol{v} - (\boldsymbol{r}\cdot\boldsymbol{v})\boldsymbol{U})\cdot\boldsymbol{a}_{SRP}\right] \tag{13.11}$$

The term $\boldsymbol{r}\cdot\boldsymbol{v}$ averages to zero, leaving the average of $\boldsymbol{r}\boldsymbol{v}$ to be carried out, which is not zero. In Appendix C it is shown that $\overline{\boldsymbol{r}\boldsymbol{v}} = -\tilde{\boldsymbol{H}}/2$, leading to

$$\dot{\boldsymbol{e}} = \frac{3}{2\mu}\tilde{\boldsymbol{a}}_{SRP}\cdot\boldsymbol{H} \tag{13.12}$$

For the final step, scale $\boldsymbol{H}$ by a factor $\sqrt{\mu a}$, and note that a is constant on average.

$$\boldsymbol{h} = \frac{1}{\sqrt{\mu a}}\boldsymbol{H} \tag{13.13}$$

With this step, and substituting $\boldsymbol{a}_{SRP} = a_{SRP}\hat{\boldsymbol{d}}$, the equations become

$$\dot{\boldsymbol{e}} = \frac{3a_{SRP}}{2}\sqrt{\frac{a}{\mu}}\tilde{\hat{\boldsymbol{d}}}\cdot\boldsymbol{h} \tag{13.14}$$

$$\dot{\boldsymbol{h}} = \frac{3a_{SRP}}{2}\sqrt{\frac{a}{\mu}}\tilde{\hat{\boldsymbol{d}}}\cdot\boldsymbol{e} \tag{13.15}$$

With this normalization the angular momentum vector equals $\boldsymbol{h} = \sqrt{1-e^2}\hat{\boldsymbol{h}}$ with the constraint $\boldsymbol{e}\cdot\boldsymbol{e} + \boldsymbol{h}\cdot\boldsymbol{h} = 1$.

13.2 No-Rotation Solution

Equations 13.14 and 13.15 are time-invariant, linear equations if the rotation of $\hat{\boldsymbol{d}}$ is neglected. As such, they can be solved in closed form [139]. To facilitate this re-write the equations in a matrix form:

$$\begin{bmatrix}\dot{\boldsymbol{e}}\\ \dot{\boldsymbol{h}}\end{bmatrix} = \frac{3a_{SRP}}{2}\sqrt{\frac{a}{\mu}}\begin{bmatrix}\boldsymbol{0} & \tilde{\hat{\boldsymbol{d}}}\\ \tilde{\hat{\boldsymbol{d}}} & \boldsymbol{0}\end{bmatrix}\begin{bmatrix}\boldsymbol{e}\\ \boldsymbol{h}\end{bmatrix} \tag{13.16}$$

The general solution for this dynamical system is:

$$\begin{bmatrix} \boldsymbol{e}(t) \\ \boldsymbol{h}(t) \end{bmatrix} = \Phi(t - t_o) \begin{bmatrix} \boldsymbol{e}_o \\ \boldsymbol{h}_o \end{bmatrix} \tag{13.17}$$

$$\begin{aligned} \Phi(t) = &\begin{bmatrix} \hat{\boldsymbol{d}}\hat{\boldsymbol{d}} & \mathbf{0} \\ \mathbf{0} & \hat{\boldsymbol{d}}\hat{\boldsymbol{d}} \end{bmatrix} - \cos\left(\frac{3a_{SRP}}{2}\sqrt{\frac{a}{\mu}}t\right) \begin{bmatrix} \tilde{\hat{\boldsymbol{d}}} \cdot \tilde{\hat{\boldsymbol{d}}} & \mathbf{0} \\ \mathbf{0} & \tilde{\hat{\boldsymbol{d}}} \cdot \tilde{\hat{\boldsymbol{d}}} \end{bmatrix} \\ &+ \sin\left(\frac{3a_{SRP}}{2}\sqrt{\frac{a}{\mu}}t\right) \begin{bmatrix} \mathbf{0} & \tilde{\hat{\boldsymbol{d}}} \\ \tilde{\hat{\boldsymbol{d}}} & \mathbf{0} \end{bmatrix} \end{aligned} \tag{13.18}$$

The state transition matrix is also an orthonormal matrix in that $\Phi^{-1} = \Phi^T$, which was the main motivation behind introducing the normalization to the angular momentum vector $\boldsymbol{h}$.

This result assumes, in its derivation, that the spacecraft remains bound to the asteroid. It also does not account for the offset in mean orbit plane along the direction $\hat{\boldsymbol{d}}$. It is interesting to note, however, that both rates of change in $\boldsymbol{h}$ and $\boldsymbol{e}$ are normal to $\hat{\boldsymbol{d}}$, meaning that a displacement of the center of orbit in this direction should not affect these averaged equations.

There are two simple applications of these secular equations. First is to find all possible frozen orbits, or orbits where the average angular momentum and eccentricity vectors remain constant. Second is to study some special solutions.

13.2.1 Frozen Orbits

Frozen orbits are defined for all $\boldsymbol{e}$ and $\boldsymbol{h}$ such that $\dot{\boldsymbol{e}} = \dot{\boldsymbol{h}} = 0$. This can be reduced to the two conditions:

$$\tilde{\hat{\boldsymbol{d}}} \cdot \boldsymbol{h} = 0 \tag{13.19}$$

$$\tilde{\hat{\boldsymbol{d}}} \cdot \boldsymbol{e} = 0 \tag{13.20}$$

or that either the magnitude of the vector be zero or it be aligned with $\hat{\boldsymbol{d}}$. Since $\boldsymbol{h}$ and $\boldsymbol{e}$ are normal to each other, they cannot both be aligned with $\hat{\boldsymbol{d}}$ at the same time, thus one of them must be zero. Not considered is the case of $h = 0$ as this is a rectilinear orbit, leaving $e = 0$ and $\boldsymbol{h} = \pm h\hat{\boldsymbol{d}}$ being the only viable frozen orbit solution. This agrees exactly with the relative equilibrium found for this orbit in the exact analysis, a circular orbit with angular momentum vector oriented towards or away from the Sun.

13.2.2 Special Solutions

Next consider the general solution for the system eccentricity when started in an initially circular orbit, resulting in the closed-form solution

$$\boldsymbol{e}(t) = \sin(\omega t)\tilde{\hat{\boldsymbol{d}}} \cdot \hat{\boldsymbol{h}}_o \tag{13.21}$$

$$\boldsymbol{h}(t) = \cos(\omega t)\boldsymbol{h}_o + (1 - \cos(\omega t))\hat{\boldsymbol{d}}\hat{\boldsymbol{d}} \cdot \boldsymbol{h}_o \tag{13.22}$$

Of special interest is the eccentricity of the orbit and the inclination, relative to the direction $\hat{\boldsymbol{d}}$. To set up these solutions note that the initial inclination is defined by $\cos i_o = \hat{\boldsymbol{h}}_o \cdot \hat{\boldsymbol{d}}$. The eccentricity solution is then found as

$$e = |\sin(\omega t)| \sin i_o \tag{13.23}$$

Thus, an orbit started in the plane perpendicular to $\hat{\boldsymbol{d}}$, or $i_o = 0, \pi$, will remain circular. As the inclination varies from these values, so does the maximum eccentricity reached, up to a 90° orbit which will pass through a rectilinear orbit every period $2\pi/\omega$. The inclination solution can be found to follow

$$\cos i = \frac{\cos i_o}{\sqrt{1 - \sin^2 i_o \sin^2(\omega t)}} \tag{13.24}$$

Again, an orbit started perpendicular to $\hat{\boldsymbol{d}}$ remains so, and the maximum excursion from this plane is the initial inclination. Of course, any other special case can be worked out, but these are of interest due to the practical applications of circular orbits.

13.3 Rotation Solution

Now consider the perturbation solution when rotation of $\hat{\boldsymbol{d}}$ occurs. The rotation rate equals the true anomaly rate of $\dot{f}$, and the rotation direction is always perpendicular to $\hat{\boldsymbol{d}}$ and defined as the direction $\hat{\boldsymbol{z}}$. Under the assumption that the orbit rate about the asteroid is fast as compared to the orbit rate of the asteroid about the Sun the same equations for $\dot{\boldsymbol{e}}$ and $\dot{\boldsymbol{h}}$ can be used. A complication occurs as $\boldsymbol{e}$ and $\boldsymbol{h}$ are specified in an inertially fixed frame but now the perturbing force direction $\hat{\boldsymbol{d}}$ rotates. To alleviate this the equations for the eccentricity and angular momentum vectors can be shifted into a rotating frame using the transport equation. Given a frame rotational velocity vector of $\dot{f}\hat{\boldsymbol{z}}$ yields

$$\dot{\boldsymbol{e}}_r + \dot{f}\tilde{\hat{\boldsymbol{z}}} \cdot \boldsymbol{e} = \frac{3a_{SRP}}{2}\sqrt{\frac{a}{\mu}}\tilde{\hat{\boldsymbol{d}}} \cdot \boldsymbol{h} \tag{13.25}$$

$$\dot{\boldsymbol{h}}_r + \dot{f}\tilde{\hat{\boldsymbol{z}}} \cdot \boldsymbol{h} = \frac{3a_{SRP}}{2}\sqrt{\frac{a}{\mu}}\tilde{\hat{\boldsymbol{d}}} \cdot \boldsymbol{e} \tag{13.26}$$

where the r subscript indicates time derivative with respect to a rotating frame and will be dropped from this point on. Thus, again, the secular equations can be written as a linear system

$$\begin{bmatrix} \dot{\boldsymbol{e}} \\ \dot{\boldsymbol{h}} \end{bmatrix} = \frac{3a_{SRP}}{2}\sqrt{\frac{a}{\mu}} \begin{bmatrix} -\dfrac{2\dot{f}}{3a_{SRP}}\sqrt{\dfrac{\mu}{a}}\tilde{\hat{\boldsymbol{z}}} & \tilde{\hat{\boldsymbol{d}}} \\ \tilde{\hat{\boldsymbol{d}}} & -\dfrac{2\dot{f}}{3a_{SRP}}\sqrt{\dfrac{\mu}{a}}\tilde{\hat{\boldsymbol{z}}} \end{bmatrix} \begin{bmatrix} \boldsymbol{e} \\ \boldsymbol{h} \end{bmatrix} \tag{13.27}$$

While this is a linear differential equation, it is no longer time-invariant as both $\dot{f}$ and a_{SRP} now vary in time. It is important to note that the ratio $\dot{f}/a_{SRP}$ is time-invariant, as both vary inversely with d^2:

$$\frac{2\dot{f}}{3a_{SRP}}\sqrt{\frac{\mu}{a}} = \frac{2B_{SC}\sqrt{\mu_{Sun}A(1-E^2)}}{3(1+\rho)P_0}\sqrt{\frac{\mu}{a}} \tag{13.28}$$

This quantity is a constant for a given asteroid, spacecraft and spacecraft orbit. As it is inconvenient to work with, define the equivalent quantity, following [106]:

$$\tan\Lambda = \frac{3(1+\rho)P_0}{2B_{SC}}\sqrt{\frac{a}{\mu\mu_{Sun}A(1-E^2)}} \tag{13.29}$$

As the SRP perturbation becomes strong, $\Lambda \to \pi/2$, and as it becomes weak $\Lambda \to 0$. Broken down into more fundamental orbital quantities it can be expressed as

$$\tan\Lambda = \frac{3(1+\rho)P_0}{2B}\frac{1}{V_{lc}H_{sun}} \tag{13.30}$$

where V_{lc} is the local circular speed of the spacecraft about the asteroid and H_{sun} is the specific angular momentum of the asteroid about the Sun.

Despite the time-invariance of the ratio, the multiplying factor of the matrix is still time-varying. This can be eliminated, however, by changing the independent parameter from time to the true anomaly of the asteroid about the Sun. To make this transformation let, for example, $\dot{\boldsymbol{e}} = \frac{d\boldsymbol{e}}{df}\dot{f} = \boldsymbol{e}'\dot{f}$. Then the factor in front of the matrix becomes the newly defined SRP strength parameter, leading to the time-invariant linear differential equations:

$$\begin{bmatrix}\boldsymbol{e}'\\ \boldsymbol{h}'\end{bmatrix} = \begin{bmatrix}-\tilde{\hat{\boldsymbol{z}}} & \tan\Lambda\tilde{\hat{\boldsymbol{d}}}\\ \tan\Lambda\tilde{\hat{\boldsymbol{d}}} & -\tilde{\hat{\boldsymbol{z}}}\end{bmatrix}\begin{bmatrix}\boldsymbol{e}\\ \boldsymbol{h}\end{bmatrix} \tag{13.31}$$

where the direction $\hat{\boldsymbol{d}}$ is now fixed in the rotating coordinate frame. It is convenient to redefine the independent variable from true anomaly f to a new angle $\psi = f/\cos\Lambda$.

The solution matrix can again be reduced to elementary functions, yielding the explicit solution:

$$\begin{bmatrix}\boldsymbol{e}(\psi)\\ \boldsymbol{h}(\psi)\end{bmatrix} = \Phi_r(\psi-\psi_o)\begin{bmatrix}\boldsymbol{e}_o\\ \boldsymbol{h}_o\end{bmatrix} \tag{13.32}$$

$$\begin{aligned}\Phi(\psi) = {} & \cos(\psi)I_{6\times 6} + (1-\cos(\psi))\times\\ & \begin{bmatrix}\cos^2\Lambda\hat{\boldsymbol{z}}\hat{\boldsymbol{z}} + \sin^2\Lambda\hat{\boldsymbol{d}}\hat{\boldsymbol{d}} & -\sin\Lambda\cos\Lambda\left(\hat{\boldsymbol{z}}\hat{\boldsymbol{d}}+\hat{\boldsymbol{d}}\hat{\boldsymbol{z}}\right)\\ -\sin\Lambda\cos\Lambda\left(\hat{\boldsymbol{z}}\hat{\boldsymbol{d}}+\hat{\boldsymbol{d}}\hat{\boldsymbol{z}}\right) & \cos^2\Lambda\hat{\boldsymbol{z}}\hat{\boldsymbol{z}} + \sin^2\Lambda\hat{\boldsymbol{d}}\hat{\boldsymbol{d}}\end{bmatrix}\\ & + \sin(\psi)\begin{bmatrix}-\cos\Lambda\tilde{\hat{\boldsymbol{z}}} & \sin\Lambda\tilde{\hat{\boldsymbol{d}}}\\ \sin\Lambda\tilde{\hat{\boldsymbol{d}}} & -\cos\Lambda\tilde{\hat{\boldsymbol{z}}}\end{bmatrix}\end{aligned} \tag{13.33}$$

The solutions are periodic in ψ, repeating every true anomaly $2\pi \cos \Lambda$. Thus, over one heliocentric orbit the solution will repeat $1/\cos \Lambda$ times. As the perturbation grows large, and Λ approaches $\pi/2$, the solution will repeat many times. Conversely, as the perturbation grows small the solution will repeat only once every heliocentric orbit. Also, the state transition matrix is again orthonormal and defines a rotation matrix in six-dimensional space.

13.3.1 Frozen Orbits

The existence of frozen orbits for this system are discussed again. There are now additional possibilities not present in the case of no-rotation. Again, to find them search for solutions to the algebraic equations

$$-\tilde{\hat{\boldsymbol{z}}} \cdot \boldsymbol{e} + \tan \Lambda \tilde{\hat{\boldsymbol{d}}} \cdot \boldsymbol{h} = 0 \tag{13.34}$$

$$-\tilde{\hat{\boldsymbol{z}}} \cdot \boldsymbol{h} + \tan \Lambda \tilde{\hat{\boldsymbol{d}}} \cdot \boldsymbol{e} = 0 \tag{13.35}$$

The vectors $\boldsymbol{e}$ and $\boldsymbol{h}$ remain orthogonal but now both must have non-zero magnitude for a frozen orbit to exist. A detailed study of this matrix and its null spaces shows that there are two classes of solutions.

Ecliptic Plane Frozen Orbits First choose $\boldsymbol{e}$ parallel to $\hat{\boldsymbol{d}}$ and $\boldsymbol{h}$ parallel to $\hat{\boldsymbol{z}}$, then the second equation is identically solved and the first equation reduces to a vector equation parallel to the unit vector $\hat{\boldsymbol{t}}$. It turns out that the direction in which these vectors point is important, thus introduce the test solutions $\boldsymbol{e} = e(\hat{\boldsymbol{d}} \cdot \hat{\boldsymbol{e}})\hat{\boldsymbol{d}}$ and $\boldsymbol{h} = h(\hat{\boldsymbol{z}} \cdot \hat{\boldsymbol{h}})\hat{\boldsymbol{z}}$. Resolving the first equation along the direction $\hat{\boldsymbol{t}}$ then yields:

$$e(\hat{\boldsymbol{d}} \cdot \hat{\boldsymbol{e}}) + \tan \Lambda h(\hat{\boldsymbol{z}} \cdot \hat{\boldsymbol{h}}) = 0 \tag{13.36}$$

Note again that $h = \sqrt{1 - e^2}$, which simplifies the expression to

$$\frac{e}{\sqrt{1 - e^2}} = -(\hat{\boldsymbol{d}} \cdot \hat{\boldsymbol{e}})(\hat{\boldsymbol{z}} \cdot \hat{\boldsymbol{h}}) \tan \Lambda \tag{13.37}$$

Thus there are two conditions for a frozen orbit to exist in this configuration:

$$-(\hat{\boldsymbol{d}} \cdot \hat{\boldsymbol{e}})(\hat{\boldsymbol{z}} \cdot \hat{\boldsymbol{h}}) = 1 \tag{13.38}$$

$$e = \sin \Lambda \tag{13.39}$$

This class of frozen orbit was originally discussed in [156] and were called ecliptic frozen orbits. The orbit lies in the same plane as the asteroid's heliocentric orbit. If the orbit normal is aligned with the heliocentric orbit, periapsis must point towards the Sun, otherwise if the orbit normal is anti-parallel to the heliocentric orbit normal, periapsis must point away from the Sun. As the perturbation strength grows the orbit approaches rectilinear, while if the perturbation strength vanishes the orbit approaches circular. Due to this, these orbits are not preferred for strongly

perturbed situations, as the periapsis has a low altitude. Also, these orbits cross through the body's shadow.

Terminator Plane Frozen Orbits A second frozen orbit solution, called solar plane-of-sky orbits in [156], also exist and are more useful for highly perturbed systems. Now, choose $\boldsymbol{e}$ parallel to $\hat{\boldsymbol{z}}$ and $\boldsymbol{h}$ parallel to $\hat{\boldsymbol{d}}$ so that Eq. 13.34 is identically solved and Eq. 13.35 reduces to a vector equation parallel to the unit vector $\hat{\boldsymbol{t}}$. Again the direction in which these vectors point is important, thus introduce the test solutions $\boldsymbol{e} = e(\hat{\boldsymbol{z}} \cdot \hat{\boldsymbol{e}})\hat{\boldsymbol{z}}$ and $\boldsymbol{h} = h(\hat{\boldsymbol{d}} \cdot \hat{\boldsymbol{h}})\hat{\boldsymbol{d}}$. Resolving Eq. 13.35 along the direction $\hat{\boldsymbol{t}}$ then yields:

$$e \tan \Lambda(\hat{\boldsymbol{z}} \cdot \hat{\boldsymbol{e}}) + h(\hat{\boldsymbol{d}} \cdot \hat{\boldsymbol{h}}) = 0 \tag{13.40}$$

which simplifies again to

$$\frac{e}{\sqrt{1-e^2}} = -(\hat{\boldsymbol{z}} \cdot \hat{\boldsymbol{e}})(\hat{\boldsymbol{d}} \cdot \hat{\boldsymbol{h}}) \cot \Lambda \tag{13.41}$$

The two conditions for a frozen orbit to exist in this configuration are:

$$-(\hat{\boldsymbol{z}} \cdot \hat{\boldsymbol{e}})(\hat{\boldsymbol{d}} \cdot \hat{\boldsymbol{h}}) = 1 \tag{13.42}$$

$$e = \cos \Lambda \tag{13.43}$$

This orbit lies in the plane perpendicular to the Sun-line, commonly referred to as the terminator plane. If the orbit normal points towards the Sun, periapsis must point above the orbit plane along the positive $\hat{\boldsymbol{z}}$-axis, otherwise if the orbit normal points away from the Sun, periapsis must point below the orbit plane. As the perturbation strength grows the orbit becomes more circular, while if the perturbation strength vanishes the orbit approaches rectilinear. Due to this, these orbits are preferred for strongly perturbed situations. Also, these orbits avoid the asteroid's shadow. Finally, these orbits are the natural continuation of the equilibrium orbits in the no-rotation case.

13.3.2 Special Solutions

Finally for this case the closed-form solution for eccentricity and inclination are studied when the orbit is initially circular. The solution is more complex in this case but can still be stated in a relatively simple form for eccentricity

$$\begin{aligned} e(\psi) = 2 \sin \Lambda |\sin(\psi/2)| \Big\{ & 1 - \sin^2 \Lambda \sin^2(\psi/2) \\ & - \left[\cos \Lambda \sin(\psi/2) \left(\hat{\boldsymbol{h}}_o \cdot \hat{\boldsymbol{t}} \right) + \cos(\psi/2) \left(\hat{\boldsymbol{h}}_o \cdot \hat{\boldsymbol{d}} \right) \right]^2 \Big\}^{1/2} \end{aligned} \tag{13.44}$$

Some special cases are considered to make this more apparent. First, assume that the orbit normal is parallel to $\hat{\boldsymbol{z}}$ (see Fig. 13.1). Then the solution reduces to:

$$e(\psi) = 2\sin\Lambda|\sin(\psi/2)|\sqrt{1-\sin^2\Lambda\sin^2(\psi/2)} \tag{13.45}$$

The maximum eccentricity behavior is actually somewhat complex for this case. If $\tan\Lambda < 1$, then the maximum occurs at $\psi = \pi$ and equals $\sin 2\Lambda$. However, if $\tan\Lambda > 1$ then the maximum is $e = 1$ and occurs at $\psi = \arccos(-\cot^2\Lambda)$.

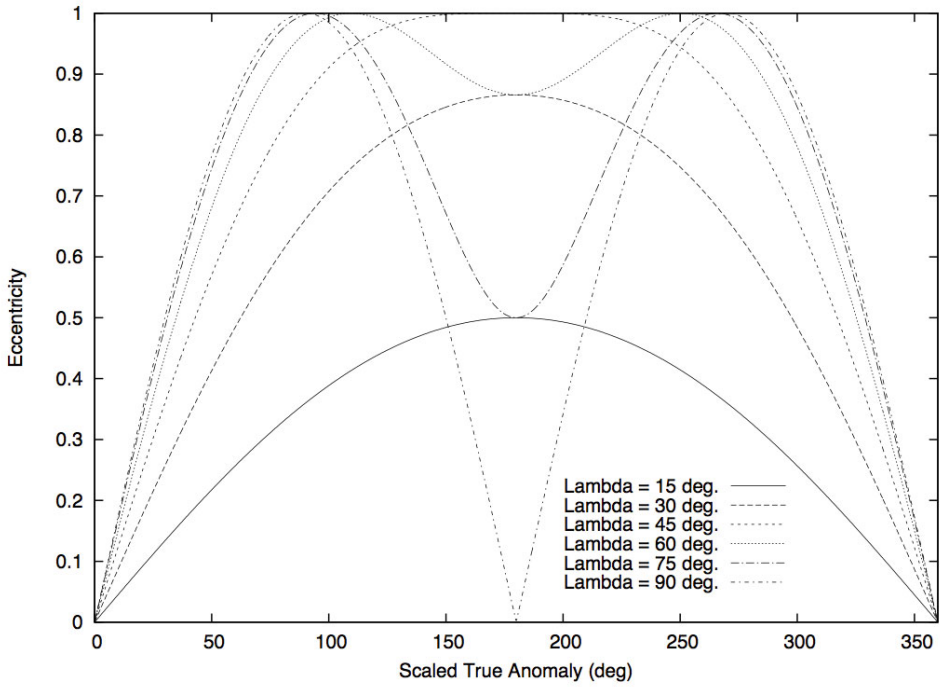


Fig. 13.1 Eccentricity as a function of the scaled true anomaly (ψ) for different values of Λ). Eccentricity is initially zero and the angular momentum vector points in the positive $\hat{\boldsymbol{z}}$-direction.

Next, assume the orbit normal is parallel to $\hat{\boldsymbol{d}}$ (see Fig. 13.2). Then the eccentricity varies as

$$e(\psi) = \frac{1}{2}\sin(2\Lambda)(1-\cos\psi) \tag{13.46}$$

The maximum now occurs at $\psi = \pi$ and equals $\sin 2\Lambda$. It is interesting to note that for a weak and a strong perturbation the maximum eccentricity approaches 0 in this case, but for a "median" perturbation of $\tan\Lambda = 1$ the maximum eccentricity is unity, leading to impact. For this case the proper choice of eccentricity vector will yield a frozen orbit.

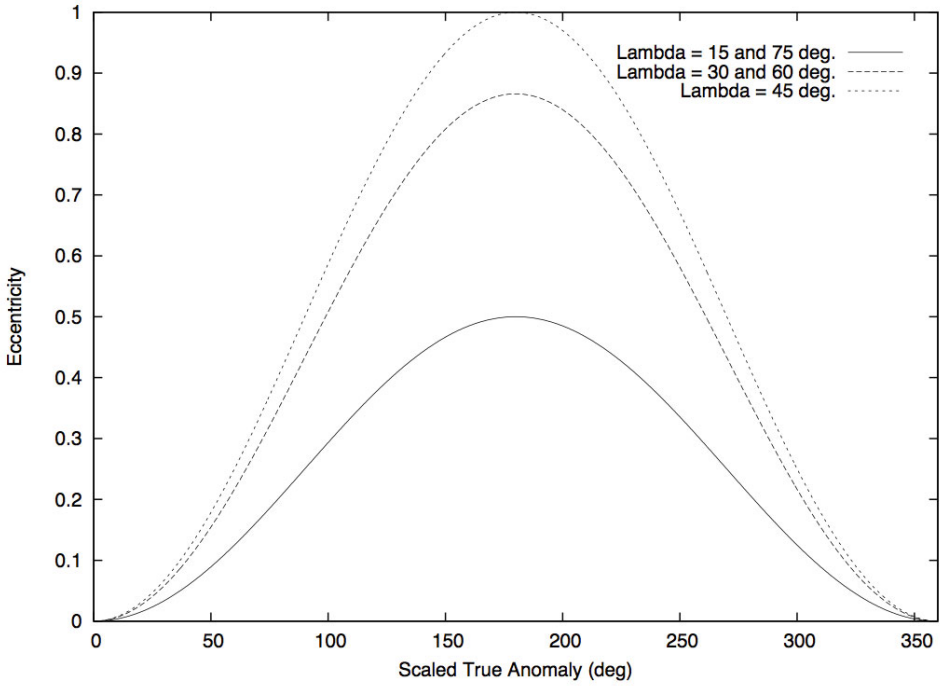


Fig. 13.2 Eccentricity as a function of the scaled true anomaly (ψ) for different values of Λ). Eccentricity is initially zero and the angular momentum vector points in the positive $\hat{\boldsymbol{d}}$-direction.

Finally, assume the orbit is initially parallel to $\hat{\boldsymbol{t}}$ (see Fig. 13.3). Then the eccentricity varies as:

$$e(\psi) = \sin\Lambda|\sin\psi| \tag{13.47}$$

and the maximum eccentricity occurs at $\psi = \pm\pi/2$ and reaches a value of $\sin\Lambda$.

For all three cases we note that the generic dynamics of an initially circular orbit are quite complex. Thus the terminator frozen orbits seem to be unique in their ability to maintain a constant eccentricity in the presence of the SRP perturbation. We do not consider the inclination variation for these problems in closed form, due to its complex dynamics.

13.4 Characteristic Values of Λ

From the above we see that the parameter Λ exerts a significant influence over the expected orbital dynamics of a spacecraft in the vicinity of a small body. Thus it is instructive to consider the expected values of this parameter across a range of small-body sizes and densities. In Fig. 13.4 the values of this parameter are shown

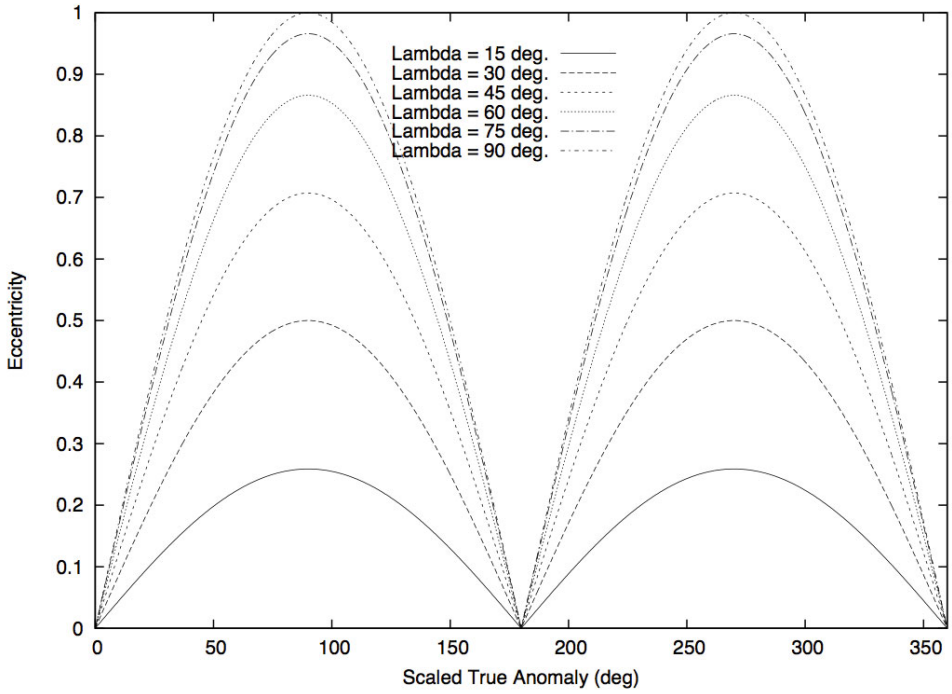


Fig. 13.3 Eccentricity as a function of the scaled true anomaly (ψ) for different values of Λ). Eccentricity is initially zero and the angular momentum vector points in the positive $\hat{\boldsymbol{t}}$-direction.

for a range of small-body sizes as a function of a/R for some characteristic values of the other parameters. These plots assume a zero reflectance, a small-body orbit parameter $A(1-E^2)$ equal to 1 AU, a spacecraft area to mass ratio of 30 kg/m^2 (see Table 12.1 for a range of plausible values of this parameter) and a small body bulk density of 2 g/cm^3. With these values the equation for Λ becomes

$$\tan\Lambda \sim \frac{1.5}{R}\sqrt{\frac{a}{R}} \tag{13.48}$$

To scale this result to other values of spacecraft reflectance multiply by $(1+\rho)$, to other values of spacecraft mass to area ratio multiply by $30/B_{SC}$ in units of kg/m^2, to scale to other values of small-body bulk density multiply by $\sqrt{2/\sigma}$ in units of g/cm^3, and to scale to other values of small-body orbit parameter multiply by $1/\sqrt{P}$ in units of AU.

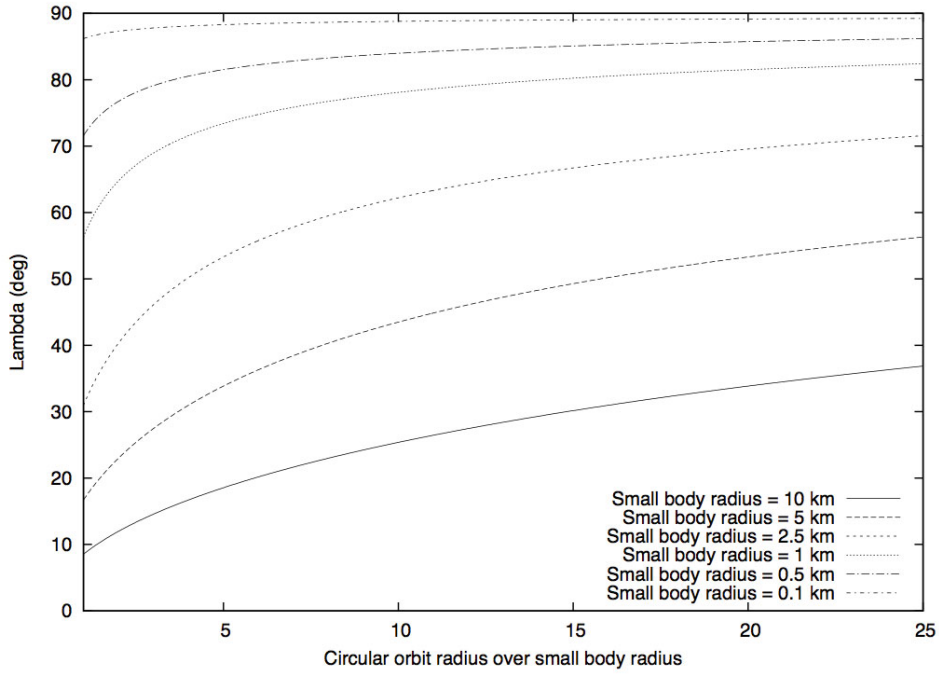


Fig. 13.4 Solar radiation pressure parameter Λ as a function of semi-major axis over small-body radius, a/R, and small-body radius. Note that the plotted upper limits of the a/R values may violate the limits for a bound orbit; see Fig. 12.8 to find these limits.

13.5 Discussion

The existence of this closed-form solution for averaged motion in the case of strongly perturbed solar radiation pressure is quite remarkable. Although the fact that this system can be integrated in closed form was known for some time, [106, 139, 156], the explicit form of the solution given here, as applied to motion about an asteroid or comet in a heliocentric orbit of arbitrary eccentricity, has only been known more recently [161]. In addition to providing a relatively precise description of motion for this case, it also provides a specific strategy to use when orbiting about a small asteroid: place the spacecraft in the terminator plane with an eccentricity specified as $e = \cos \Lambda$ with the appropriate direction of the argument of periapsis. Such an orbit will then naturally follow the Sun and can be stable over extremely long timespans, barring adverse interactions with the asteroid's non-spherical gravity field.

14. Small Bodies: Asteroid 25143 Itokawa

The previous two chapters are now applied to the specific problem of orbital mechanics about a small body. For definiteness we focus on the asteroid Itokawa (with a mean radius of 160 meters), for which a shape model, rotation state and mass is available. Itokawa was visited by the Hayabusa spacecraft which collected samples from that body before its successful return to Earth in June 2010. That spacecraft did not enter sustained orbit about the asteroid, but instead implemented a hovering approach for all close proximity dynamics. Still, it is a question of practical interest whether such an orbital approach could have been implemented at such a small body. In this chapter we apply the theories outlined previously and derive approximate limits for *minimum* orbit radii about such bodies. A general overall question which this analysis poses is whether there are minimum size limits on an asteroid beyond which it cannot be orbited. This question was recently investigated by Shupe [177], where for a spacecraft modeled after the Orion crewed vehicle it was found that the main barrier to achieving stable orbit was not solar radiation pressure but the central-body gravity field and shape, and the size and shape of the orbiting spacecraft.

14.1 Model of Itokawa

The necessary parameters and models for this analysis are the total mass of the asteroid, the rotation period of the asteroid, its rotation pole, its heliocentric orbit and the shape of the body (as we will apply a constant density approximation to find its gravitational field). Also needed is the spacecraft mass-to-area ratio for the SRP calculations.

The shape model of Itokawa used in this analysis was obtained from imaging of the asteroid during the rendezvous mission. The highest precision version of the shape model has a surface resolution of $\sim$10 cm, but we use a significantly lower-resolution model for this work [45]. The shape model has again been transformed into a triangular plate model centered at its volumetric center and oriented along its principal axes of inertia (assuming constant density). For gravitational calcula-

D.J. Scheeres, *Orbital Motion in Strongly Perturbed Environments*, Springer Praxis Books,
DOI 10.1007/978-3-642-03256-1_14, © Springer-Verlag Berlin Heidelberg 2012

Table 14.1 Itokawa and Hayabusa parameters

Parameter	Symbol	Value	Units
Itokawa Semi-Major Axis	A	1.324	AU
Itokawa Eccentricity	E	0.280	–
Gravitational Parameter	μ	2.36×10^{-9}	km^3/s^2
Rotation Period	T_r	12.132	hours
Pole Declination	δ	~ -90	degrees
Mean Radius	r_o	0.162	km
Resonance Radius	r_s	0.484	km
Oblateness Gravity Coefficient	$R_o^2 C_{20}$	-8.513×10^{-3}	km^2
Ellipticity Gravity Coefficient	$R_o^2 C_{22}$	3.713×10^{-3}	km^2
Spacecraft Mass to Area Ratio	B	30	kg/m^2

tions outside of the circumscribing sphere of the asteroid we used a gravity field expanded up to degree and order 16, while when close to the body we used the polygonal gravitational field algorithm of Werner, described previously. For analytical calculations the analysis is restricted to the second-degree and -order gravity field coefficients.

14.2 Maximum Orbital Stability Limits on Semi-Major Axis

The first item of discussion is the bounds under which a spacecraft is trapped in orbit about a small asteroid, summarized in Chapter 12. The analysis by Dankowicz yields a simple limit on semi-major axis that generally ensures escape from the small body if violated [27]. If the semi-major axis is larger than a_{Max}, given below, escape will generally occur

$$a_{Max} = \frac{\sqrt{3}}{4}\sqrt{\frac{\mu B}{(1+\rho)P_0}}d \tag{14.1}$$

Using a very different approach, Scheeres and Marzari develop a sufficient condition for stability about a small body [171]. If the semi-major axis is less than a_{Min}, then escape is prohibited. In the general problem it is sometimes possible for the semi-major axis to evolve from a value less than a_{Min} to one larger than this limit

$$a_{Min} = \frac{1}{4}\sqrt{\frac{\mu B}{(1+\rho)P_0}}d \tag{14.2}$$

These limiting semi-major axes scale linearly with the small-body distance from the sun. Thus, escape is most likely to occur at perihelion. Figure 14.1 shows an example of an initial orbit that lies within this bound but, as the small body approaches perihelion, crosses it and escapes. They are also proportional to $\sqrt{B}$,

meaning that as the mass-to-area ratio of a body increases, its limiting semi-major axis also increases. For Itokawa (assuming no photon reflection, or $\rho = 0$) the semi-major axis for the necessary condition for the spacecraft to be bound to the asteroid is ∼1 km at periapsis, while the sufficient distance for escape is ∼1.73 km, and at this limit a spacecraft is barely bound. The numerically determined limit for the spacecraft to be bound to the asteroid, assuming an initial orbit in the Sun terminator plane, is found to be 1.75 km at perihelion. Lower bounds exist due to gravitational interactions with the body mass distribution, and are discussed later.

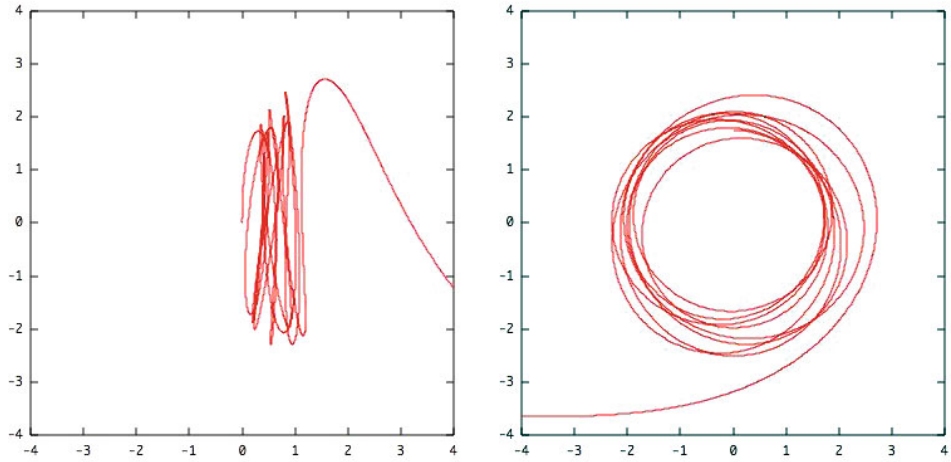


Fig. 14.1 Spacecraft orbit about an asteroid in an elliptic orbit about the Sun. The orbit is initially stable, but due to the decreasing asteroid–Sun distance becomes unstable and escapes. The figures show the view down onto the asteroid orbit plane (left) and the view from the Sun (right).

14.3 Nominal Stable Orbit Design

This chapter tacitly assumes that the spacecraft has been placed in one of the stable frozen orbits about the asteroid discussed in Chapter 13. The specification of these orbits are quite simple and are reviewed here. First, the orbiter is given a semi-major axis less than the escape limit, and above the estimated lower limits for stability. Second, the orbit plane is placed in the Sun terminator plane relative to the asteroid. The eccentricity of the orbit is chosen to be $e = \cos \Lambda$, where Λ is computed in Eq. 13.29. For Hayabusa at Itokawa we find

$$\tan \Lambda \sim 21\sqrt{a} \tag{14.3}$$

Thus across the range of semi-major axis from 0.75–1.75 km, this yields an angle of ∼86.8 to ∼87.9°, meaning that the frozen orbit eccentricity is on the order of 0.055–0.037, all close to circular.

The argument of periapsis is chosen depending on whether the spacecraft's angular momentum vector points towards or away from the Sun. If the angular momentum vector points towards the Sun the argument of periapsis should point directly above the orbital plane, in the direction of the asteroid's heliocentric angular momentum vector, while if the spacecraft angular momentum vector points away from the Sun the periapsis vector must point below the orbital plane. The ecliptic plane frozen orbit solutions are not discussed here, as these become rectilinear for strong SRP perturbations and thus are not preferred.

There exists a minimum semi-major axis for stability as well, although this is more difficult to uniquely identify. The lower stability limit arises due to two different effects, each related to the asteroid's gravity field. First, the oblateness of the asteroid can cause the spacecraft orbit to precess relative to its preferred terminator plane orientation. If the precession is great enough it can overpower the SRP precession rate, which then allows the eccentricity to go through much larger variations, potentially impacting the asteroid surface or inducing the second type of instability. The other instability is due to interaction of the spacecraft with the C_{22} gravity field parameter, which can lead to large fluctuations in the orbiter angular momentum and energy from orbit to orbit, as analyzed previously. In this regime the semi-major axis and eccentricity begin to evolve chaotically and thus may either impact with the asteroid, escape, or at least follow a path that is difficult to predict due to its fluctuations. In Fig. 14.2 example orbits are presented at the extreme orbit sizes for stability.

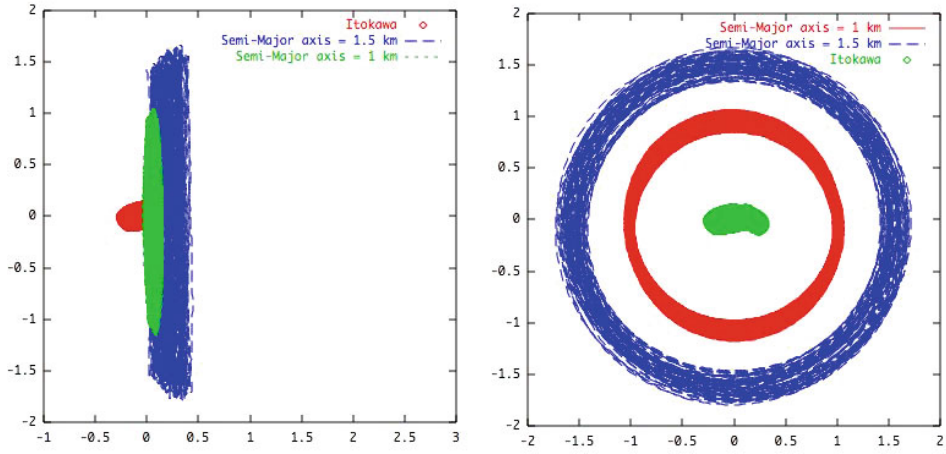


Fig. 14.2 Minimum and maximum orbit radii for stable terminator orbits about Itokawa. The figure on the left shows the view perpendicular to the Sun-line from within the asteroid orbit plane; the figure on the right shows the view from the Sun.

14.4 Mass Distribution Perturbation and Averaging

Now the classical results for the averaged effect of a second-degree and -order gravity field are reviewed. In order to relate this analysis to the SRP frame requires that the oblateness effect be specified in a general orbit frame, and not one chosen so that the inclination is measured from the symmetry axis. If the maximum moment of inertia axis of the mass distribution is stipulated as $\hat{\mathbf{p}}$ and the mass is assumed to have a symmetric distribution about this axis the perturbation function is

$$R_{20} = -\frac{\mu C_{20}}{2r^3}\left[1 - 3\left(\hat{\boldsymbol{r}} \cdot \hat{\boldsymbol{p}}\right)^2\right] \tag{14.4}$$

The orbit frame does not need to be specified when stated in this form, yet averaging can still be carried out (see Appendix C). Doing so yields

$$\bar{R}_{20} = \frac{n^2 C_{20}}{4(1-e^2)^{3/2}}\left[1 - 3\left(\hat{\boldsymbol{h}} \cdot \hat{\boldsymbol{p}}\right)^2\right] \tag{14.5}$$

where $\hat{\mathbf{h}}$ is the unit vector along the orbit normal.

Analysis of this potential shows that $\bar{R}_{20} =$ constant, implying that $\hat{\mathbf{h}} \cdot \hat{\mathbf{p}}$ is constant, which means that the orbit plane inclination relative to the rotation pole $\hat{\mathbf{p}}$ is a fixed quantity. Also derivable from this form of the equations is that the precession rate of the orbit plane about the orbit pole is a constant equal to $3nC_{20}/2p^2\hat{\mathbf{h}} \cdot \hat{\mathbf{p}}$.

Now specify this potential in the frame of choice, assuming that the rotation pole of the asteroid, $\hat{\boldsymbol{p}}$, is specified by its declination angle δ and its right ascension α as $\hat{\boldsymbol{p}} = \cos\delta\sin\alpha\hat{\boldsymbol{x}} - \cos\delta\cos\alpha\hat{\boldsymbol{y}} + \sin\delta\hat{\boldsymbol{z}}$, where α is measured 90° in advance of the projection of the rotation pole on the $\hat{\boldsymbol{x}}$–$\hat{\boldsymbol{y}}$ plane. The rotation pole is nominally fixed in inertial space, thus as the asteroid travels about the Sun the right ascension will vary as $\alpha - \nu$. Given this formulation and $\hat{\boldsymbol{h}} = \sin i\sin\Omega\hat{\boldsymbol{x}} - \sin i\cos\Omega\hat{\boldsymbol{y}} + \cos i\hat{\boldsymbol{z}}$ the general potential in this frame becomes:

$$\bar{R}_{20} = \frac{n^2 C_{20}}{4(1-e^2)^{3/2}}\left[1 - 3\left(\sin\delta\cos i + \cos\delta\sin i\cos(\Omega - \alpha)\right)^2\right] \tag{14.6}$$

14.5 Robustness of the Frozen Orbit Solutions

14.5.1 Stability of Relative Equilibria

To study the stability of the solar radiation pressure relative equilibria defined above may seem to be a redundant exercise, given that the general solution for motion in these systems has been defined and is oscillatory. However, it is important to understand how these oscillations manifest themselves, especially when the effect

of joint perturbations between the solar radiation pressure and mass distribution effects are considered. For this stability analysis it is easier to work with the orbital elements themselves, as they do not have the nonlinear constraints on magnitudes, etc., that the angular momentum and eccentricity vectors inherit.

For completeness we consider the two frozen orbit solutions in turn. In both cases the averaged potential is given in the form

$$\bar{R}_{SRP} = -\frac{3aea_{SRP}}{2}\hat{\mathbf{d}}\cdot\hat{\mathbf{e}} \tag{14.7}$$

$$\begin{aligned}\hat{\mathbf{e}} = &[\cos\omega\cos\Omega - \sin\omega\sin\Omega\cos i]\,\hat{\mathbf{x}} \\ &+ [\cos\omega\sin\Omega + \sin\omega\cos\Omega\cos i]\,\hat{\mathbf{y}} + \sin\omega\sin i\hat{\mathbf{z}}\end{aligned} \tag{14.8}$$

and the vector $\hat{\mathbf{e}}$ is stated in reference frame where $\hat{\mathbf{d}} = \hat{\mathbf{x}}$. The independent variable is not shifted to the true anomaly in the current discussion, as this would complicate the inclusion of the asteroid oblateness perturbation later.

As a first step, restate the Lagrange equations for this perturbing potential function

$$\frac{da}{dt} = 0 \tag{14.9}$$

$$\frac{de}{dt} = -\frac{3a_{SRP}}{2na}\sqrt{1-e^2}\,[\sin\omega\cos\Omega + \cos\omega\sin\Omega\cos i] \tag{14.10}$$

$$\frac{di}{dt} = -\frac{3a_{SRP}}{2na}\frac{e}{\sqrt{1-e^2}}\cos\omega\sin\Omega\sin i \tag{14.11}$$

$$\frac{d\omega}{dt} = -\frac{3a_{SRP}}{2na}\frac{1}{e\sqrt{1-e^2}}\left[(1-e^2)\cos\omega\cos\Omega - \sin\omega\sin\Omega\cos i\right] \tag{14.12}$$

$$\frac{d\Omega}{dt} = -\frac{3a_{SRP}}{2na}\frac{e}{\sqrt{1-e^2}}\sin\omega\sin\Omega - \dot{f} \tag{14.13}$$

$$\frac{d\sigma}{dt} = \frac{3a_{SRP}}{2na}\frac{1+e^2}{e}\,[\cos\omega\cos\Omega - \sin\omega\sin\Omega\cos i] \tag{14.14}$$

where this is stated in a rotating reference frame, captured by including the angular rate, $\dot{f}$, of the asteroid about the sun.

14.5.2 Stability of Ecliptic Solutions

In the current orbit frame, ecliptic frozen orbit solutions have an inclination of 0 or π, which necessitates the use of the longitude of periapsis, $\varpi = \omega \pm \Omega$, the sign depending on the inclination. Evaluate the Lagrange equations and this equation assuming $\cos i = \pm 1$

$$\frac{de}{dt} = -\frac{3a_{SRP}}{2na}\sqrt{1-e^2}\sin\varpi \tag{14.15}$$

$$\frac{di}{dt} = 0 \tag{14.16}$$

$$\frac{d\varpi}{dt} = -\frac{3a_{SRP}}{2na}\frac{\sqrt{1-e^2}}{e}\cos\varpi \mp \dot{f} \tag{14.17}$$

$$\frac{d\sigma}{dt} = \frac{3a_{SRP}}{2na}\frac{1+e^2}{e}\cos\varpi \tag{14.18}$$

Given this, the frozen orbit solutions for this case are expressed as

$$\cos(i)\cos(\varpi) = -1 \tag{14.19}$$

$$e = \sin\Lambda \tag{14.20}$$

meaning that $i = 0$, $\varpi = \pi$ is one solution and $i = \pi$, $\varpi = 0$ is another, both with the same eccentricity. It can be easily verified that this solution yields stationary values in all the orbit elements except for the mean anomaly rate.

Now linearize the Lagrange equations about this point to evaluate the equations in the vicinity of this frozen orbit solution, with elements e, i, and ϖ nominally being the free variables. The inclination is nominally zero, however, if no out-of-plane perturbations occur. Thus just focusing on the in-plane perturbation yields

$$\frac{d\delta e}{dt} = -(\pm)_\varpi \frac{3a_{SRP}}{2na}\cos\Lambda\delta\varpi \tag{14.21}$$

$$\frac{d\delta\varpi}{dt} = (\pm)_\varpi \frac{3a_{SRP}}{2na}\frac{1}{\cos\Lambda\sin^2\Lambda}\delta e \tag{14.22}$$

$$(\pm)_\varpi = \cos\varpi \tag{14.23}$$

and the local dynamics can be reduced to a harmonic oscillator in eccentricity

$$\frac{d^2\delta e}{dt^2} = -\left(\frac{3a_{SRP}}{2na}\right)^2 \frac{1}{\sin^2\Lambda}\delta e \tag{14.24}$$

It can be shown that out-of-plane perturbations will yield a similar result.

14.5.3 Stability of Terminator Solutions

The terminator plane frozen orbit solution in the current frame is a bit easier to specify as it does not correspond to a singular set of orbit elements

$$i = \pi/2 \tag{14.25}$$

$$\sin\Omega\sin\omega = -1 \tag{14.26}$$

$$e = \cos\Lambda \tag{14.27}$$

Linearizing the Lagrange equations about this solution in terms of variables e, i, ω, and Ω yields

$$\frac{d\delta e}{dt} = -\frac{3a_{SRP}}{2na}\sin\Lambda\delta\Omega \tag{14.28}$$

$$\frac{d\delta i}{dt} = -\frac{3a_{SRP}}{2na}\cot\Lambda\delta\omega \tag{14.29}$$

$$\frac{d\delta\omega}{dt} = \frac{3a_{SRP}}{2na}\frac{1}{\sin\Lambda\cos\Lambda}\delta i \tag{14.30}$$

$$\frac{d\delta\Omega}{dt} = \frac{3a_{SRP}}{2na}\frac{1}{\sin^3\Lambda}\delta e \tag{14.31}$$

Thus, the system splits into two uncoupled harmonic oscillators again. Expressing these in terms of eccentricity and inclination

$$\frac{d^2\delta e}{dt^2} = -\left(\frac{3a_{SRP}}{2na}\right)^2\frac{1}{\sin^2\Lambda}\delta e \tag{14.32}$$

$$\frac{d^2\delta i}{dt^2} = -\left(\frac{3a_{SRP}}{2na}\right)^2\frac{1}{\sin^2\Lambda}\delta i \tag{14.33}$$

with both having the same oscillation frequency. If true-anomaly is made the independent variable the equations simplify to

$$\delta e'' = -\frac{1}{\cos^2\Lambda}\delta e \tag{14.34}$$

$$\delta i'' = -\frac{1}{\cos^2\Lambda}\delta i \tag{14.35}$$

with the oscillation period being $2\pi\cos\Lambda$, which is to be expected given the general solution found for motion in this system.

Writing out the solutions for the terminator orbits in state-space form in the time domain yields

$$\begin{bmatrix}\delta e\\ \delta\Omega\end{bmatrix} = \begin{bmatrix}\cos(\sigma t) & -\sin^2\Lambda\sin(\sigma t)\\ \frac{1}{\sin^2\Lambda}\sin(\sigma t) & \cos(\sigma t)\end{bmatrix}\begin{bmatrix}\delta e_o\\ \delta\Omega_o\end{bmatrix} \tag{14.36}$$

$$\begin{bmatrix}\delta i\\ \delta\omega\end{bmatrix} = \begin{bmatrix}\cos(\sigma t) & -\cos\Lambda\sin(\sigma t)\\ \frac{1}{\cos\Lambda}\sin(\sigma t) & \cos(\sigma t)\end{bmatrix}\begin{bmatrix}\delta i_o\\ \delta\omega_o\end{bmatrix} \tag{14.37}$$

$$\sigma = \frac{3a_{SRP}}{2na}\frac{1}{\sin\Lambda} \tag{14.38}$$

These solutions are of use later.

14.5.4 Oblateness Perturbations

Now consider the effect of central body oblateness on these frozen orbits. A reasonable assumption is that if the orbit semi-major axis is large, the effect of the asteroid shape may be negligible. When the asteroid is small, however, the maximum semi-major axis becomes small and must lie close to the body, raising the

possibility for a destabilizing interaction between the SRP and oblateness perturbations. This interaction will be evaluated by assuming that the spacecraft lies in a frozen orbit and then including the secular rates from oblateness as a constant perturbation in the Lagrange equations. The resulting system can be solved to find the maximum oscillation amplitude in the orbit element of interest, in this case the eccentricity and inclination.

Of primary interest is the terminator frozen orbits, due to their favorable properties, and thus the effect of oblateness is only studied on this class of frozen orbit. The rate of change of each of the orbit elements due to this perturbation at the frozen orbit conditions, stated previously, are evaluated as

$$\frac{da}{dt} = 0 \tag{14.39}$$

$$\frac{de}{dt} = 0 \tag{14.40}$$

$$\frac{di}{dt} = -\frac{3}{4}\frac{nC_{20}}{a^2 \sin^4 \Lambda} \cos^2 \delta \sin 2\alpha \tag{14.41}$$

$$\frac{d\omega}{dt} = \frac{3}{4}\frac{nC_{20}}{a^2 \sin^4 \Lambda} \left[1 - 3\cos^2 \delta \sin^2 \alpha\right] \tag{14.42}$$

$$\frac{d\Omega}{dt} = (\pm)_\Omega \frac{3}{4}\frac{nC_{20}}{a^2 \sin^4 \Lambda} \sin 2\delta \sin \alpha \tag{14.43}$$

$$(\pm)_\Omega = \sin \Omega|_{\Omega = \pm \pi/2} \tag{14.44}$$

These nominally constant terms can be included on the right-hand side of Eqs. 14.28 to 14.31 to form a set of non-homogeneous equations of the form $\dot{x} = Ax + B$. Assuming that the homogeneous system has a solution $x_H(t) = \Phi(t)x_o$ the full solution is

$$x(t) = \Phi(t)x_o + \Phi(t) \int_0^t \Phi(-\tau)B \, d\tau \tag{14.45}$$

Then, assuming zero initial conditions (i.e., evaluated at the frozen orbit conditions), the general solution for these deviations due to oblateness can be found by quadratures and will have the form $\int_0^t \Phi(t-\tau)d\tau B$. In the following assume that the right ascension α is fixed (although it will really drift in time due to the motion of the asteroid about the Sun).

$$\begin{bmatrix} \delta i \\ \delta \omega \end{bmatrix} = -\frac{n^2 C_{20}}{2aa_{SRP} \sin^3 \Lambda} \begin{bmatrix} \cos \Lambda (1 - 3\cos^2 \delta \sin^2 \alpha)(1 - \cos(\sigma t)) + \cos^2 \delta \sin(2\alpha) \sin(\sigma t) \\ \frac{1}{\cos \Lambda} \cos^2 \delta \sin(2\alpha)(1 - \cos(\sigma t)) - (1 - 3\cos^2 \delta \sin^2 \alpha) \sin(\sigma t) \end{bmatrix} \tag{14.46}$$

and

$$\begin{bmatrix} \delta e \\ \delta \Omega \end{bmatrix} = (\pm)_\Omega \frac{n^2 C_{20}}{2aa_{SRP} \sin^3 \Lambda} \sin(2\delta) \sin \alpha \begin{bmatrix} -\sin^2 \Lambda (1 - \cos(\sigma t)) \\ \sin(\sigma t) \end{bmatrix} \tag{14.47}$$

These results can be used to evaluate the fluctuation amplitude in these orbit elements from their nominal values, especially in eccentricity, in order to determine whether the variation is acceptable. If the amplitude becomes too large the linearization assumptions made will be violated, and a full nonlinear evaluation should be made. The amplitude of oscillation will change as the asteroid moves about the Sun.

The variation in eccentricity from its nominal value, $\cos \Lambda$, is bounded by:

$$|\delta e| < \frac{n^2 |C_{20}| \sin(2\delta) \sin \alpha}{a \sin \Lambda} \frac{B d^2}{(1+\rho) P_o} \tag{14.48}$$

As the asteroid draws closer to the Sun the deviations in eccentricity decrease, and conversely increase as the asteroid travels farther from the Sun. This is due to the contest between the restoring torque of the solar radiation pressure and the perturbing torque of the oblateness on the orbit. When the solar radiation pressure grows weak the oblateness can yield larger variations. The deviation in eccentricity will also decrease as semi-major axis increases, due primarily to the increased distance from the mass distribution. There is a limit on the size of the semi-major axis, which allows us to identify the minimum perturbation due to that term. The maximum effect occurs when the asteroid pole declination is at $\pm\pi/4$, while the predicted fluctuations in eccentricity disappear for declinations of 0 and $\pm\pi/2$. Finally, the fluctuation amplitude will also go through two extrema and two zero periods through an asteroid year, as the right ascension α changes relative to the Sun-line. The two asteroids visited to date, Eros and Itokawa, had pole declinations of 0 and -90 degrees, respectively, and thus should not show a significant eccentricity perturbation on the terminator orbits from the oblateness.

14.5.5 Perturbation from Ellipticity

For the ellipticity effect, or the influence of the C_{22} gravity coefficient, the numerical analysis given in [73] is referenced, where it is empirically determined that orbits outside of 1.5 resonance radii are not as subject to destabilization due to the orbit ellipticity. The resonance radius is the orbit semi-major axis where the orbit period equals the rotation period. Thus, for rapidly rotating bodies one can in principle orbit more closely, while for slowly rotating bodies one must in general maintain a greater distance. The stated 1.5 resonance radius limit does not make any assumption about the orientation of the rotation pole, and is only sharp if the orbit inclination is less than 45 degrees in general. For high inclination orbits, especially for retrograde orbits with inclinations greater than 135 degrees, it becomes possible to orbit much more closely to the body without suffering any effects from the ellipticity. In these situations, however, the oblateness becomes a major perturbation.

Under these provisions the orbit limit to guard against ellipticity effects can be stated as:

$$a > \frac{3}{2} \left(\frac{T^2 \mu}{4\pi^2} \right)^{1/3} \tag{14.49}$$

where T is the rotation period of the asteroid. In general, once the orbit pole of the asteroid is known it is possible to immediately map out when the terminator orbits will have to take special care relative to their interaction with the asteroid gravity field distribution.

To be more specific two plots from [71] are presented in Fig. 14.4 which show the structure of stable and unstable orbits about a uniformly rotating asteroid (in this case neglecting solar radiation pressure). In these plots the parameter $\omega_T = 2\pi/T$, the rotation frequency of the asteroid. The standard relationship between the moments of inertia of a body and its second-degree and -order gravity field are

$$C_{20} = -\frac{1}{2}\left(2I_{zz} - I_{xx} - I_{yy}\right) \tag{14.50}$$

$$C_{22} = \frac{1}{4}\left(I_{yy} - I_{xx}\right) \tag{14.51}$$

A final definition, the parameter γ is defined in [73] as $\gamma = (I_{yy} - I_{xx})/(I_{zz} - I_{xx}) = -4C_{22}/(C_{20} - 2C_{22})$, and is a measure of the body's equatorial ellipticity. The stability plots in Fig. 14.4 are derived by starting a large series of integrations at each point on the plane. If the spacecraft orbit impacts or escapes, a dot is placed and that initial state is unstable. If the orbit persists for some time, several orbits in the [73] paper, then it is considered stable and an "X" is placed. For asteroid Itokawa $\gamma \sim 1$, $(I_{zz} - I_{xx})/r_s^2 = 0.04$ and $r_s = \left(\mu/\omega_T^2\right)^{1/3} \sim 0.6$ km, predicting the onset of instability for a planar orbit at a distance of ∼0.75 km. Figure 14.3

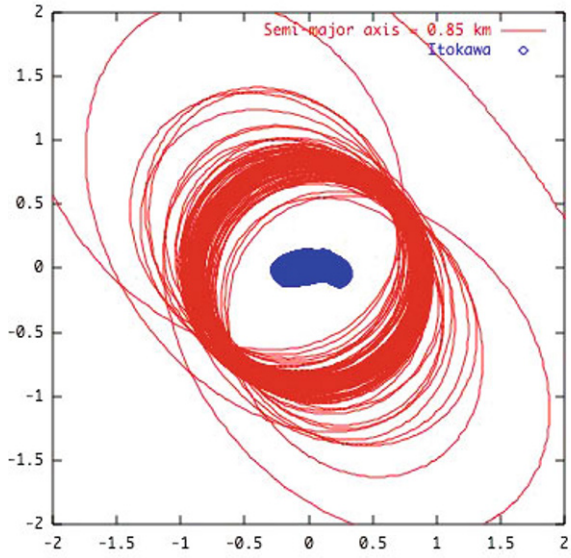


Fig. 14.3 View from the Sun of an orbit started just inside the minimum orbit radius limit. The closer proximity of the satellite to the asteroid's mass distribution destabilizes the orbit, leading to eventual escape.

shows an orbit for the Hayabusa spacecraft at Itokawa started within this lower limit on orbit stability. We clearly see that the previously stable orbit configuration is destabilized by interaction with the body's C_{22} gravity coefficient.

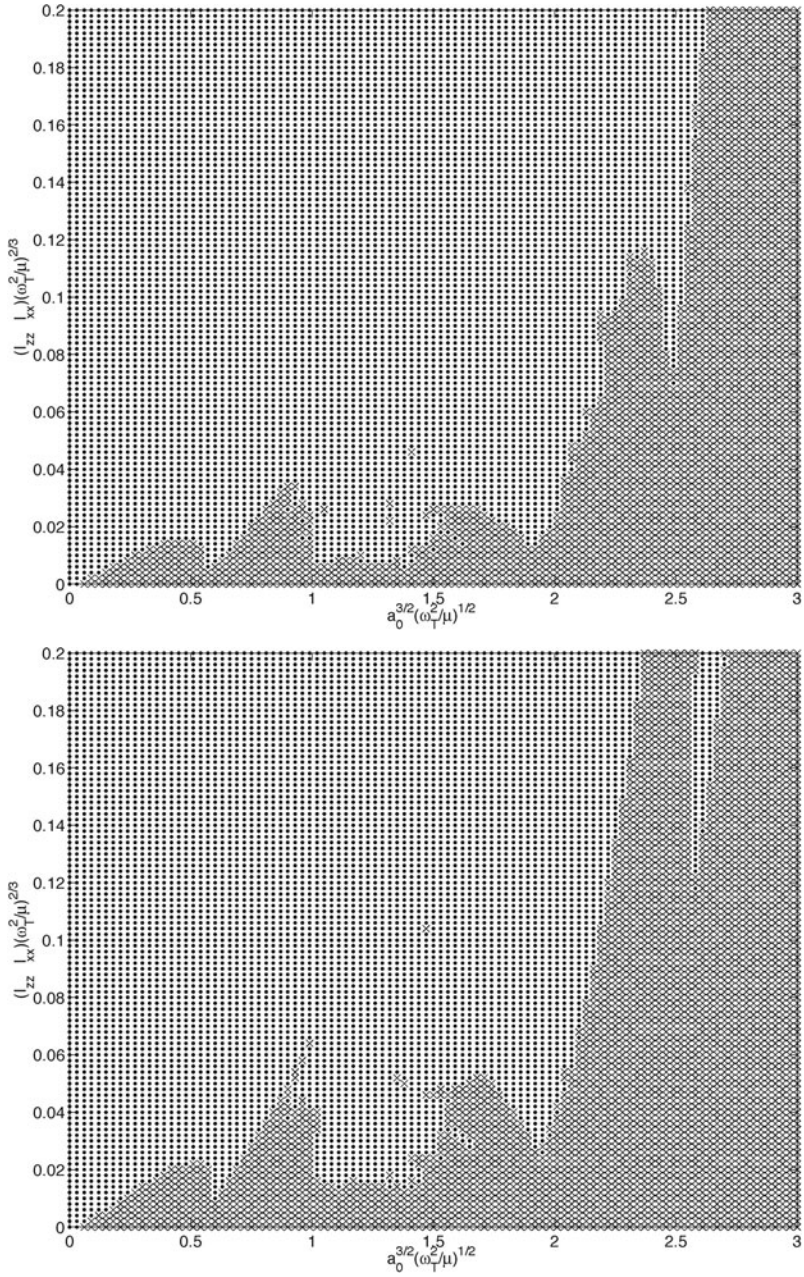


Fig. 14.4 Regions of stable and unstable orbital motion for uniformly rotating asteroids, taken from [71]. The upper plot shows the diagram for $\gamma = 1$, while the lower shows it for $\gamma = 0.5$.

15. Comet Outgassing

This chapter discusses the interaction of a spacecraft orbiter at a comet using simple models of comet outgassing. The current model assumes that outgassing jets provide an impulse to an orbiter that acts mainly in the radial direction, based on the classical model of comet outgassing. It must be noted that other comet outgassing models have been developed more recently [163, 142], some of which involving significant non-radial components. We acknowledge that these models may be more precise, yet do not analyze them in detail here. An initial analysis can be found in [115] which we recommend to the reader. It is significant to note that our conception of comets and their outgassing environment is on the cusp of a major evolution once the Rosetta spacecraft arrives at comet Churyumov-Gerasimenko and characterizes this body during its active phase.

In the following we model the interaction of a spacecraft with two distinct "end members" of cometary models, a continuously outgassing model and one with discrete jets, both acting in a radial direction. For these simple analytical models there are a number of results concerning how a spacecraft responds to this environment. There have been only a relative few analyses of spacecraft with outgassing comets, essentially starting with Scheeres et al. [172], the thesis and related study by Byram [21, 23, 22], and the more recent analyses by Mysen [115]. Once the actual model for the comet outgassing environment is confirmed, the continuation of these studies will be a fruitful area of analysis.

15.1 Continuously Outgassing Fields

Continuous field outgassing at a comet can be described by an acceleration term

$$a_O = \frac{\mu_d}{r^2}\left[1 - \alpha + \alpha\cos\theta\right] \tag{15.1}$$

where μ_d is the effective outgassing parameter, $\alpha \leq 0.5$ is the outgassing field asymmetry parameter and θ is the spacecraft–comet–Sun phase angle. Figure 15.1 shows a cartoon of this model. Alternate models that allow for regions of zero

D.J. Scheeres, *Orbital Motion in Strongly Perturbed Environments*, Springer Praxis Books,
DOI 10.1007/978-3-642-03256-1_15, © Springer-Verlag Berlin Heidelberg 2012

outgassing in the shadowed portion of the comet are described in [21]. This model is symmetric about the Sun-line, essentially assuming a low thermal inertia of the comet surface. A higher value of thermal inertia will displace the maximum outgassing point as a function of the rotation state of the comet relative to the Sun and create an additional asymmetry about the Sun-line.

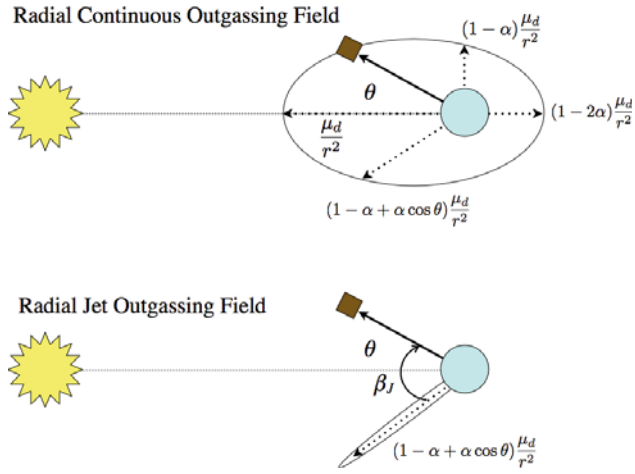


Fig. 15.1 Cartoons showing a simple radial outgassing field for a continuous distribution (top) and a discrete jet (bottom).

The outgassing field and solar radiation pressure should be modeled simultaneously for this case. The solar radiation pressure is stated using the averaged Lagrange equations derived previously and incorporates the outgassing acceleration using the Gauss equations. Averaging these lead to the equations

$$\frac{da}{dt} = \frac{2C_O ae}{1-e^2}\left[\sin\omega\cos\lambda + \cos i\cos\omega\sin\lambda\right] \tag{15.2}$$

$$\frac{de}{dt} = \left[C_O - C_{SRP}\sqrt{1-e^2}\right]\left[\sin\omega\cos\lambda + \cos i\cos\omega\sin\lambda\right] \tag{15.3}$$

$$\frac{di}{dt} = -\frac{C_{SRP}e}{\sqrt{1-e^2}}\sin i\cos\omega\sin\lambda \tag{15.4}$$

$$\frac{d\lambda}{dt} = -\frac{C_{SRP}e}{\sqrt{1-e^2}}\sin\omega\sin\lambda - \dot{f} \tag{15.5}$$

$$\frac{d\omega}{dt} = \frac{1}{e}\left[C_O - C_{SRP}\sqrt{1-e^2}\right]\left[\cos\omega\cos\lambda - \cos i\sin\omega\sin\lambda\right] - \cos i\left(\dot{\lambda}+\dot{f}\right) \tag{15.6}$$

where

$$C_{SRP} = \frac{3a_{SRP}}{2}\sqrt{\frac{a}{\mu}} \quad \text{and} \quad C_O = \frac{\alpha n\mu_d}{2\mu}$$

The semi-major axis is no longer constant, a fact which arises from the averaging process in the Gauss equations.

It is instructive to consider what happens to the previously stable terminator orbits in this model. To test this substitute $i = \pi/2$, $\lambda = \pm\pi/2$ and $\omega = \mp\pi/2$, which orients the orbit according to the frozen orbit solution, to find

$$\frac{da}{dt} = 0 \tag{15.7}$$

$$\frac{de}{dt} = 0 \tag{15.8}$$

$$\frac{di}{dt} = 0 \tag{15.9}$$

$$\frac{d\lambda}{dt} = \dot{f}\left[\tan\Lambda\frac{e}{\sqrt{1-e^2}} - 1\right] \tag{15.10}$$

$$\frac{d\omega}{dt} = 0 \tag{15.11}$$

Thus, for such a symmetric outgassing field the previously found stable terminator orbit is still satisfied as a frozen orbit, with $e = \cos\Lambda$ nulling out the rate in the relative node λ. An entirely analogous result holds for the ecliptic plane frozen orbits discussed in Chapter 13.

This results solely from the symmetry of the above model, and inclusion of an asymmetric model relative to the sub-solar point will destroy these results in general. Still, the terminator environment of a comet is expected to be relatively more benign than orbits that cross jets at low solar phase angles due to the weak solar insolation of the surface.

15.2 Jet Outgassing Fields

A simple generalization of the continuous outgassing model can be made for application to jets. The models discussed in [23] are more sophisticated, yet the following captures the main items of interest (see Fig. 15.1).

$$a_J = \frac{\mu_d}{r^2}\left[1 - \alpha + \alpha\cos\theta\right]\delta(\beta_J) \tag{15.12}$$

where now $\delta(-)$ is the Dirac delta function and β_J is the jet–spacecraft angle. The interpretation of μ_d is now a bit different, and it represents the total pressure force that the jet delivers at a scaled distance of $r = 1$ at a given phase angle from the Sun.

The corresponding Gauss equations can be stated with spacecraft true anomaly as the independent parameter

$$\frac{da}{df} = \frac{2ae\mu_d}{\mu(1-e^2)}\left[1-\alpha+\alpha\cos\theta\right]\sin f\delta(\beta_J) \tag{15.13}$$

$$\frac{de}{df} = \frac{\mu_d}{\mu}\left[1-\alpha+\alpha\cos\theta\right]\sin f\delta(\beta_J) \tag{15.14}$$

$$\frac{d\omega}{df} = -\frac{\mu_d}{e\mu}\left[1-\alpha+\alpha\cos\theta\right]\cos f\delta(\beta_J) \tag{15.15}$$

with only the jet model included.

It can be shown that $dp/df = 0$, which is expected as the jets are modeled as a radial acceleration and thus will not affect the total angular momentum. Despite this, there can still be substantial changes in semi-major axis and eccentricity. Integrating these equations over an encounter yields

$$\Delta a = \frac{2a_o e_o \mu_d}{\mu(1-e_o^2)}\left[1-\alpha+\alpha\cos\theta_o\right]\sin f_o \tag{15.16}$$

$$\Delta e = \frac{\mu_d}{\mu}\left[1-\alpha+\alpha\cos\theta_o\right]\sin f_o \tag{15.17}$$

$$\Delta\omega = -\frac{\mu_d}{e_o\mu}\left[1-\alpha+\alpha\cos\theta_o\right]\cos f_o \tag{15.18}$$

Of specific interest is what happens to the periapsis radius during a jet fly-through, as this determines whether the next orbit has a more intense interaction with the body or not. Given $q = a(1-e)$, considering change in this element yields $\Delta q = (1-e)\Delta a - a\Delta e$, which results in

$$\Delta q = -\frac{a_o\mu_d}{\mu}\left(\frac{1-e_o}{1+e_o}\right)\left[1-\alpha+\alpha\cos\theta_o\right]\sin f_o \tag{15.19}$$

Thus, to ensure positive changes in periapsis radius requires jet flyovers to be situated such that $\sin f_o < 0$, or that the spacecraft should be transitioning from apoapsis towards periapsis. This level of control may not always be feasible, but could be achieved if the active areas of the comet nucleus were mapped out. Whether the mission would allow a spacecraft to fly through an active jet is another issue entirely.

For this model again consider the robustness of a terminator orbit to cometary outgassing jets. For the given model, and what usually occurs for all cometary outgassing models, the strength of outgassing is minimized at the sunrise terminator, and may be reducing again at the sunset terminator. In the stated model this arises from the $\cos\theta$ term which may be zero on the night side of the comet, depending on how α is chosen. Under this consideration, the spacecraft is expected to be more robust to such jets if placed in a terminator orbit.

16. Planetary Satellites: Exact Analysis

This chapter considers orbital dynamics about planetary satellites, a topic distinct from orbital dynamics relative to asteroids or comets. We cover this topic in the remaining chapters as it represents another example of a strongly perturbed orbital system for spacecraft. Europa is chosen as the nominal central body for large planetary satellites due to the great interest of sending an orbiter mission to this body to probe its interior ocean, and due to the fact that its orbital dynamics environment is quite interesting and has been considered in a number of papers [167, 188, 186, 127, 126, 128, 146, 148, 147, 88, 89, 87] and theses [187, 125, 123]. At heart, the theory of large planetary orbiters is the study of dynamics in the Hill problem, previously derived in Chapters 2 and 4. This chapter introduces a number of different analyses that can be applied to this system, and provides the tools for scaling these results to planetary satellites in general.

16.1 Model of Europa

To provide a physical context for the derivation of this model, the Europa environment model parameters are given in Table 16.1. These constants are approximate and are taken from [167] and are for use in the current and following chapter.

Table 16.1 Europa orbital environment models.

Parameter	Symbol	Value	Units
Europa Gravitational Parameter	μ_E	3.201×10^3	km^3/s^2
Jupiter Gravitational Parameter	μ_J	1.267×10^8	km^3/s^2
Mass Ratio	$(\mu_E/\mu_J)^{1/3}$	0.029	–
Europa Radius	R_E	1565	km
Europa Oblateness	J_2	1051.315	km^2
Europa Orbit Period	T_E	3.552	days
Europa Orbit Rate	N_E	2.05×10^{-5}	rad/sec
Orbiter Nominal Altitude	h	200	km
Orbiter Nominal Period	T	2.28	hours
Orbiter Nom. Orbit rate	n	7.66×10^{-4}	rad/sec

D.J. Scheeres, *Orbital Motion in Strongly Perturbed Environments*, Springer Praxis Books,
DOI 10.1007/978-3-642-03256-1_16, © Springer-Verlag Berlin Heidelberg 2012

16.2 Equations of Motion

The physical situation modeled consists of a spacecraft orbiting in the vicinity of a planetary satellite which is in orbit about a larger planet (Fig. 16.1). Formally, this can be written as a restricted three-body problem (assuming the planetary satellite is in a circular orbit). However, this physical situation also has all the elements necessary to apply the Hill approximation. Specifically, it can be assumed that the planetary satellite is in a near-circular orbit about the primary, that the orbiter is orbiting in the vicinity of the planetary satellite, and that the mass of the planetary satellite is sufficiently small as compared to the planet's mass. This last condition is stated specifically as $(M_S/M_P)^{1/3} \ll 1$, and allows us to expand the effect of the planet's gravitational attraction up to second order in this small parameter. For the Europa–Jupiter system this parameter equals ∼0.03, indicating the validity of the smallness assumption.

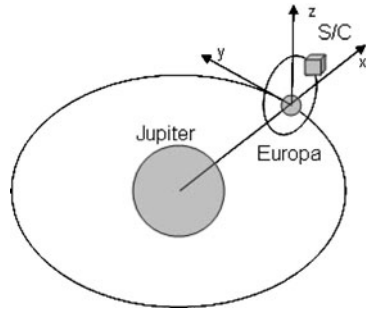


Fig. 16.1 Graphic describing the geometry of a planetary satellite orbiter.

Then, the perturbing effect of the planet's gravity can be stated in potential form as

$$\mathcal{R} = \frac{3}{2}N_S^2 x^2 - \frac{1}{2}N_S^2 r^2 \tag{16.1}$$

where x is measured along the line from the planet to its satellite, r is the orbital radius from the satellite center, and N_S is the orbit angular rate of the planetary satellite about the planet and is approximately $\sqrt{(\mu_P/a_S^3)}$, where μ_P is the planet's gravitational parameter and a_S is the semi-major axis of the planetary satellite's orbit about the planet. See Fig. 16.1 for a picture of the coordinate geometry. The equations of motion of the orbiter can be specified in a Cartesian frame rotating with the planetary satellite about the planet:

$$\begin{aligned} \ddot{x} - 2N_S\dot{y} &= -\frac{\mu_S}{r^3}x + 3N_S^2 x \\ \ddot{y} + 2N_S\dot{x} &= -\frac{\mu_S}{r^3}y \\ \ddot{z} &= -\frac{\mu_S}{r^3}z - N_S^2 z \end{aligned} \tag{16.2}$$

where μ_S is the gravitational parameter of the attracting satellite, x is measured from the planetary satellite center away from the planet, z is measured from the planetary satellite center out of its orbital plane, y completes the triad and points in the direction of the planetary satellite's motion, and $r = \sqrt{x^2 + y^2 + z^2}$.

16.2.1 Normalized Equations of Motion

It is instructive to normalize these equations of motion, as this is the form in which the Hill problem is usually presented, and as this also allows the applications of these results to almost all planetary satellites in the solar system. First define a new time parameter $\tau = N_S t$, so that $d(-)/dt = N_S d(-)/d\tau$. Physically, τ is just the orbit angle that the planetary satellite makes around the planet. The length scale for the Hill problem is different than the one used for the restricted three-body problem, which is the distance between the primaries. Instead, the length scale is the "resonance radius" for the mean motion of the planetary satellite relative to the mass of the satellite, $\left(\mu_S/N_S^2\right)^{1/3} = (M_S/M_P)^{1/3} a_S$. Substitution of this as the length scale then yields the non-dimensional equations of motion

$$\ddot{x} - 2\dot{y} = -\frac{x}{r^3} + 3x \tag{16.3}$$

$$\ddot{y} + 2\dot{x} = -\frac{y}{r^3} \tag{16.4}$$

$$\ddot{z} = -z - \frac{z}{r^3} \tag{16.5}$$

The non-dimensional Hill problem has no parameter, and thus any dynamical result from these equations of motion can be arbitrarily scaled to other systems modeled by the Hill problem. However, not all motion in systems described by the Hill problem is equivalent, as the initial conditions from these other systems must be scaled into or out of the Hill problem. Thus, given an initial condition to a physically defined system modeled by the dimensional Hill problem, the positions must be divided by $\left(\mu_S/N_S^2\right)^{1/3}$ and the velocities by $\left(\mu_S N_S\right)^{1/3}$. The combination of this scaling along with the physical size of a planetary satellite can place very different restrictions on the nature of motion in orbit about a body, or in its proximity.

To appropriately capture this, it is most instructive to provide the normalized radius of a planetary satellite, computed by dividing the body's mean radius by the Hill length scale. The value of the normalized size of a planetary satellite provides a direct indication of the regime of position space in the normalized equations where an initial condition can be specified. In Table 16.2 a partial list of known planetary satellites is presented which fit, or are close to fitting, the physical assumptions needed to apply the Hill approximation. This table lists the main expansion criterion for the Hill problem, that the mass ratio between the satellite and the planet raised to the one-third power is small, the length and time scales for the system, and the satellite radius in the normalized system. The normalized satellite radius proves to be the most important physical parameter in the Hill problem, as it specifies the range of radii for "real" motion in a given system. Some satellites almost

fully fill their Hill sphere, whose extent is delimited by the location of the libration points along the x-axis at $\pm(1/3)^{1/3} \sim \pm 0.69336$. Similar scalings can be applied to all of the planets relative to the Sun, as the physical assumptions to apply the Hill approximation are all valid for these systems. In [116] a list of the normalized planet radii are given and applied for spacecraft transfers in the solar system.

Table 16.2 Select planetary satellite orbit and Hill scaling parameters, including the scaled radius of the body.

Planet	Satellite	Mass Ratio Criterion $\left(\frac{M_S}{M_P}\right)^{1/3}$	Length Scale l $\left(\frac{M_S}{M_P}\right)^{1/3} a_S$ (km)	Time Scale $1/N_S$ (hours)	Normalized Radius R/l
Earth	Moon	0.231[1]	88.704E3	104.85	0.0196
Mars	Phobos	0.0025	23.4	1.22	0.4734
	Deimos	0.0014	32.8	4.81	0.1888
Jupiter	Io	0.036	15.198E3	6.76	0.1199
	Europa	0.0293	19.600E3	13.56	0.0796
	Ganymede	0.0427	45.690E3	27.31	0.0576
	Callisto	0.0384	72.307E3	63.75	0.0333
	Amalthea	0.0015	271.5	1.90	0.3074
	Thebe	0.00073	162.1	2.58	0.3042
	Adrastea	0.00021	27.1	1.14	0.3027
	Metis	0.00037	47.4	1.13	0.4540
Saturn	Mimas	0.0043	797.7	3.60	0.2484
	Enceladus	0.0051	1213.9	5.24	0.2077
	Tethys	0.0109	3211.8	7.21	0.1659
	Dione	0.0123	4642.0	10.45	0.1210
	Rhea	0.0164	8643.5	17.26	0.0884
	Titan	0.0619	75.631E3	60.91	0.0340
	Hyperion	0.0031	4591.4	81.28	0.0294
Uranus	Ariel	0.0249	4756.4	9.64	0.1217
	Umbriel	0.0238	6337.9	15.88	0.0922
	Titania	0.0343	14.951E3	33.25	0.0528
	Oberon	0.0326	19.023E3	51.49	0.0400
	Miranda	0.0092	1190.4	5.38	0.1981
Neptune	Triton	0.0593	21.037E3	22.41	0.0643
Pluto	Charon	0.603[1]	11.819E3	24.09	0.0511

[1]Formally violates mass restriction for the application of Hill's approximation.

16.3 Classical Analysis of the Problem

The "classical" analysis of the Hill problem is first presented, which involves a discussion of the Jacobi integral, zero-velocity curves, and the equilibrium points and the dynamical structure in their vicinity.

16.3.1 Jacobi Integral and Zero-Velocity Curves

The Jacobi integral for the Hill problem is computed as

$$J = \frac{1}{2}v^2 - V(\boldsymbol{r}) \tag{16.6}$$

$$V(\boldsymbol{r}) = \frac{1}{r} + \frac{1}{2}\left(3x^2 - z^2\right) \tag{16.7}$$

where v is the speed relative to the rotating frame and the normalized form of the equations are used. The Jacobi integral can be restated in terms of Keplerian energy and angular momentum

$$J = E - G + \frac{1}{2}\left(r^2 - 3x^2\right) \tag{16.8}$$

where E is the Keplerian energy, G is the z-component of the angular momentum relative to the planetary satellite, and the additional term is just a function of position relative to the planetary satellite.

In the rotating frame it becomes possible to discuss the zero-velocity curves, or specifically those regions where motion is allowed for a give value of J. To delineate these regions rewrite the Jacobi integral in terms of the speed, $\frac{1}{2}v^2 = J + V(\boldsymbol{r}) \geq 0$ and note that this must be strictly zero or positive. Thus, for positive values of J all planar regions of space are accessible to a spacecraft. For negative values of J, however, there are restrictions. First take the case of $J \to -\infty$. For the inequality to hold either $r \to 0$ or $x \to \pm\infty$. Thus the limiting curves are a circular region about the planetary satellite or two vertical lines perpendicular to the x-axis and at an arbitrarily far distance from the origin. Figure 16.2 shows the zero-velocity curves for the Hill problem at a range of energies.

As the value of J increases from this extreme, the circular region around the planetary satellite expands and the vertical lines approach the planetary satellite. As this happens one can show that the vertical lines preferentially distort towards the satellite and that the circular region distorts away from the satellite towards the vertical lines. Due to this the zero-velocity curves first intersect each other along $y = 0$ symmetrically about the origin. At the point of intersection the equation $J = -V(\boldsymbol{r})$ has multiple solutions in its vicinity and thus its linear expansion about this point is zero, or $V_{\boldsymbol{r}} = 0$, yielding an equilibrium point.

As J continues to increase, the zero-velocity curves pull back from the x-axis and begin to recede along the $\pm y$-axis till, at $J = 0$, they disappear. Note that in the z-direction that as long as $J < 0$ that there exists a region, far from the origin, where the $\pm z$ regions are connected to each other through the x–y plane. At $J = 0$ this connection disappears and there are disjoint regions where motion cannot occur along the z regions. For J positive, there always exist regions in the $\pm z$ regions where motion is not allowed, although these regions become arbitrarily far from the $z = 0$ plane as J grows large.

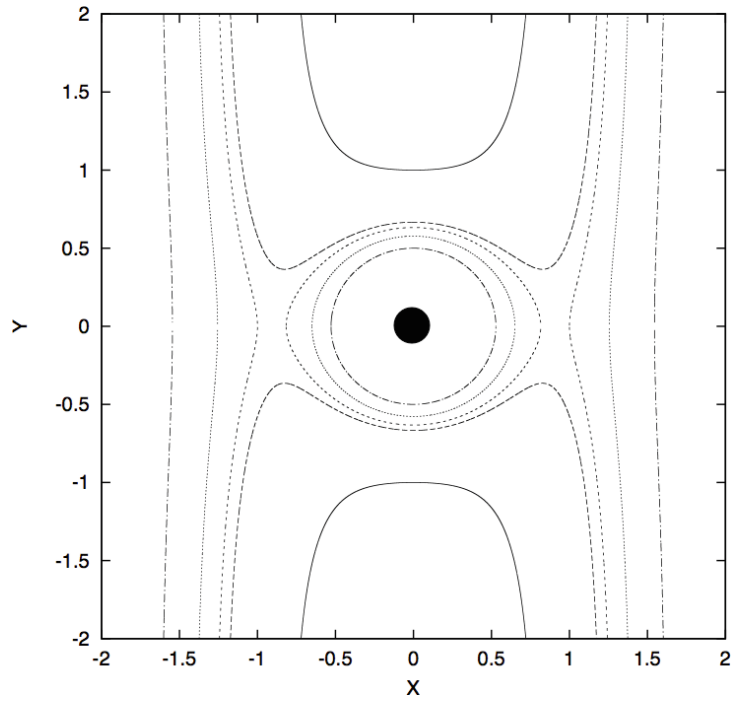


Fig. 16.2 Zero-velocity curves for the Hill three-body problem.

16.3.2 Equilibrium Points

The Hill problem has two equilibrium points defined for it, analogous to the L_1 and L_2 points in the restricted three-body problem, except in the Hill problem these points are exactly symmetric about the body. These occur at the location where the zero-velocity curves intersect, corresponding to the gradient of V being singular, and can be solved for exactly

$$x_{1,2} = \pm \left(\frac{1}{3}\right)^{1/3} \tag{16.9}$$

$$y_{1,2} = z_{1,2} = 0 \tag{16.10}$$

In our derivation convention the planet is off to the left and hence the subscript 1 corresponds to the L_1 point which has the negative sign, and the L_2 point has the positive sign. In the dimensional equations the x coordinates are equal to

$$x_{1,2} = \pm \left(\frac{\mu_S}{3\mu_P}\right)^{1/3} a_S \tag{16.11}$$

Thus they scale linearly with the distance between the planet and the satellite. It is interesting to note that these solutions equal the first-order solutions for the

expansion formula for the location of the $L_{1,2}$ equilibrium points in the restricted three-body problem.

16.3.3 Local Motion about Equilibria

Motion in the vicinity of the equilibrium points is analyzed through a stability analysis. The stability of these points are computed by forming the linearized equations of motion about these points and studying their characteristics. The second-order linear equations about these equilibrium points is

$$\delta\ddot{\boldsymbol{r}} + 2\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\delta\dot{\boldsymbol{r}} = \begin{bmatrix} 9 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -4 \end{bmatrix}\delta\boldsymbol{r} \tag{16.12}$$

with an associated characteristic equation of

$$\left(\lambda^2 + 4\right)\left(\lambda^4 - 2\lambda^2 - 27\right) = 0 \tag{16.13}$$

The out-of-plane eigenvalues are $\lambda = \pm 2i$, where i is the imaginary unit. These represent oscillations that are decoupled with the in-plane motion and yield oscillatory motion with a frequency twice that of the satellite's orbit about the planet, or of half the period. Thus, for an out-of-plane oscillation about the Europa $L_{1,2}$ equilibrium points their period will be on the order of 1.75 days, half the 3.5-day period of Europa about Jupiter. The position eigenvectors of this mode are $\boldsymbol{u}_{\pm z} = [0, 0, 1]$ and the corresponding velocity eigenvectors are $\boldsymbol{u}_{\pm \dot{z}} = [0, 0, \pm 2i]$ and thus consist of pure oscillation along a one-dimensional position manifold in the out-of-plane direction. The full solution for out-of-plane motion is then constructed as $\delta z = \delta z_o \cos(2\tau) + \frac{1}{2}\delta \dot{z}_o \sin(2\tau)$.

The in-plane eigenvalues decouple into a pair of imaginary eigenvalues,

$$\lambda = \pm i\sqrt{\sqrt{28} - 1} \sim \pm 2.071i$$

and a positive and negative real pair

$$\lambda = \pm\sqrt{1 + \sqrt{28}} \sim \pm 2.508$$

The imaginary roots represent a center manifold of periodic orbits that enclose the equilibrium points with frequency $\omega_C = 2.071$ and thus period slightly shorter than the out-of-plane oscillations. The real eigenvalues correspond to stable and unstable one-dimensional manifolds that emanate from the equilibria. The characteristic time of these manifolds is $\tau_c \sim 0.399$ time units. Since 2π time units represent one orbit period of the satellite about the planet, this corresponds to a characteristic time that is $\sim$6% of the orbit period. Thus, for Europa the characteristic time along these manifolds is a short 0.22 days, or 5.3 hours. This is much more rapid than

the corresponding characteristic times at the Earth–Moon $L_{1,2}$ points (1.7 days) or the Sun–Earth $L_{1,2}$ points (22 days).

The position eigenvectors for the in-plane oscillation modes are

$$\boldsymbol{u}_{\pm C} = \begin{bmatrix} 1 \\ \pm i\frac{2\sqrt{\sqrt{28}-1}}{\sqrt{28}-4} \\ 0 \end{bmatrix} \sim \begin{bmatrix} 1 \\ \pm i3.208 \\ 0 \end{bmatrix} \tag{16.14}$$

Resolving this into motion as a function of initial conditions yields

$$\delta x(\tau) = \cos(\omega_C \tau)\delta x_o + \frac{1}{3.208}\sin(\omega_C \tau)\delta y_o \tag{16.15}$$

$$\delta y(\tau) = \cos(\omega_C \tau)\delta y_o - 3.208\sin(\omega_C \tau)\delta x_o \tag{16.16}$$

with the initial velocities found by time differentiation. Thus the motion along these periodic orbits are centered on an ellipse with semi-major axes of 1:3.208 along the x:y-direction, respectively.

The eigenvectors along the unstable and stable manifolds are found to be

$$\boldsymbol{u}_{\pm} = \begin{bmatrix} 1 \\ \mp\frac{\sqrt{\sqrt{28}+1}}{\sqrt{7}+2} \\ 0 \end{bmatrix} \sim \begin{bmatrix} 1 \\ \mp 0.540 \\ 0 \end{bmatrix} \tag{16.17}$$

with the upper sign corresponding to unstable motion and the lower to stable motion. Resolving the motion along these manifolds yields

$$\delta x(\tau) = e^{\pm\tau/\tau_c}\delta x_o \tag{16.18}$$

$$\delta y(\tau) = \mp 0.540\, e^{\pm\tau/\tau_c}\delta x_o \tag{16.19}$$

with the initial velocity again found by time differentiation. The unstable and stable manifolds make an angle of $\mp 28.4^\circ$ with respect to the x-axis direction, respectively.

16.3.4 Periodic Orbits in the Hill Problem

Given that it is simple to specify and has a range of applications, yet remains a non-integrable dynamical problem, the Hill Problem is a natural system to be studied using advanced tools such as are found in celestial mechanics. One of the prime methods of analysis is the study of periodic orbits in such a system. There have been classical studies of the Hill problem using these techniques, most notably the paper by Hénon studying planar families of periodic orbits in the Hill problem [67, 68]. Hénon defines 5 distinct classes of periodic orbits in the planar Hill problem, and studies their stability properties along with their analytic approximation in some specific regimes.

When considering the three-dimensional, or non-planar, Hill problem there are additional families of periodic orbits that can be found and studied. A good

overview of these can be found in [146], where they use the restricted three-body problem, but have mass ratios so small that their results are similar to the Hill problem. As an example, we present a description of halo orbits in the Hill problem. These are orbits that bifurcate from planar Lyapunov orbits, defined above for linear variations about the equilibria, when they grow large. In the following figures the evolution of the halo orbit family is shown as a function of Jacobi integral values from their bifurcation from the planar family up to larger values. These are presented to specifically point out a few points of interest regarding the intersection of periodic orbits with other families and stability transitions due to extrema with respect to the Jacobi integral values. In Fig. 16.3 several halo orbits are shown, selected from the computed family. Note that associated with every motion in the Hill problem, and hence with every periodic orbit, are symmetric orbits reflected about the z-axis and about the origin in the x-axis.

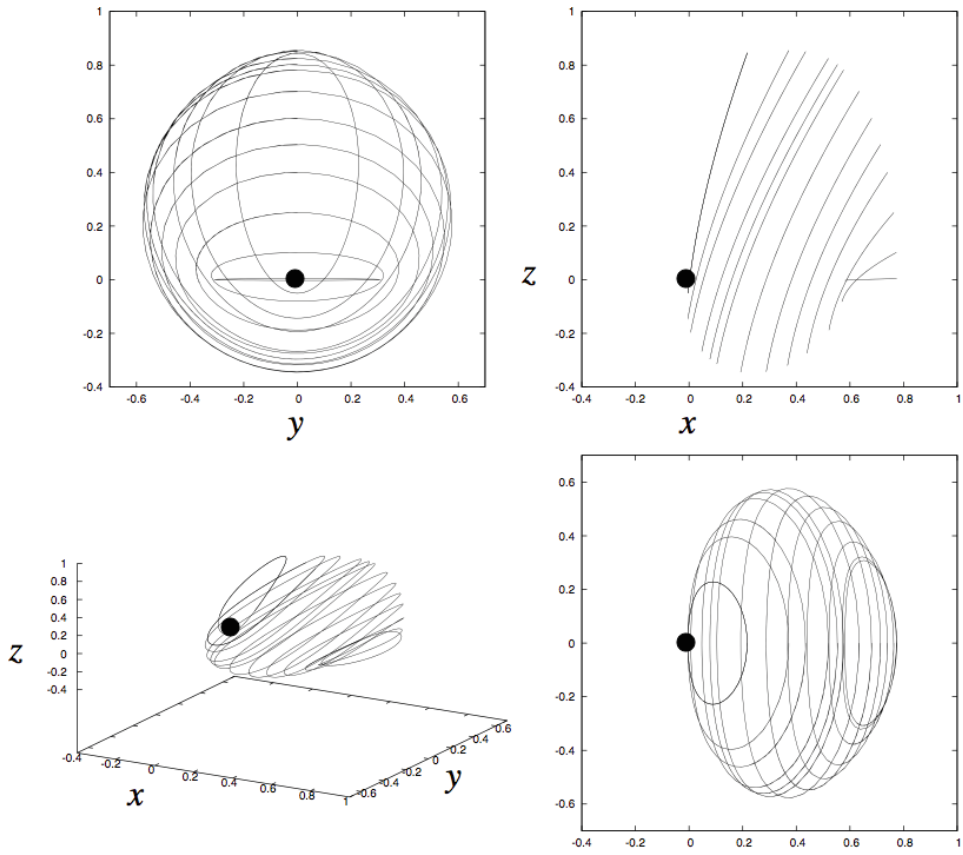


Fig. 16.3 Halo orbits in the Hill problem.

Figures 16.4 and 16.5 present a portion of a halo periodic orbit family and a planar periodic orbit family, from where they intersect to higher values of the Jacobi constant and, for the halo orbit family, through one extremum of that family. The surface of section for the halo orbit family computation is along the $y = 0$ plane with the $\dot{y}$ coordinate used to eliminate the Jacobi integral unity eigenvalue. Thus, to define a halo orbit in this situation requires a state with 4 terms, $x, z, \dot{x}, \dot{z}$ and a Jacobi constant value. Due to the symmetry of the Hill problem and the halo orbits, the $\dot{x}$ and $\dot{z}$ states are both zero when these orbits cross the $y = 0$ plane. The same surface of section technique is used for the planar problem, except now the z components are uniformly zero. Thus, Fig. 16.4 completely defines the different

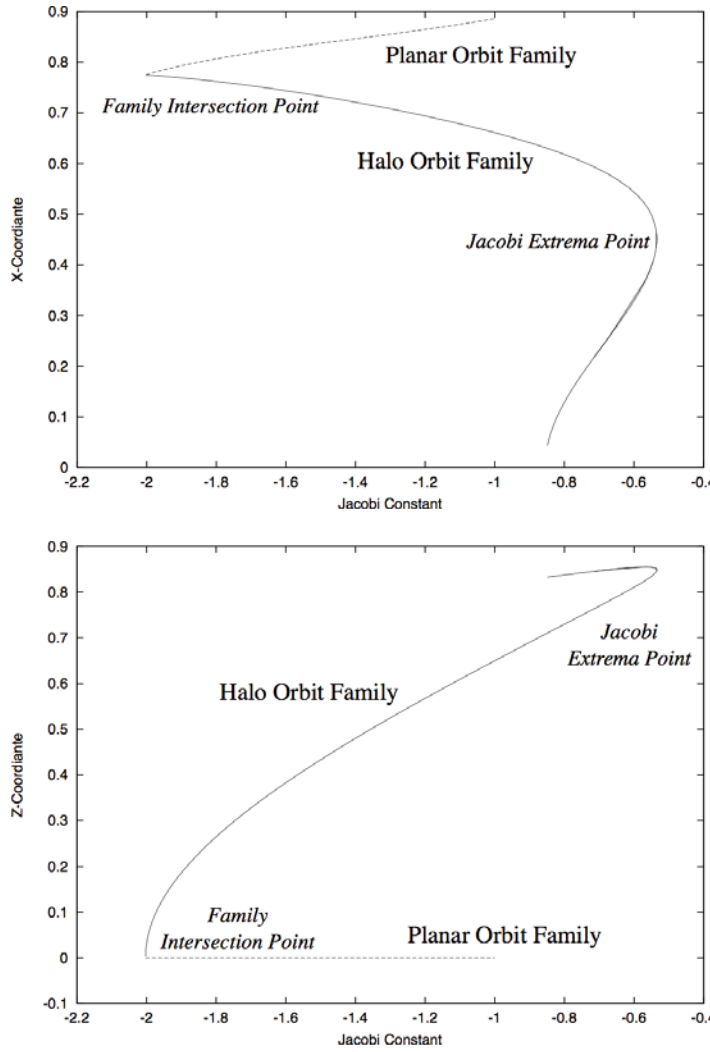


Fig. 16.4 x and z coordinates of the periodic orbit families as a function of Jacobi constant.

members of these periodic orbit families. Two distinctive points along these plots are denoted, where the halo orbit has unity eigenvalues in its monodromy matrix (visible in Fig. 16.5). One is where the two families intersect with each other, and the other where the halo orbit family goes through a local extremum in the Jacobi integral.

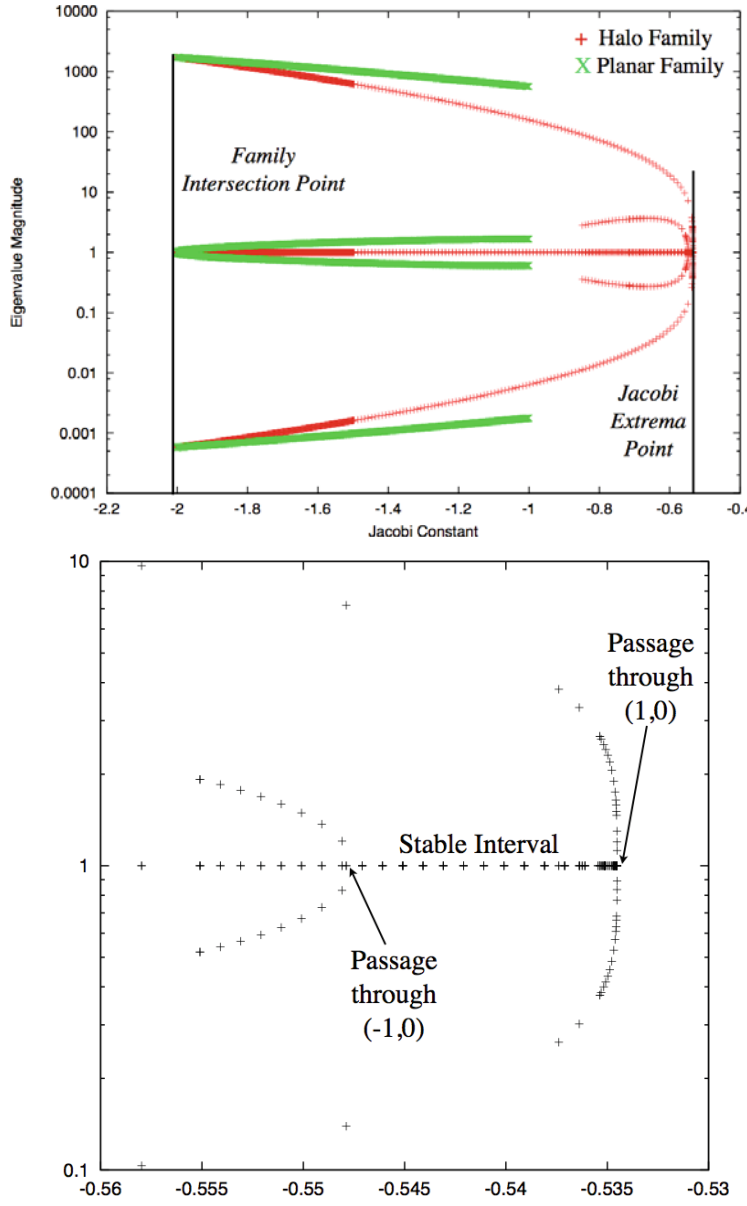


Fig. 16.5 Eigenvalue magnitudes as a function of Jacobi constant. Detail of the halo orbit eigenvalues around the Jacobi value extremum.

At the family intersection the eigenvalues follow the general rules outlined previously in Chapter 5, thus the halo and planar orbits each have one set of unity eigenvalues, with eigenvectors in the out-of-plane direction. Note that the in-plane eigenvalues are strongly unstable at these points; however, the halo orbit retains one set of eigenvalues on the unit circle (hence with unity magnitude). The planar orbits remain unstable in the out-of-plane direction, hence their eigenvalues are only unity at the intersection point and leave along the real axis after that. The halo orbit family has a decrease in their eigenvalue magnitudes, hence becoming less unstable as their Jacobi values increase. As the family reaches its extremum in the Jacobi integral they must converge to $(1, 0)$ in the complex plane and break off onto the unit circle. This leads to an interval of halo orbits which are stable, with all of their eigenvalues on the unit circle. This does not persist as two of these unit eigenvalues migrate over to the negative x-axis and converge to $(-1, 0)$ on the complex plane and become unstable. This instability corresponds to an intersection with a family of orbits of twice the period. Thus, for a short interval after the extremum these orbits are stable but eventually lose their stability through a period doubling bifurcation.

Figure 16.6 plots the periods of these periodic orbits as a function of the Jacobi integral. When families reach an extremum in the Jacobi integral, its value cannot be simply used as a continuation parameter as the families take on two distinct values in the neighborhood of the extrema. The period of the halo orbit family is well defined through this value and thus can still be used as a continuation parameter. To carry this out, one must shift from the Jacobi integral value to period

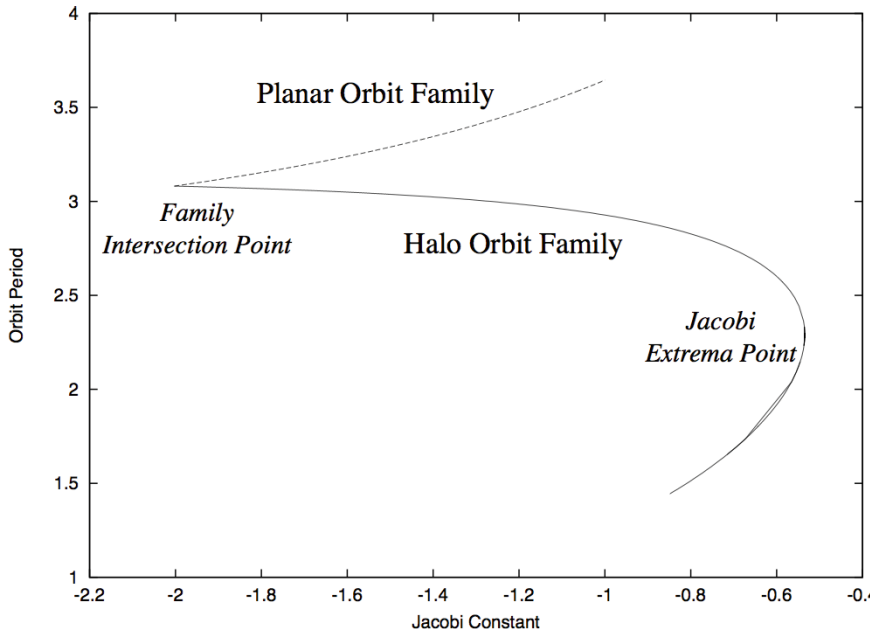


Fig. 16.6 Period of the periodic orbit families as a function of Jacobi constant.

as the independent parameter. In [31] this fact is used to develop an extremely general continuation scheme for periodic orbit families that relies on the intrinsic geometry of the families to trace them out.

16.4 Advanced Dynamics in the Hill Three-Body Problem

The planetary satellite orbiter problem as defined above is of the same cardinality as the circular restricted three-body problem (CR3BP), in that it is a non-integrable problem with the same degrees of freedom. In some ways, due to the absence of any parameter in the reduced form of this problem, it can be more difficult to analyze than the CR3BP as one cannot assume a small parameter to find a nominally integrable problem in the limit. For the H3BP one can only find integrable problems by treating the initial conditions as parameters. This leads to the classical Clohessy–Wiltshire equations for linearized motion about a circular orbit when $|y| \gg 1$ and the two-body problem when $|r| \ll 1$. Another point concerning this problem is that in the full derivation of the H3BP it is not necessary to make the assumption that one of the three bodies is infinitesimally small, and thus these solutions also represent the dynamics of two bodies of arbitrary relative mass with respect to each other, mutually orbiting about a larger body [155]. The current applications still tacitly make the assumption that the orbiter's mass is infinitesimal with respect to the planetary satellite.

Due to the complexity of the H3BP, we do not attempt to provide a complete description of dynamics in this problem. However, in the next subsections two novel approaches to characterizing and understanding dynamics of an orbiter in this problem are described, motivated by the analyses found in [187, 125, 123].

16.4.1 Transit, Capture and Escape

In Chapter 5 Poincaré maps were defined, and in particular the Poincaré periapsis map was introduced as a way to reduce the dimensionality of dynamical systems. In this reduction, only the closest approach of a trajectory to a particular body is tracked, and can be easily represented as the coordinates of the periapsis (technically defined as the point where $\dot{r} = 0$ and $\ddot{r} > 0$). In the planar problem, given the two coordinates and the Jacobi constant for the trajectory the total speed can be computed and, as the point is defined at closest approach, the orientation of the velocity vector is automatically known (perpendicular to the radius) fully defining the state. For the three-dimensional problem it is also necessary to track the rotation of the velocity vector around the radius vector (see [125]). The discussion here will only focus on the planar problem.

The following examples are from the thesis of Marci Paskowitz [125]. The dynamics of sets of trajectories that enter into the zero-velocity region close to the central body are shown by tracking their Poincaré periapsis maps. Figure 16.7 shows a single trajectory that enters the central region and has a number of periapsis passages relative to the body. Of interest is the totality of all trajectories that enter into the central zone at a given energy level and their subsequent orbital evolution. Figure

16.8 shows the first periapsis passage of all trajectories entering the central region from the right. The set that remains close to the entry point are associated with the Lyapunov orbit that exists for the given energy level in the throat region. Of those close to the neck, some will escape at the next iteration while others will continue down into the central region. Figure 16.9 shows multiple iterations of the map, with black being the first passage, blue the second, red the third and green the fourth. There exists a "following set" that comes from the throat region yet lands within the similar groups. Not shown are those trajectories that start near the throat and escape. It is important to note that all of the trajectories that come within this region cannot escape immediately and must follow the different mapping sets.

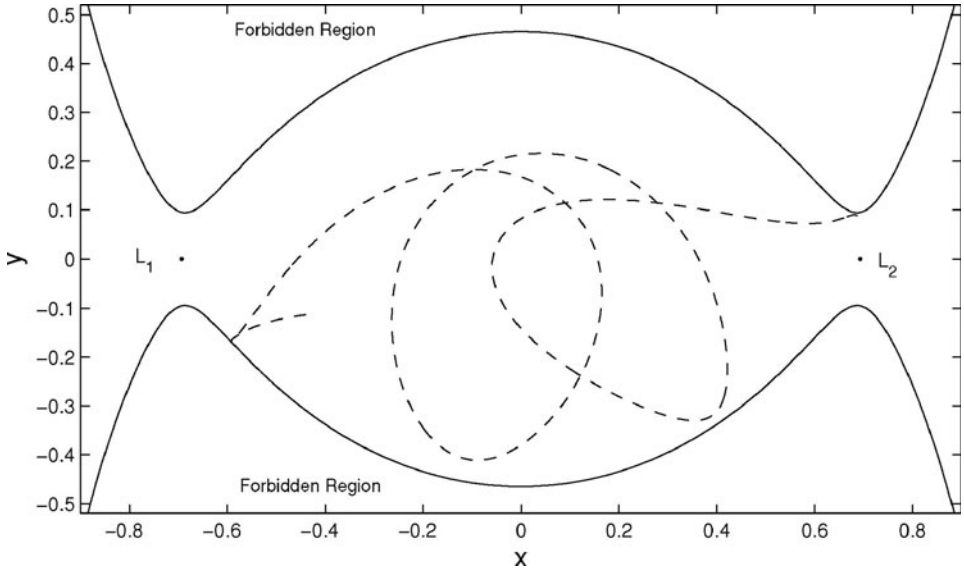


Fig. 16.7 Trajectory close to the central body in the Hill three-body problem. Note that the closest approaches between the particle and the central body are well defined. Credit: [125].

The previous figures are shown for a single Jacobi constant value. As the Jacobi constant is adjusted to larger or smaller values, the extent of the entry region and, correspondingly, the set of the periapsis points, will expand or contract. When the Jacobi constant is chosen to equal the value at the equilibrium points, there is only a single manifold that emanates from that point into the interior region of the zero-velocity curve (forward in time). As the constant value is increased, the entry region expands along with the volume of initial states, shown in Fig. 16.10. Due to the varying Jacobi constant levels, the subsequent trajectory groups will have different dynamics, although they follow the same general pathway around the central body as is evident in Fig. 16.9. There is a peculiarity associated with the Europa system, the lowest Jacobi constant that emanates from the equilibrium point (i.e., the unstable manifold of the equilibrium point) has its first periapsis passage at the surface of Europa. For higher values of the constant a large portion

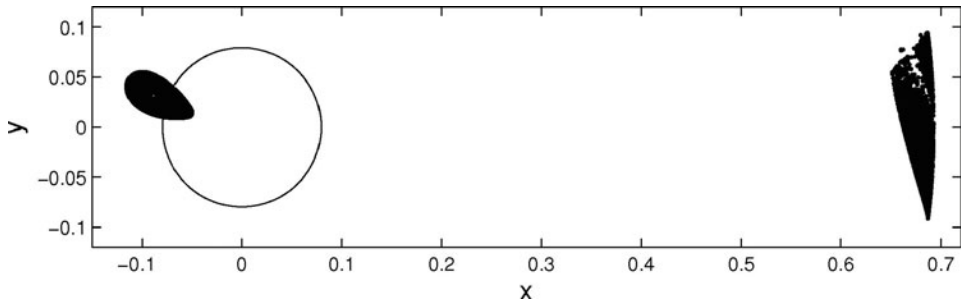


Fig. 16.8 First periapsis passages of all trajectories entering the central zero-velocity region. Most trajectories have their first passage close to the central body, but a significant portion have their first passage near the throat. The central circle represents the radius of Europa. Credit: [125].

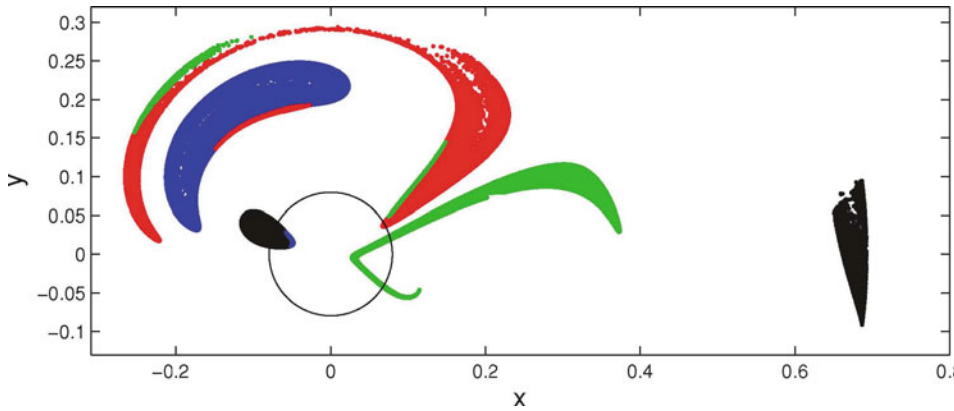


Fig. 16.9 Multiple periapsis passages. Black = first passage, blue = second passage, red = third passage, green = fourth passage. Credit: [125].

of the initial set continues to have their first periapsis passage beneath the asteroid surface.

There are several symmetries in the Hill problem which make it simple to take the current periapsis Poincaré maps and understand other pathways and trajectories through the system. In the planar problem there are two main symmetries. First, changing $(x, y, \dot{x}, \dot{y}, t) \rightarrow (-x, -y, -\dot{x}, -\dot{y}, t)$ transforms solutions across the origin. To prove that the solutions have this property, one needs to show that the equations of motion remain satisfied by the transformation. In Fig. 16.11 this transformation of the initial periapsis set is represented by the set across the central body in Quadrant IV. This set will move forward in time and shadow the sets already presented in Fig. 16.9, and indeed represents the set of all trajectories entering the interior zone from the left.

The other symmetry of interest transforms $(x, y, \dot{x}, \dot{y}, t) \rightarrow (-x, y, \dot{x}, -\dot{y}, -t)$, and in conjunction with the previous transform also yields $(x, y, \dot{x}, \dot{y}, t) \rightarrow (x, -y, -\dot{x}, \dot{y}, -t)$. These solutions go "backwards" in time and represent the solutions that escape, or leave, from the central zone. Given the uniqueness of the periapsis maps, whenever two sets intersect with each other, the trajectories must

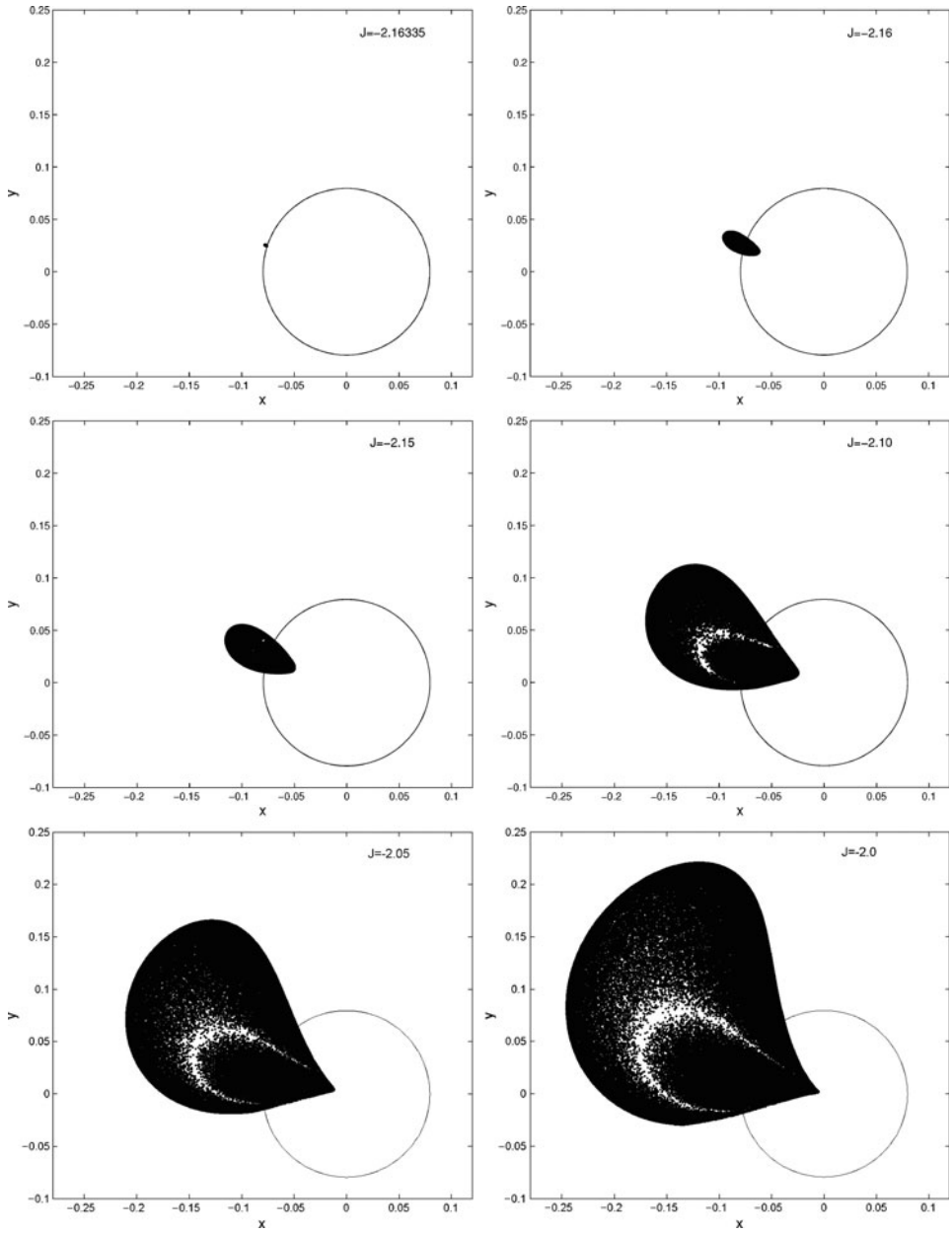


Fig. 16.10 First periapsis passage maps for varying levels of the Jacobi constant. The circle represents the surface of Europa. Credit: [125].

be consistent and follow each other either forward or backwards in time. Thus, when a forward in time set intersects with a backwards in time set, the region of overlap must escape from the system in the requisite number of maps. This can be seen explicitly in Fig. 16.12, where on the second (blue) iteration of the forwards

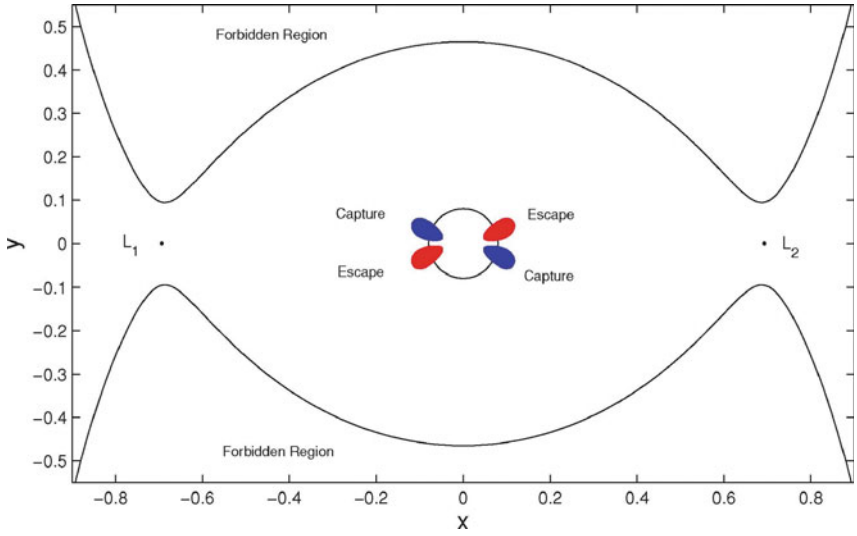


Fig. 16.11 The sets of all first and last periapsis passages for a given Jacobi constant, found by applying the symmetry transformations to the initially computed periapsis set. Credit: [125].

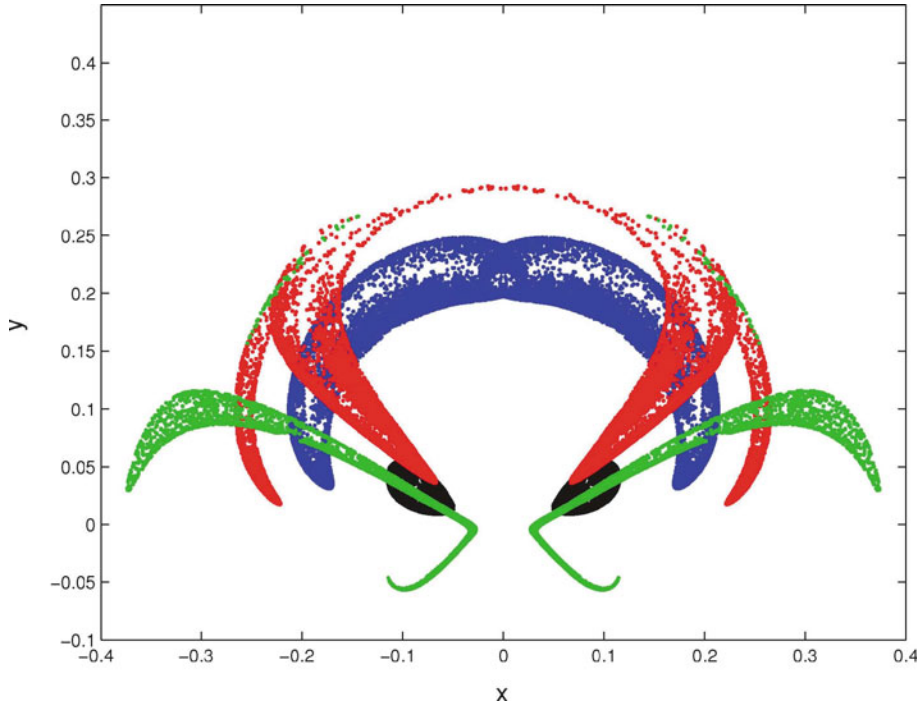


Fig. 16.12 The set of forward periapsis sets (entering from the right) and backwards periapsis sets (exiting on the left) over four iterations. Note that the existence of overlap regions implies that some of the trajectories will transit. Credit: [125].

and backwards periapsis sets that there is a region of overlap. Thus, this region of overlap must then map to a red region (forwards in time) that intersects with a black region (backwards in time), meaning that in the next iteration it will escape from the central region (in this case to the left). Figure 16.13 shows an explicit trajectory from this overlap set.

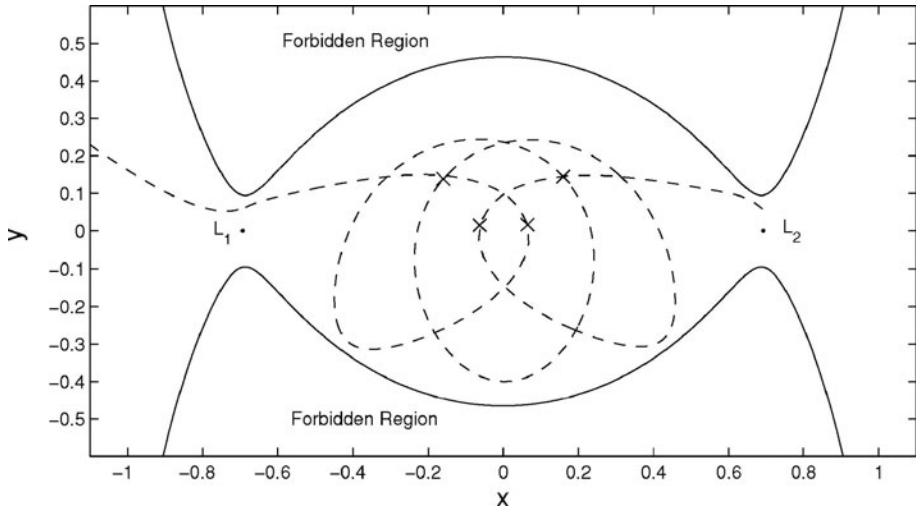


Fig. 16.13 A specific trajectory that enters from the right and exits on the left after 4 periapsis passages. For this Jacobi constant level this is the minimum number of orbits before transit can occur. Credit: [125].

As the periapsis sets define unique points in phase space, they can also be used to define and delineate the final outcomes in a unique way. This is shown in Figure 16.14, where the impacting orbits (red) and escaping orbits (black) are delineated across the entire first periapsis plot region. Continuing the iterations will divide the region into finer and finer delineations, providing targets that are guaranteed to be "safe" over some finite timespan.

16.4.2 Higher-Order Expansion

As the second example for the Hill problem consider the higher-order expansion theory presented earlier in Chapter 6. Writing the normalized Hill equations into the standard form of $\dot{\boldsymbol{x}} = f_{\boldsymbol{x}}(\boldsymbol{x})$ yields

$$\dot{\boldsymbol{x}} = \left[\begin{array}{ccc|ccc} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 3 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \end{array}\right] \boldsymbol{x} + \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ \hline \\ U_{\boldsymbol{r}} \\ \\ \end{array}\right] \tag{16.20}$$

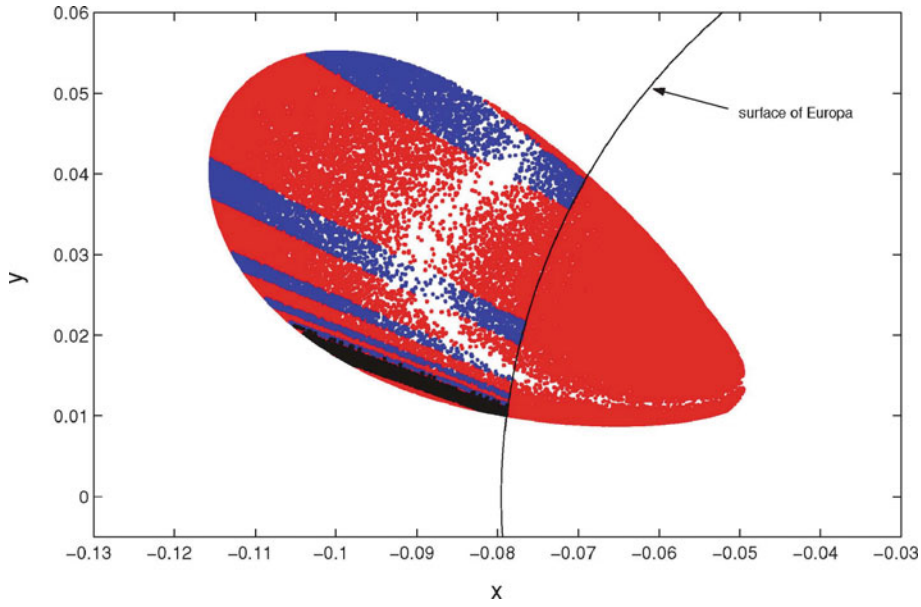


Fig. 16.14 Trajectory outcomes indicated on the first periapsis set. Red denotes impact at some future point, black denotes escape, evaluated over several iterations. Credit: [125].

where $\boldsymbol{x} = [\boldsymbol{r}, \boldsymbol{v}]$ and $U = 1/r$. From this form of the equations, it is simple to note all of the higher-order expansions of the equations of motion

$$f_{\boldsymbol{r},\boldsymbol{r}} = \boldsymbol{0} \tag{16.21}$$

$$f_{\boldsymbol{r},\boldsymbol{v}} = \boldsymbol{I} \tag{16.22}$$

$$f_{\boldsymbol{v},\boldsymbol{r}} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} + U_{\boldsymbol{r}} \tag{16.23}$$

$$f_{\boldsymbol{v},\boldsymbol{v}} = \begin{bmatrix} 0 & 2 & 0 \\ -2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{16.24}$$

Thus, the only higher-order expansions that must be considered are

$$f_{\boldsymbol{v},\boldsymbol{r}\cdots\boldsymbol{r}} = U_{\boldsymbol{r}\boldsymbol{r}\cdots\boldsymbol{r}} \tag{16.25}$$

and all other higher-order partials are identically zero. Thus, the equations of motion for the higher-order state transition tensors have a simplification and do not require all orders of the internal expansions to be computed.

In his thesis and a related series of papers, Park investigated the use of these higher-order expansions for describing motion in the vicinity of a nominally integrated trajectory in the Hill three-body problem [123, 124]. Some of those results are presented here as an example of this method for describing the local dynamics

about a nominal trajectory to higher order. Figure 16.15 shows a nominal trajectory in position and velocity space for a trajectory in close proximity to Europa. The trajectory extends over 6.5 days (almost two Europa orbit periods) and consists of several orbits about the primary, with a strongly changing orbit due to the Jovian perturbations.

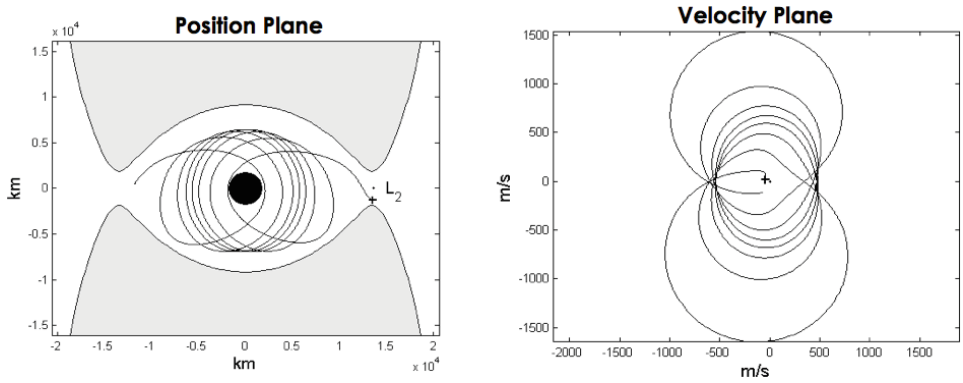


Fig. 16.15 Nominal trajectory used for the computation of higher-order expansions for the theory. Credit: [123].

Figures 16.16 and 16.17 present two computations made with expansions up to fourth order integrated along the nominal trajectory. Both of these examples are computed at 5.74 days, near the end of the trajectory and taken at a point where the higher-order expansions are crucial for capturing the local dynamics accurately. Figure 16.16 shows the projected propagation of a set of initial conditions that consist of a simple circle drawn in position space at a radius of 10 km relative to the nominal solution, and with zero velocity offset. At this later epoch the position extent has grown to hundreds of kilometers and several meters per second in velocity. The different lines on the plot represent different-order contributions to the solution, with the black line representing the numerically integrated solution taken at several hundred points around the initial circle. It can be explicitly seen how the higher order terms of the solution correct the lower-order terms. Classical spacecraft navigation theory relies on linear expansions about a nominal trajectory for describing relative motion (represented here as the blue line). It is significant to note how far off the mark the linear expansion is, whereas even the second-order terms show a much stronger convergence to the true orbital motion. Once the state transition tensors are computed they can be used to generate the relative trajectories of arbitrary initial distributions without recomputing the expansion terms again.

With such an analytic representation of the solution, it is also possible to directly compute the mean, covariance and higher-order statistical moments of the trajectory assuming some initial distribution [124]. The nonlinear mapping of the dynamics allows one to capture the non-zero mean of the true statistical trajectory, and provides a much more robust description of the orbit covariance. Figure 16.17

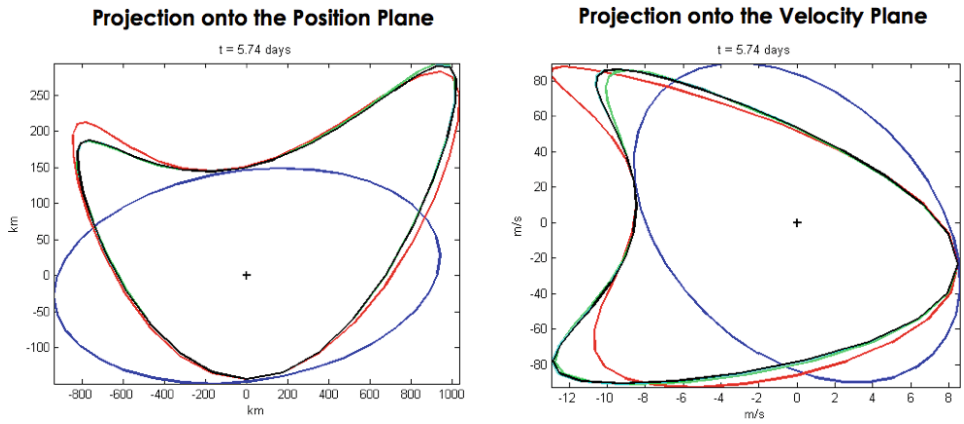


Fig. 16.16 Relative dynamics of a set of initial conditions consisting of a line of radius 10 km drawn about the initial position, with zero deviations in initial velocity. The blue line represents the first-order (linear) expansion, red the second-order, green the third-order, cyan the fourth-order, and black a set of directly integrated points. Credit: [123].

shows the projection of the computed orbit means and covariance ellipsoids relative to the nominal trajectory at the same point in the orbit as chosen for the previous figure. Here the initial distribution is assumed to be Gaussian with a spherical covariance of radius 10 km in position and 0.1 m/s in velocity with no initial correlation. The total extent of the position and velocity covariance is increased by a few orders of magnitude after the mapping, and at this point of the orbit the higher-order expansion terms are starting to exhibit some divergence from the true

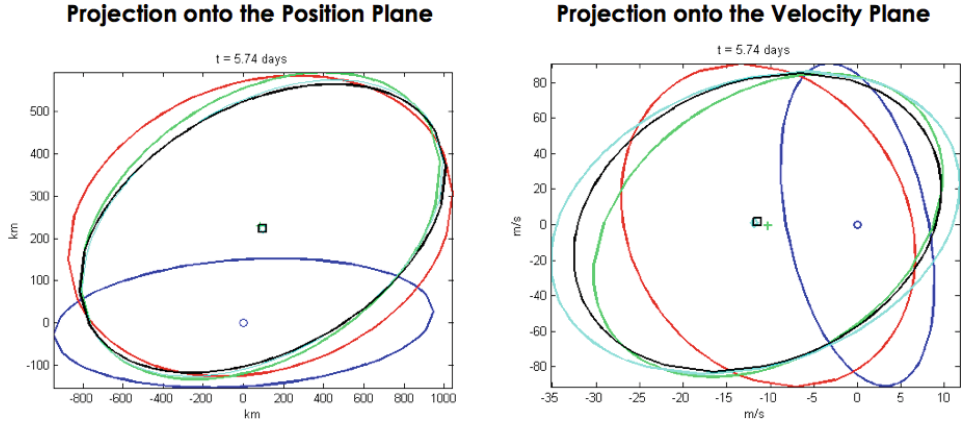


Fig. 16.17 Projections of the mean and covariance of an initially Gaussian distribution consisting of a spherical distribution of radius 10 km in position and a spherical distribution of 0.1 m/s in velocity, with no initial correlation. The blue line represents the first-order (linear) expansion, red the second-order, green the third-order, cyan the fourth-order, and black the mean and covariance computed using a Monte Carlo computation. Credit: [123].

value (earlier computations are highly convergent in general for this example). In this plot the blue is the linear mean and covariance (note that the mean remains at zero, i.e., does not deviate from the nominal), a generic property of linearly mapped Gaussian distributions. This example shows how the actual mean deviates by over 200 km in position and by tens of m/s in velocity from the nominal, and is accurately captured by the higher-order expansions. The black lines correspond to a Monte Carlo computation of the mean and covariance, consisting of thousands of points drawn from the initial distribution and mapped forward in time. Again, the analytical theory can be immediately applied to any initial covariance desired, whereas the numerical result must be re-run for even the smallest changes in the initial distribution.

17. Planetary Satellites: Averaged Analysis

This chapter studies the orbit dynamics of low-altitude, near-circular orbits above an oblate planetary satellite. The mathematical model used in the analysis includes the tidal perturbation from the planet and the effect of the planetary satellite's oblateness and, as such, corresponds to the investigation of a particular "region" of initial condition space of the classical Hill problem with the modification of an oblate central body. The analysis described in the following is largely taken from [167].

17.1 Motivation and Model

The motivation for this analysis is the proposed Europa orbiter mission, which has been contemplated for over two decades as of the writing of this book. The dynamics of low-altitude orbiters about Europa are found to be quite complex where orbits with inclinations within $\sim 50^\circ$ of polar can suffer impacts with that planetary satellite's surface in timespans of days to weeks. This strong instability is actually found for all planetary satellites, with the characteristic time to impact controlled by the orbit period of the planetary satellite and the scaled radius of the planetary satellite defined in the previous chapter. In Figs. 17.1 to 17.3 numerical integrations of a low-altitude Europa orbiter are shown, started at a polar inclination. It can be explicitly seen that the orbital dynamics suffer an instability that would cause the orbiter to impact with the Europa surface within a short time span.

This dynamical instability can be completely explained using an approximate, averaged model for the motion of a satellite about a planetary satellite, and is described herein. The averaging techniques applied in this chapter are relatively simple and of first-order only. Others have performed detailed, higher-order averaging analysis (see [87]), which are of great interest and utility for the mission design and analysis of dynamics in this class of problem. The interested reader wishing to delve deeper into an analytical understanding of the dynamics of planetary orbiters is encouraged to consult these texts.

D.J. Scheeres, *Orbital Motion in Strongly Perturbed Environments*, Springer Praxis Books,
DOI 10.1007/978-3-642-03256-1_17, © Springer-Verlag Berlin Heidelberg 2012

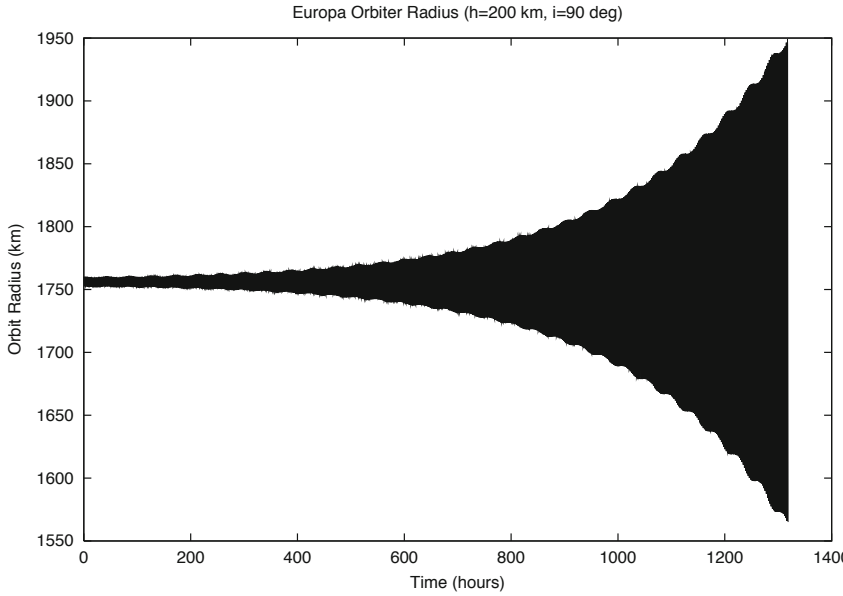


Fig. 17.1 Radius of a Europa orbiter started in a circular orbit at an altitude of 200 km. If uncontrolled, impact occurs in a little over 50 days.

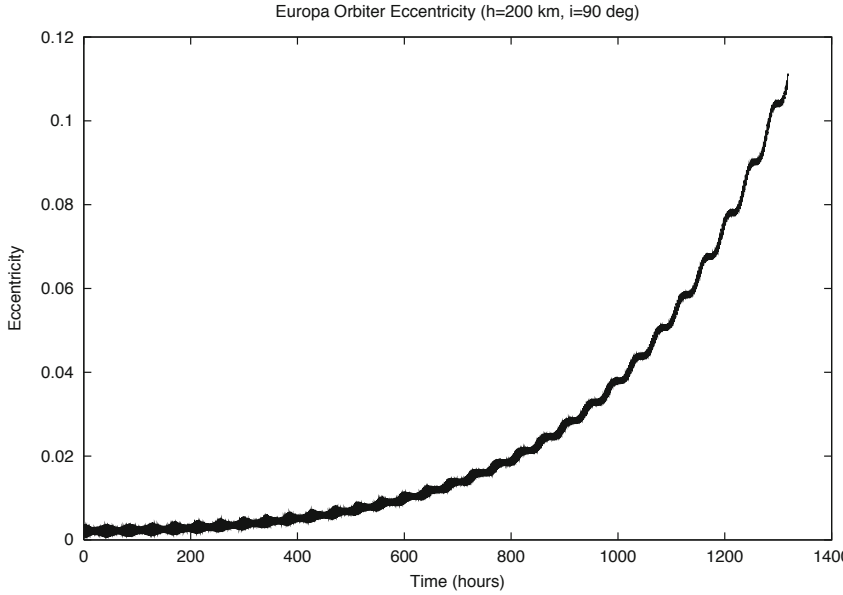


Fig. 17.2 Eccentricity of a Europa orbiter started in a circular orbit at an altitude of 200 km. The eccentricity increases exponentially while the semi-major axis is approximately constant, leading to impact.

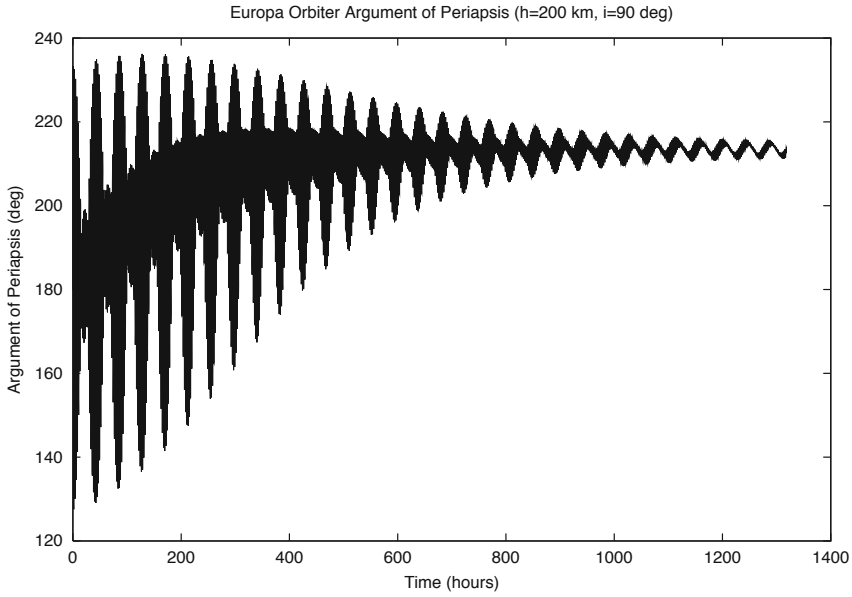


Fig. 17.3 Argument of periapsis of a Europa orbiter started in a circular orbit at an altitude of 200 km. The argument of periapsis is asymptotic to a constant value, defined as the unstable manifold emanating from a circular orbit.

17.2 Dynamics of Planetary Satellite Orbiters

17.2.1 Tidal Perturbation Model

The fundamental equations of motion and the application of the Hill approximation were discussed in Chapter 16. Of primary interest for the current analysis is the tidal perturbation potential, stated as

$$\mathcal{R} = \frac{3}{2}N_S^2 x^2 - \frac{1}{2}N_S^2 r^2 \tag{17.1}$$

where x is measured along the line from the planet to its satellite, r is the orbital radius from the satellite center, and N_S is the orbit angular rate of the planetary satellite about the planet. Figure 16.1 presents a picture of the coordinate geometry.

In terms of inertially referenced orbital elements the coordinate x is:

$$x = r\left[\cos(\omega + f)\cos(\lambda) - \sin(\omega + f)\sin(\lambda)\cos i\right] \tag{17.2}$$

$$\lambda = \Omega - N_S(t - t_o) \tag{17.3}$$

where ω is the argument of periapsis, f is the true anomaly, i is the inclination, Ω is the longitude of the ascending node relative to an inertial frame, t is the time, t_o is an initial epoch, and λ is the longitude of the ascending node relative to the rotating coordinate frame. Using these definitions the perturbing potential can be

expressed in terms of osculating orbital elements alone

$$\mathcal{R} = \frac{N_S^2 r^2}{2}\left\{3\left[\cos(\omega+f)\cos(\lambda) - \sin(\omega+f)\sin(\lambda)\cos i\right]^2 - 1\right\} \tag{17.4}$$

This potential can be used in the Lagrange planetary equations to compute the effect of the tidal acceleration on the osculating orbital elements.

17.2.2 Averaging Assumptions

To simplify the analysis of this problem one can concentrate only on secular changes in the orbital elements. The secular effect of the tidal perturbation can be found by averaging the perturbing potential over the orbit mean anomaly, and then substituting this averaged potential into the Lagrange equations. The perturbing potential expressed in inertial orbital elements has a time variation in the longitude of the ascending node corresponding to the motion of the planetary satellite about the planet. Thus, in addition to the "smallness" assumption of the perturbing acceleration (which is valid here in general) also assume that the orbiter's mean motion about the planetary satellite, $n = \sqrt{\mu_S/a^3}$, where a is the semi-major axis of the spacecraft, is much greater than the planetary satellite's mean motion N_S. This implies that in the time it takes the orbiter to make one revolution about the satellite, the satellite will only have moved a relatively small angle about the planet. The ratio of Europa's mean motion over a low-altitude Europa orbiter's mean motion is $\sim$0.02, implying that Europa will move an angle of only 7° over one period of the orbiter.

Average the potential over one spacecraft orbit, assuming that Ω (or the term $N_S t$) is constant

$$\overline{\mathcal{R}} = \frac{1}{2\pi}\int_0^{2\pi} \mathcal{R}\, dM \tag{17.5}$$

To perform the averaging transform from mean anomaly to true anomaly and write the averaged potential as:

$$\overline{\mathcal{R}} = \frac{1}{2\pi}\frac{a^2(1-e^2)^4 N_S^2}{4\sqrt{1-e^2}}\int_0^{2\pi}\frac{1}{(1+e\cos f)^4}\left\{\left(1-3\sin^2\lambda\sin^2 i\right)\right.$$
$$\left.+3\left(\cos 2\lambda + \sin^2\lambda\sin^2 i\right)\cos 2(f+\omega) - 3\sin 2\lambda\cos i\sin 2(f+\omega)\right\} df \tag{17.6}$$

The expression $1/(1+e\cos f)^4$ can be expanded as a cosine series:

$$\frac{1}{(1+e\cos f)^4} = \sum_{m=0}^{\infty} b_m \cos(mf) \tag{17.7}$$

where closed-form expressions of its coefficients are given in Appendix B. Of specific interest are the coefficients

$$b_o = \frac{\sqrt{1-e^2}}{(1-e^2)^4}\left(1+\frac{3}{2}e^2\right) \tag{17.8}$$

$$b_2 = 5\frac{\sqrt{1-e^2}}{(1-e^2)^4}e^2 \tag{17.9}$$

Using these results the quadrature can be completed to find:

$$\begin{aligned}\overline{\mathcal{R}} = \frac{N_S^2 a^2}{4}\Bigg[&\left(1-\frac{3}{2}\sin^2 i+\frac{3}{2}\cos 2\lambda \sin^2 i\right)\left(1+\frac{3}{2}e^2\right)\\ &+\frac{15}{4}e^2\cos 2\omega\left\{\sin^2 i+\cos 2\lambda\left(1+\cos^2 i\right)\right\}\\ &-5e^2\sin 2\omega\sin 2\lambda\cos i\Bigg]\end{aligned} \tag{17.10}$$

While the averaged potential is now simplified, it still has a time-varying term (λ) due to the motion of the planetary satellite. A further simplification can be made by averaging the potential over one orbit about the planet. For the Europa orbiter problem this second averaging is justified as there are order of magnitude differences between the period of the Europa orbiter and the period of Europa about Jupiter (3.5 days). This second averaging is motivated more by intuitive reasoning than by mathematical argument and in [167] was justified by comparing the results of the following analysis with precision numerical integrations. Performing this second averaging yields a time invariant potential:

$$\widetilde{\mathcal{R}} = \frac{1}{2\pi}\int_0^{2\pi}\bar{\mathcal{R}}d(N_S t) \tag{17.11}$$

$$= \frac{N_S^2 a^2}{4}\left[\left(1-\frac{3}{2}\sin^2 i\right)\left(1+\frac{3}{2}e^2\right)+\frac{15}{4}e^2\cos 2\omega\sin^2 i\right] \tag{17.12}$$

Since the dynamics of a low-altitude orbiter are being considered, the effect of planetary satellite oblateness may also be an important element. The form of this perturbing potential can be averaged over a single orbit to yield

$$\overline{\mathcal{R}}_{J_2} = \frac{\mu_S J_2}{2a^3(1-e^2)^{3/2}}\left(1-\frac{3}{2}\sin^2 i\right) \tag{17.13}$$

where J_2 is the Europa oblateness term, given here in dimensional units of km^2. Due to the linearity of the Lagrange planetary equations with respect to the potential, this perturbing potential can be added to the doubly-averaged tidal perturbation potential to yield the final, secular, perturbing potential.

17.2.3 The Secular Lagrange Equations

These equations are first stated only incorporating the tidal term (Eq. 17.12).

$$\frac{da}{dt} = 0 \tag{17.14}$$

$$\frac{di}{dt} = -\frac{15}{16}\frac{N_S^2}{n}\frac{e^2}{\sqrt{1-e^2}}\sin 2i \sin 2\omega \tag{17.15}$$

$$\frac{d\Omega}{dt} = -\frac{3}{8}\frac{N_S^2}{n}\frac{\cos i}{\sqrt{1-e^2}}\left[2 + 3e^2 - 5e^2\cos 2\omega\right] \tag{17.16}$$

$$\frac{de}{dt} = \frac{15}{8}\frac{N_S^2}{n}e\sqrt{1-e^2}\sin^2 i \sin 2\omega \tag{17.17}$$

$$\frac{d\omega}{dt} = \frac{3}{8}\frac{N_S^2}{n}\frac{1}{\sqrt{1-e^2}}\left[5\cos^2 i - 1 + 5\sin^2 i \cos 2\omega + e^2\left(1 - 5\cos 2\omega\right)\right] \tag{17.18}$$

The semi-major axis is constant, as is the usual case for secular perturbations arising from a potential. The longitude of the ascending node, Ω, is ignorable in that its motion does not affect any other secular element. Similarly, the modification of the mean anomaly is ignorable and is not presented here. Thus, the secular effect of the tidal perturbation creates a coupled evolution of the orbit inclination, eccentricity, and argument of periapsis.

Since only low-altitude orbiters are considered, the analysis can be restricted to small eccentricities. Specifically, assuming that the semi-major axis of the orbit is expressed as $R_S + h$, where R_S is the satellite radius and h is the altitude, then the eccentricity of the orbit is constrained by:

$$e < \frac{h}{R_S + h} \tag{17.19}$$

which is small, given the assumption on h. Neglecting higher orders of eccentricity in the coupled motion of i, e and ω yields

$$\frac{di}{dt} \sim \mathcal{O}(e^2) \tag{17.20}$$

$$\frac{de}{dt} \sim \frac{15}{8}\frac{N_S^2}{n}e\sin^2 i \sin 2\omega + \mathcal{O}(e^2) \tag{17.21}$$

$$\frac{d\omega}{dt} \sim \frac{3}{8}\frac{N_S^2}{n}\left[4 - 5\sin^2 i + 5\sin^2 i \cos 2\omega\right] + \mathcal{O}(e^2) \tag{17.22}$$

Thus, the inclination should be constant on average (to $\mathcal{O}(e)$), reducing this to a two-dimensional problem.

At this point the effect of the planetary satellite oblateness in incorporated. Classical analyses of the J_2 effect on low-altitude orbits show that, on average, the semi-major axis, inclination, and eccentricity will not be affected, and that the longitude of the ascending node and argument of periapsis will have a constant

secular change due to this term. As before, the longitude of the ascending node is ignorable as it has no effect on the remaining elements and is not considered further. The effect of J_2 on the argument of periapsis is represented as:

$$\dot{\omega}_{J_2} = \frac{3nJ_2}{p^2}\left(1 - \frac{5}{4}\sin^2 i\right) \tag{17.23}$$

Including the effect of J_2 on the argument of periapsis and ignoring higher orders of eccentricity yields the final form of the secular Lagrange equations for a Europa orbiter

$$\frac{de}{dt} = \frac{15}{8}\frac{N_S^2}{n} e \sin^2 i \sin 2\omega \tag{17.24}$$

$$\frac{d\omega}{dt} = \frac{3}{8}\frac{N_S^2}{n}\left[4 - 5\sin^2 i + 5\sin^2 i \cos 2\omega\right] + \frac{3nJ_2}{p^2}\left(1 - \frac{5}{4}\sin^2 i\right) \tag{17.25}$$

For notational convenience the equation for the argument of periapsis are rewritten as:

$$\dot{\omega} = \frac{15}{8}\frac{N_S^2}{n}\sin^2 i\left[\cos 2\omega + \alpha\right] \tag{17.26}$$

$$\alpha = (1 + 2\chi)\frac{4 - 5\sin^2 i}{5\sin^2 i} \tag{17.27}$$

$$\chi = \left(\frac{n}{N_S}\right)^2 \frac{J_2}{a^2} \tag{17.28}$$

This form of the equation is valid only for non-equatorial orbits ($i \neq 0, \pi$), but neither of these inclinations will be considered in the following. For a Europa orbiter the parameter $\chi \sim 0.47$ for an altitude of 200 km. Note that the parameter χ is not necessarily small, for example the same quantity evaluated for a low-Earth orbiter considering the tidal perturbation of the Sun yields a value of $\chi \sim 1.9 \times 10^4$; for a geosynchronous orbit it yields a value of $\chi \sim 3$.

17.2.4 Analytical Integration of the Equations

As a linear, time invariant system, Eqn. 17.26 for the argument of periapsis can be solved to find (note that λ is redefined here)

$$\tan\omega = \begin{cases} \sqrt{\dfrac{1+\alpha}{1-\alpha}}\tanh\left(\lambda t + \phi\right) & \alpha^2 < 1 \\ \lambda t + \tan(\omega_o) & \alpha = 1 \\ \sqrt{\dfrac{\alpha+1}{\alpha-1}}\tan\left(\lambda t + \phi\right) & \alpha^2 > 1 \end{cases} \tag{17.29}$$

$$\lambda = \begin{cases} \dfrac{15}{8}\dfrac{N_S^2}{n}\sin^2 i\sqrt{1-\alpha^2} & \alpha^2 < 1 \\ \dfrac{15}{4}\dfrac{N_S^2}{n}\sin^2 i & \alpha = \pm 1 \\ \dfrac{15}{8}\dfrac{N_S^2}{n}\sin^2 i\sqrt{\alpha^2-1} & \alpha^2 > 1 \end{cases} \tag{17.30}$$

$$\phi = \begin{cases} \tanh^{-1}\left[\sqrt{\dfrac{1-\alpha}{1+\alpha}}\tan(\omega_o)\right] & \alpha^2 < 1 \\ \tan^{-1}\left[\sqrt{\dfrac{\alpha-1}{\alpha+1}}\tan(\omega_o)\right] & \alpha^2 > 1 \end{cases} \tag{17.31}$$

Note that the solution for the case $\alpha < 1$ can also be described by replacing all instances of tanh with coth, either one yielding an appropriate solution. This duality exists to allow all possible values of ω to occur. Similarly, for the case when $\alpha = -1$ the functions $\tan(\omega)$ can be replaced with $\cot(\omega)$.

This result clearly shows different behavior as a function of the parameter α. Specifically, for $\alpha^2 < 1$ the argument of periapsis approaches the limiting value:

$$\omega = \tan^{-1}\sqrt{\frac{1+\alpha}{1-\alpha}} \tag{17.32}$$

while for $\alpha^2 > 1$ the argument of periapsis circulates.

Given this solution for ω it is possible to solve for e as well

$$e = \begin{cases} e_o\sqrt{\dfrac{\cosh 2(\lambda t+\phi)-\alpha}{\cosh 2\phi-\alpha}} & \alpha^2 < 1 \\ e_o\sqrt{\dfrac{1+(\tan\omega_o+\lambda t)^2}{1+\tan\omega_o^2}} & \alpha = 1 \\ e_o\sqrt{\dfrac{\alpha-\cos 2(\lambda t+\phi)}{\alpha-\cos 2\phi}} & \alpha^2 > 1 \end{cases} \tag{17.33}$$

Again, there is a change in solution properties as a function of the parameter α. When $\alpha^2 > 1$ the eccentricity will oscillate while when $\alpha^2 < 1$ the eccentricity will grow exponentially. Thus, this explicitly shows that it is possible for eccentricity to evolve in an unstable fashion, which could lead to impact of a low-altitude orbit with the planetary satellite surface. For this case impact with the surface will usually precede eccentricity becoming so large as to violate the smallness assumptions made earlier.

17.2.4.1 Solution in Terms of h and k

Although the derivation of the solution is simpler in terms of orbital elements ω and e, the form and interpretation of the solution is eased in terms of the orbit elements:

$$h = e\sin\omega \tag{17.34}$$
$$k = e\cos\omega \tag{17.35}$$

For $\alpha^2 < 1$ the solution is

$$h = \frac{e_o}{\sqrt{1-\alpha}}\left[\sqrt{1+\alpha}\cos\omega_o\sinh(\lambda t) + \sqrt{1-\alpha}\sin\omega_o\cosh(\lambda t)\right] \tag{17.36}$$
$$k = \frac{e_o}{\sqrt{1+\alpha}}\left[\sqrt{1+\alpha}\cos\omega_o\cosh(\lambda t) + \sqrt{1-\alpha}\sin\omega_o\sinh(\lambda t)\right] \tag{17.37}$$

while for $\alpha^2 > 1$ the solution is

$$h = \frac{e_o}{\sqrt{\alpha-1}}\left[\sqrt{\alpha+1}\cos\omega_o\sin(\lambda t) + \sqrt{\alpha-1}\sin\omega_o\cos(\lambda t)\right] \tag{17.38}$$
$$k = \frac{e_o}{\sqrt{\alpha+1}}\left[\sqrt{\alpha+1}\cos\omega_o\cos(\lambda t) + \sqrt{\alpha-1}\sin\omega_o\sin(\lambda t)\right] \tag{17.39}$$

and for $\alpha = 1$ the system degenerates to

$$h = e_o\left[\sin\omega_o + \cos\omega_o\lambda t\right] \tag{17.40}$$
$$k = e_o\cos\omega_o \tag{17.41}$$

Figure 17.4 shows a sketch of the evolution of eccentricity and argument of periapsis for the main cases. For $\alpha^2 > 1$ circular orbits correspond to stable, fixed points of the secular equations. Conversely, for $\alpha^2 < 1$ circular orbits correspond to unstable, fixed points of the secular equations. Thus the system and solutions can also be viewed in context of the stability of circular orbits at low altitudes.

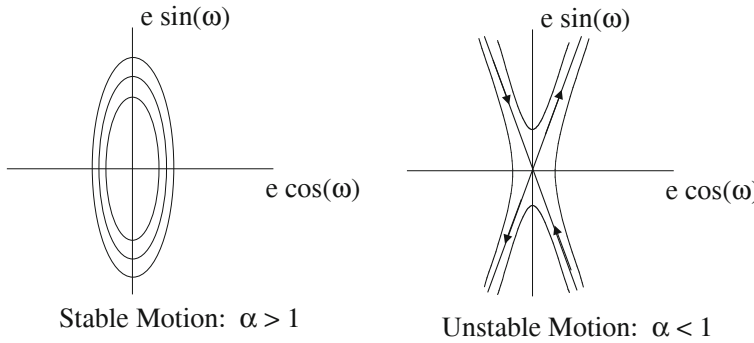


Fig. 17.4 Qualitative sketches of linearized motion relative to the nominal orbit for stable and unstable solutions.

17.2.5 Stability and Time-to-Impact

From these solutions it is possible to explicitly determine the stability conditions for an orbiter, and for the unstable case determine the time-to-impact. The stability condition is obviously $\alpha^2 > 1$ which can be reduced to a set of limits on the inclination of the orbiter:

$$\sin^2 i < \frac{2}{5}\left(\frac{1+2\chi}{1+\chi}\right) \tag{17.42}$$

or

$$\sin^2 i > \frac{2}{5}\left(\frac{1+2\chi}{\chi}\right) \tag{17.43}$$

The χ term is due to the oblateness of the planetary satellite and involves the semi-major axis of the orbit and the mean motion of the planetary satellite about the planet. The second condition cannot occur if $\chi < 2$, meaning that it does not come into play for the Europa orbiter; however, it does come into play for planetary orbiters where (as noted) the values of χ can become large. If the oblateness term is neglected ($\chi \sim 0$) the stability condition becomes independent of physical parameter values and reduces to:

$$\sin^2 i < 2/5 \tag{17.44}$$

which corresponds to inclinations in the regions: $0 \leq i < 39.23$ and $140.77 < i \leq 180$. As χ increases from zero the range of these stability intervals will increase.

For $\chi > 2$ the second condition (Eq. 17.43) becomes active and creates an additional stability region symmetrically placed about a polar orbit ($i = 90°$). If $\chi \gg 1$ then these stability intervals approach each other and leave the stability condition as $\sin^2 i \neq 4/5$. This situation occurs when N_S is extremely small, such as a planet orbiting about the Sun where the period is on the order of years, or when the J_2 perturbation dominates over the tidal perturbation. Thus, the analogous eccentricity instability for a low Earth orbiter would lie in small intervals around the critical inclinations of 63.43° and 116.56°. For a geosynchronous orbit where $\chi \sim 3$ this interval of unstable inclinations would range from 56.8° to 75° and from 105° to 123.2°. It is important to note, however, that the time constant of this instability for an orbiter about the Earth will be extremely long, on the order of 39 years for a geosynchronous satellite. In the remainder of the chapter only the case when $\chi < 2$ is considered, effectively eliminating planetary orbiters from consideration.

The solution for eccentricity can also predict the time-to-impact for an unstable orbiter. Assume a near-circular orbit above the planetary satellite with semi-major axis $a = h + R_S$, where h is the altitude and R_S is the radius of the planetary satellite (or the top of its atmosphere). Since the semi-major axis is constant on average, the condition for impact is that the periapsis radius drop below R_S, or

$a(1-e) \leq R_S$. Solving for the eccentricity at impact yields:

$$e_{imp} = \frac{h}{R_S + h} \tag{17.45}$$

Then, from Eq. 17.33, the corresponding time-to-impact can be explicitly computed as:

$$T_{imp} = \frac{1}{2\lambda}\left\{\cosh^{-1}\left[\alpha + \left(\frac{e_{imp}}{e_o}\right)^2 (\cosh 2\phi - \alpha)\right] - 2\phi\right\} \tag{17.46}$$

This gives an explicit prediction of impact time as a function of initial eccentricity, orbit altitude, inclination, and argument of periapsis. Plotted in Figs. 17.5 and 17.6 are predicted times-to-impact as a function of altitude, eccentricity, inclination and argument of periapsis. These estimates are not necessarily precise, but provide useful relative and qualitative predictions for impact times. The details of impact times are strongly influenced by the short-term oscillations seen in Figs. 17.1 to 17.3.

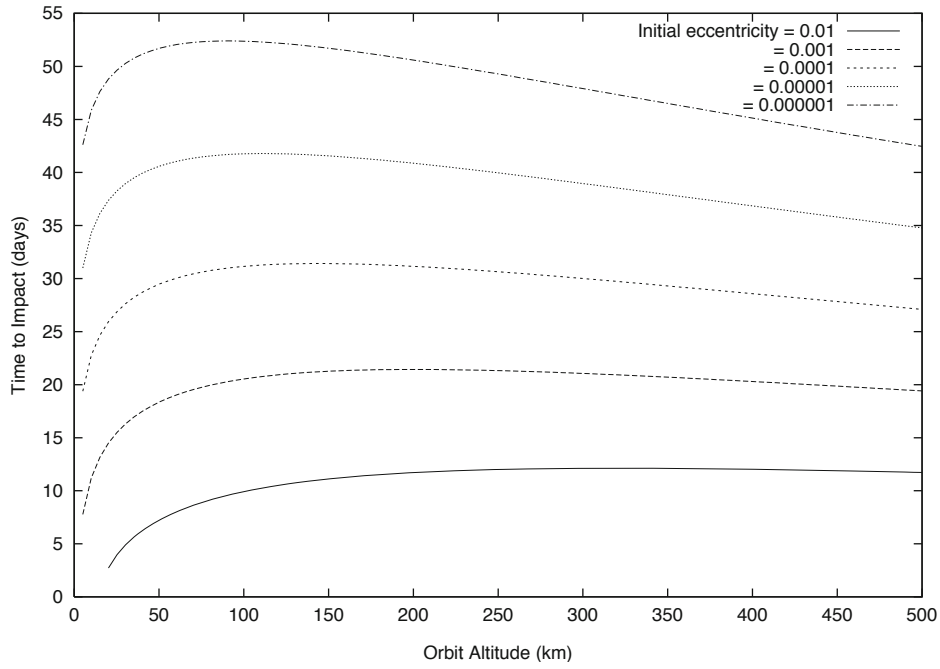


Fig. 17.5 Time-to-impact as a function of altitude at different initial eccentricities. Assumes an inclination of 90° and an initial argument of periapsis of 0°.

An interesting phenomenon can be noted in Eq. 17.33 as, by inspection of Eq. 17.46, it can be deduced that the time-to-impact can be maximized by choosing $\phi \to -\infty$. With this value for an initial condition it will ideally take an infinite length of time for the ratio e/e_o to increase. The condition for $\phi \to -\infty$ is

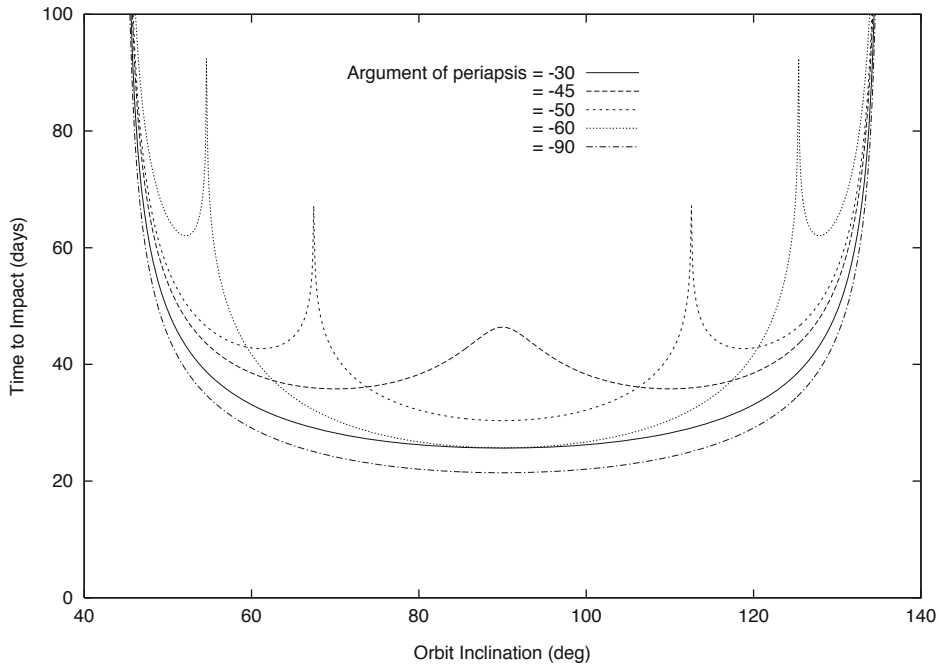


Fig. 17.6 Time to impact as a function of inclination at different initial argument of periapsis. Assumes an altitude of 200 km and an initial eccentricity of 0.001.

$$\tan\omega_o = -\sqrt{\frac{1+\alpha}{1-\alpha}} \tag{17.47}$$

$$= -\tan\omega_{lim} \tag{17.48}$$

$$\omega_o = -\omega_{lim}, \pi - \omega_{lim} \tag{17.49}$$

Thus, choosing the initial argument of periapsis according to this relation can theoretically increase the time-to-impact of the orbiter. Figure 17.7 plots these "optimal" initial arguments of periapsis as a function of inclination for an orbiter about Europa at a 200 km altitude.

In terms of the elements h and k the solution for this special case reduces to:

$$h^* = \mp\frac{e_o}{\sqrt{2}}\sqrt{1+\alpha}\, e^{-\lambda t} \tag{17.50}$$

$$k^* = \pm\frac{e_o}{\sqrt{2}}\sqrt{1-\alpha}\, e^{-\lambda t} \tag{17.51}$$

The argument of periapsis is:

$$\tan\omega^* = h^*/k^*$$

$$= -\sqrt{\frac{1+\alpha}{1-\alpha}}$$

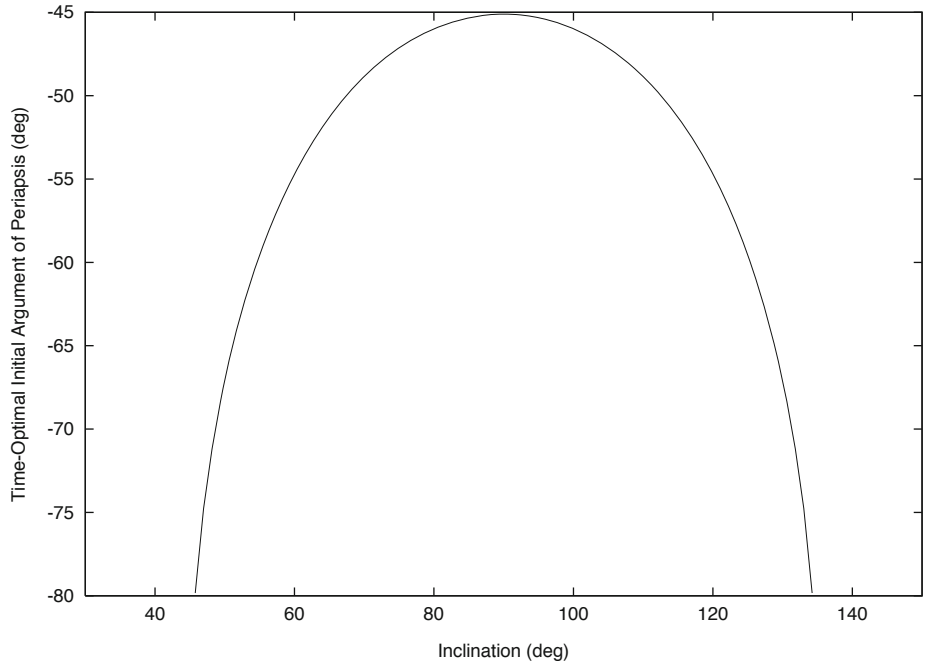


Fig. 17.7 Initial argument of periapsis to prolong the time to impact.

and does not change from its initial value. The eccentricity becomes:

$$\begin{aligned} e^* &= \sqrt{h^{*2} + k^{*2}} \\ &= e_o \, e^{-\lambda t} \end{aligned} \tag{17.52}$$

and asymptotically reduces to zero.

Even though this particular solution is asymptotically stable to a zero eccentricity, the neighborhood of this solution does not have this stability property, and in fact small perturbations will excite the unstable manifold and cause the eccentricity to eventually increase exponentially in time. Thus, while it is possible to use this result to gain additional time before an impact occurs, an uncontrolled orbit will eventually become unstable and impact. Should active control of the orbit be possible, a reasonable approach (based on this solution) would be to occasionally reset the orbit argument of periapsis to a value close to this asymptotic solution. In [126] these manifolds are derived incorporating higher-order gravity field coefficients and used to design longer-life orbits suitable for the design of science orbits about a planetary satellite.

17.3 Extension to Other Planetary Satellite Orbiters

Given the simple form and the universality of the analytic results, it is fitting that they be applied to a larger range of solar system bodies. To perform these comparisons we took a published list of planetary satellites for which mass estimates exist and computed a number of relevant constants for them.

The basic Hill assumptions only require a near-circular orbit for the planetary satellite and a relative mass ratio $(M_S/M_P)^{1/3} \ll 1$. These conditions are satisfied for almost all planetary satellites in the solar system, exceptions being Earth's Moon and Pluto's moon Charon, for which the mass ratios are not small. The averaging assumption made is that the ratio of the planetary satellite orbit mean motion over the spacecraft orbit mean motion is a small number, or $(N_S/n) \ll 1$. To enable comparison between different bodies the spacecraft mean motion is evaluated at the surface of the planetary satellite, designated as $n_S = \sqrt{\mu_S/R_S^3}$. This can be scaled to higher altitudes with the relation:

$$n = n_S \left(1 + h/R_S\right)^{-3/2} \tag{17.53}$$

$$\sim n_S \left(1 - \frac{3h}{2R_S}\right) \tag{17.54}$$

To compute a measure of the importance of this instability for near-polar orbits the characteristic exponent λ (Eq. 17.30) of a polar orbit ($i = 90°$) is computed for each body. For these computations we assume $\chi = 0$ and $i = 90°$ leading to $\alpha = -1/5$, which gives a characteristic exponent of:

$$\lambda = \frac{3\sqrt{6}}{4} \left(\frac{N_S}{n_S}\right) N_S \tag{17.55}$$

$$\sim 1.837\ldots \left(\frac{N_S}{n_S}\right) N_S \tag{17.56}$$

Thus, whenever the condition $(N_S/n_S) \ll 1$ holds, the condition $\lambda/N_S \ll 1$ should also hold – this assumption was explicitly made when the second averaging was performed for elimination of the longitude of the ascending node. Of specific interest is the characteristic time of the eccentricity instability, $\tau = 1/\lambda$. From these relations it can be seen that the characteristic time of the instability is a function only of the planet's mass parameter, the satellite's orbit radius, and the ratio of the satellite's mass parameter with the planet's mass parameter and the ratio of the satellite's physical radius with the satellite's orbit radius. The characteristic time can also be expressed in terms of satellite orbital revolutions about the planet

$$\tau/(2\pi/N_S) = \frac{2}{3\pi\sqrt{6}} \frac{n_S}{N_S} \tag{17.57}$$

and the characteristic time measured in satellite orbit periods is only a function of the ratio of mass parameters and the ratio of the satellite's physical radius with its orbital radius.

Table 17.1 presents the results of these computations for orbiters about a number of planetary satellites of interest. For each planet with satellites, the moons, the averaging conditions (N_S/n_S), the orbit period of the moons ($2\pi/N_S$), and the characteristic times of the instability at the surface of the planetary satellite (τ_S) are listed. When the averaging conditions are violated the modeling assumptions will begin to break down and coupling between the time variations can begin to become important. Included in the table for completeness are some planetary satellites of interest which do not satisfy these conditions. From the table it is apparent that this instability can be a significant concern for most planetary satellite orbiters. Of specific interest is that the instability for a Titan orbiter is relatively mild, on the order of 220 days, meaning that it may only become significant over longer periods of time and that it may be easily controlled by occasional maneuvers. In contrast, an orbiter about Enceladus has a characteristic instability time of $\sim$1.3 days, and thus would most likely require frequent control maneuvers to maintain a given near-polar orbit.

Table 17.1 Summary of select planetary satellite quantities of interest for planetary satellite orbiters [167]. The mean motion ratio is computed for a satellite orbiter at the surface of the planetary satellite. The characteristic instability time is computed for a polar orbit at the surface of the planetary satellite.

Planet	Satellite	Mean Motion ratio (N_S/n_s)	Satellite Orbit Period $2\pi/N_S$ (day)	Characteristic Instability Time τ_S (day)
Earth	Moon	0.00274	27.45	880.19
Mars	Phobos	0.3328[1]	0.32	0.08
	Deimos	0.0806	1.26	1.35
Jupiter	Io	0.04123	1.77	3.72
	Europa	0.0225	3.55	13.66
	Ganymede	0.0138	7.15	44.87
	Callisto	0.00604	16.69	239.11
	Amalthea	0.1931[1]	0.497	0.22
	Thebe	0.1763[1]	0.676	0.33
	Adrastea	0.2089[1]	0.299	0.12
	Metis	0.2762[1]	0.296	0.09
Saturn	Mimas	0.1214[1]	0.943	0.67
	Enceladus	0.0944	1.371	1.26
	Tethys	0.0669	1.888	2.44
	Dione	0.042	2.737	5.64
	Rhea	0.0263	4.518	14.85
	Titan	0.00627	15.946	220.28
	Hyperion	0.00548	21.28	336.19
Uranus	Ariel	0.04225	2.525	5.18
	Umbriel	0.0281	4.157	12.82
	Titania	0.0121	8.705	62.28
	Oberon	0.008	13.48	145.81
	Miranda	0.0893	1.408	1.36
Neptune	Triton	0.0163	5.867	31.20
Pluto	Charon	0.0112	6.307	48.70

[1]Formally violates the frequency condition for application of the second averaging assumption.

18. Small Planetary Satellites: Deimos

We finish with perhaps one of the most challenging dynamical environments that can be encountered. That of a planetary satellite which has a strongly non-spherical shape. The prototypical examples of such satellites are the Martian moons Phobos and Deimos. These bodies are also of significant scientific interest, if for no other reason than their formation and subsequent evolution persists as a significant scientific mystery. These bodies exist in a dynamical environment where both the strong tidal perturbations from the planet and their own non-spheroidal gravity fields combine to place stringent limits on feasible orbital operations at these bodies.

Orbital mechanics about the Martian planetary satellite Phobos has been of considerable interest to researchers for some time, with some of the earliest papers discussing dynamics about strongly non-spheroidal bodies occurring for this body. The earliest study of dynamics about Phobos was given by Dobrovolskis [32], where he pointed out many interesting features of this system relevant to understanding its current state and presumed natural evolution. Wiesel published the first detailed study of spacecraft orbital dynamics in the vicinity of Phobos [196]. More recently, the Russian Phobos-Grunt mission [98], which had a failed orbit insertion after its launch in November 2011 planned to visit the Martian moon Phobos and place instruments on its surface for its detailed scientific study. Associated with these plans were studies of the orbital mechanics about this specific system [176, 2].

The satellite Deimos has also been of interest to space scientists, and has even been proposed as a potential way-station for astronauts leading up to a visit to the surface of Mars. In this chapter we provide a discussion of the peculiar dynamics encountered at the Martian moon Deimos. While the environment of this moon is not as strongly perturbed as that of Phobos, the same essential dynamics are encountered at both bodies. Given that orbital dynamics about Deimos has not been explicitly studied in the published literature, it is fitting to focus the final analysis of this book on this particular body. The following work was largely performed in support of earlier NASA Discovery mission proposals to this specific body. While none of these missions have been supported for further development

D.J. Scheeres, *Orbital Motion in Strongly Perturbed Environments*, Springer Praxis Books,
DOI 10.1007/978-3-642-03256-1_18, © Springer-Verlag Berlin Heidelberg 2012

to date, missions to Deimos remain a perennial contender for space science mission proposals.

18.1 Model of Deimos

The main parameters used for the current analysis are the shape model of Deimos, its gravitational parameter and its spin period. Deimos is locked in synchronous rotation with its orbit about Mars, and thus its rotation period and orbit period are equal. Due to this we assume it has a zero obliquity angle with respect to its orbit plane and that it rotates about its maximum moment of inertia. In Table 18.1 we present the basic information needed for this study. The shape model is available at the PDS-SBN [117]. The gravity field is modeled using the Deimos shape model with a uniform density assumption to provide the total gravitational parameter listed in Table 18.1. When sufficiently far from the surface of the asteroid, nominally outside of 10 km from its center of mass, a 16×16 gravity field is used.

Table 18.1 Deimos parameters.

Parameter	Symbol	Value	Units
Gravitational Parameter	μ	1.354×10^{-4}	km^3/s^2
Rotation Period	T	32.2986	hours
Rotation Rate	ϖ	5.404×10^{-5}	rad/s
Mean Radius	R_o	6.234	km
Oblateness Gravity Coefficient	$R_o^2 C_{20}$	-4.208	km^2
Ellipticity Gravity Coefficient	$R_o^2 C_{22}$	1.251	km^2

18.2 Equations of Motion

The equations of motion of a particle about Deimos are modeled by the Hill model developed in Chapter 16 with the addition of a general gravity field instead of a point potential. The dimensional version of the equations are exclusively used for the current analysis. In the following the x-axis is aligned with the minimum moment of inertia of Deimos and is pointed towards Mars, the z-axis is aligned with the maximum moment of inertia and is normal to the orbit plane, and the y-axis completes the triad. The equations of motion in scalar form are then

$$\ddot{x} - 2\varpi\dot{y} = 3\varpi^2 x + U_x \tag{18.1}$$

$$\ddot{y} + 2\varpi\dot{x} = U_y \tag{18.2}$$

$$\ddot{z} = -\varpi^2 z + U_z \tag{18.3}$$

where ϖ is the rotation rate and $U(x, y, z)$ is the gravitational potential.

We consider one approximation to this model when discussing retrograde periodic orbits, related to the neglect of the attraction of Deimos. Specifically, when far from Deimos the gravitational attraction quickly becomes negligible relative to the tidal accelerations. For a specific example, consider displacing a particle along the x-axis away from Deimos. The tidal and gravitational accelerations will balance at a distance $\left(\mu/3\varpi^2\right)^{1/3} \sim 25$ km. As the distance of the particle from Deimos is doubled to ∼50 km, the tidal acceleration doubles and the gravitational acceleration decreases by a factor of 4, yielding a relative change between the two accelerations of a factor of 8. A further doubling to 100 km makes the tidal acceleration a factor of 64 larger than the gravitational acceleration, etc. Thus, as one considers orbits far from Deimos, on the order of hundreds of kilometers, the gravitational attraction of Deimos becomes negligible and can be ignored. Applying the approximation ($U_x \sim 0$, etc.) yields a simplified set of equations

$$\ddot{x} - 2\varpi\dot{y} = 3\varpi^2 x \tag{18.4}$$

$$\ddot{y} + 2\varpi\dot{x} = 0 \tag{18.5}$$

$$\ddot{z} = -\varpi^2 z \tag{18.6}$$

which are sometimes called the Hill equations (not to be confused with the Hill problem derived previously in this book) or the Clohessy–Wiltshire equations, especially when applied to satellites. These equations, as given above, represent the linearized motion of a particle relative to a circular orbit, written in the frame rotating with the nominal circular orbit about the planet. As these are time invariant, linear equations they can be solved in closed form.

18.3 Dynamics about Deimos

In the following several different aspects of orbital motion in the Deimos system are considered. All of the following is based on numerical evaluation of algebraic equations and numerical integration of the dynamical equations of motion, except as noted.

18.3.1 Zero-Velocity Curves and Equilibrium Points

As the equations of motion are time invariant a Jacobi integral exists for this system. A traditional way to derive this integral directly from the equations of motion is to multiply them by $\dot{x}$, $\dot{y}$ and $\dot{z}$, respectively, and add them. The Coriolis acceleration terms cancel and the resulting quantity can be reduced to an exact differential. Integrating this then yields the Jacobi integral

$$J = \frac{1}{2}\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) - \frac{1}{2}\varpi^2\left(3x^2 - z^2\right) - U(x, y, z) \tag{18.7}$$

As has been discussed earlier, this can be used to define zero-velocity surfaces that separate regions of allowable motion. For a specified value of $J = C$ the constraint is

$$C + \frac{1}{2}\varpi^2\left(3x^2 - z^2\right) + U(x, y, z) \geq 0 \tag{18.8}$$

In Fig. 18.1 the zero-velocity curves for Deimos are presented along the $z = 0$ plane. Note that they are similar to the zero-velocity surfaces in the Hill problem (Fig. 16.2), although they have some asymmetry due to the Deimos mass distribution.

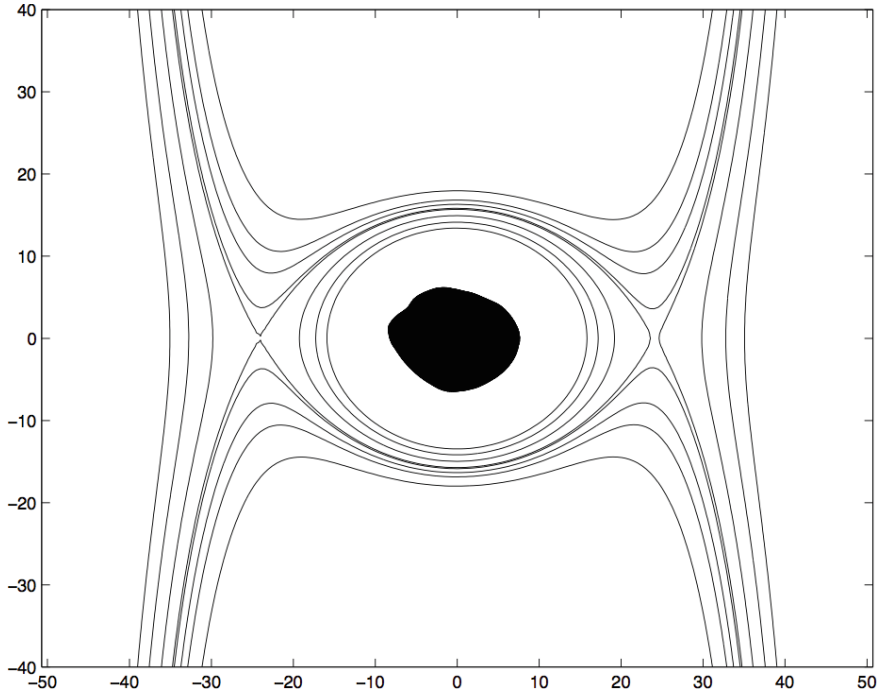


Fig. 18.1 Zero-velocity curves for Deimos along the $z = 0$ plane.

Evident in the figure are the equilibrium points along the x-axis, analogous to the equilibrium points in the Hill problem. It is relevant to point out that Deimos has no equilibrium points along the y-axis, unlike its uniformly rotating counterpart Eros. The tidal potential of Mars provides a sufficiently large perturbation so that it effectively destroys these equilibrium points. While this is clear in the Hill problem derivation, it is still a bit surprising when applied to a uniformly rotating general shape. The two Deimos equilibrium points share the same stability properties as those in the Hill problem, and given our complete discussion in Chapter 16 we do not consider them in detail here.

18.3.2 General Trajectories

To start our more detailed discussion of orbits about Deimos, we first point out the strongly unstable nature of motion at this body. The location of the equilibrium points, and hence the Hill radius for Deimos, is approximately 25 km. From numerical and analytical studies carried out in the restricted three-body problem it has been shown that direct orbits outside of half the Hill radius have a strong tendency to escape [61]. Thus, in the following initially circular orbits of radius 10–15 km were considered.

Starting these orbits with an inclination ranging between 0 and 90 degrees, they generally will escape or impact with Deimos in less than a single orbit period. A detailed analysis of these dynamics would require a combination of techniques, including the discrete maps developed for the analysis of Eros in Chapter 7 and the conditions for escape found for the Hill problem in Chapter 12. As there is seemingly little hope for developing long-term direct stable orbits about this body, these topics are not pursued further. As an example, Fig. 18.2 shows two different initially circular, polar orbits about Deimos. The only difference between the two is that one has its line of nodes along the y-axis and escapes from the body after a few orbits, and the other has its line of nodes along the x-axis and impacts.

As inclination is further increased, regions of orbital stability are found. In general, if the orbital inclination maintains a sustained value above ∼140 degrees, then an orbit can persist for long periods of time. If the inclination is lower than this, however, the eccentricity tends to grow and impact usually ensues. It is significant to note that this inclination limit is consistent with the averaged analysis in Chapter 17 for when the eccentricity becomes unstable ($i > 140^\circ$, specifically). Retrograde orbiters take advantage of having their orbital motion go against the rotation of the mass distribution, and thus can maintain stability if not destabilized by the tide. Figure 18.3 plots two orbits with an initial inclination of 135 degrees. One has its node situated so that the inclination increases to a higher value, while the other's inclination decreases to a lower value and is seen to impact rapidly. As orbital inclinations are moved to higher values, including retrograde, these initially circular orbits are stable in general, and are candidates for sustained spacecraft orbits.

18.3.3 Periodic Orbits

Moving from general initial conditions for orbits, it is instructive to consider periodic orbits about Deimos. In the following we present some example direct orbits, all unstable, that exhibit some interesting geometry. Following these examples, a discussion and analysis of retrograde periodic orbits is provided, as these are of most interest for space science missions to such bodies.

The following periodic orbits about Deimos were computed using the computational algorithm for periodic orbits presented in Chapter 6. The $y = 0$ plane is taken to be the surface of section and the $\dot{y}$ velocity is eliminated using the Jacobi integral.

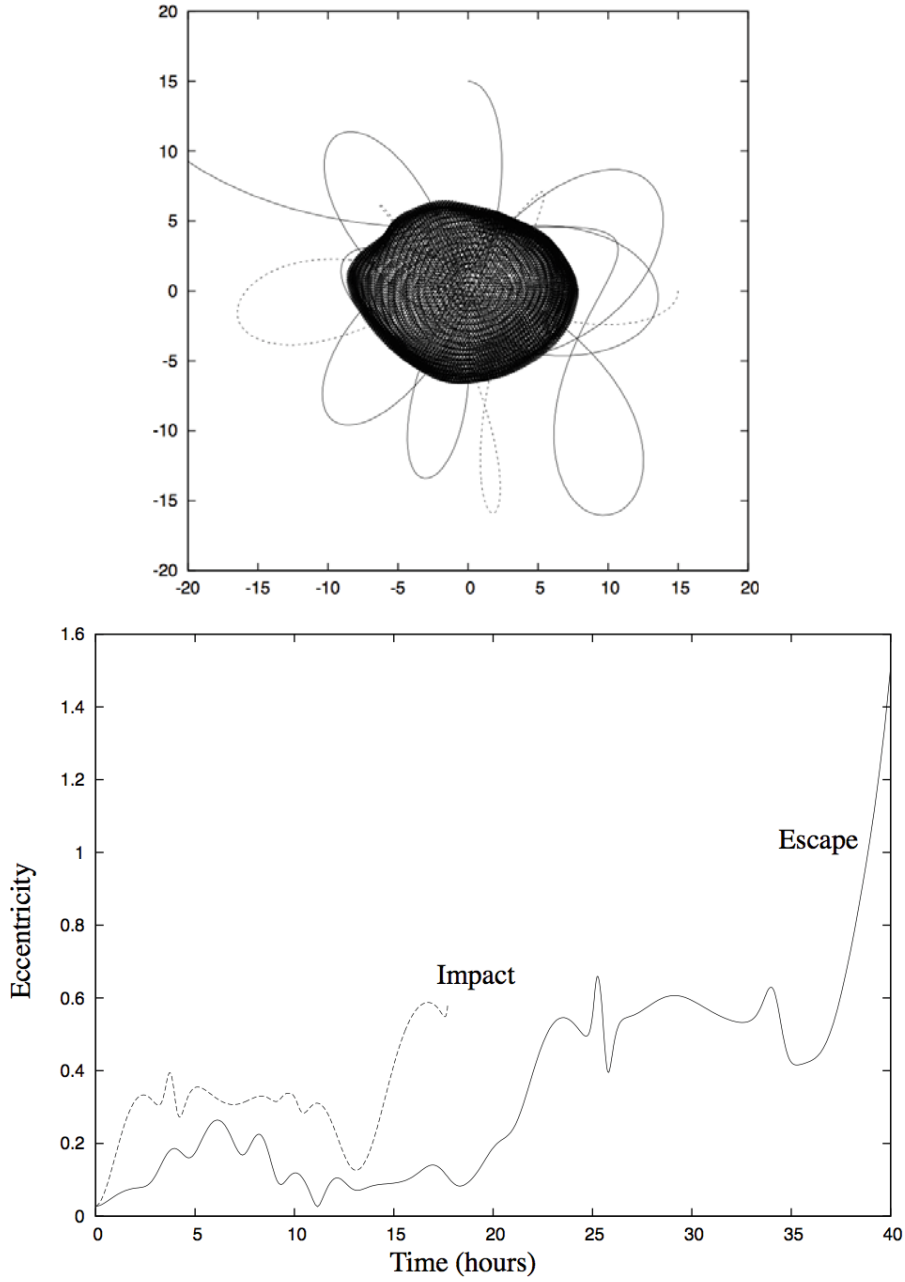


Fig. 18.2 Initially polar, circular orbits about Deimos with semi-major axis of 15 km. Top: View looking down the Deimos z-axis. Bottom: Eccentricity of the orbits as a function of time.

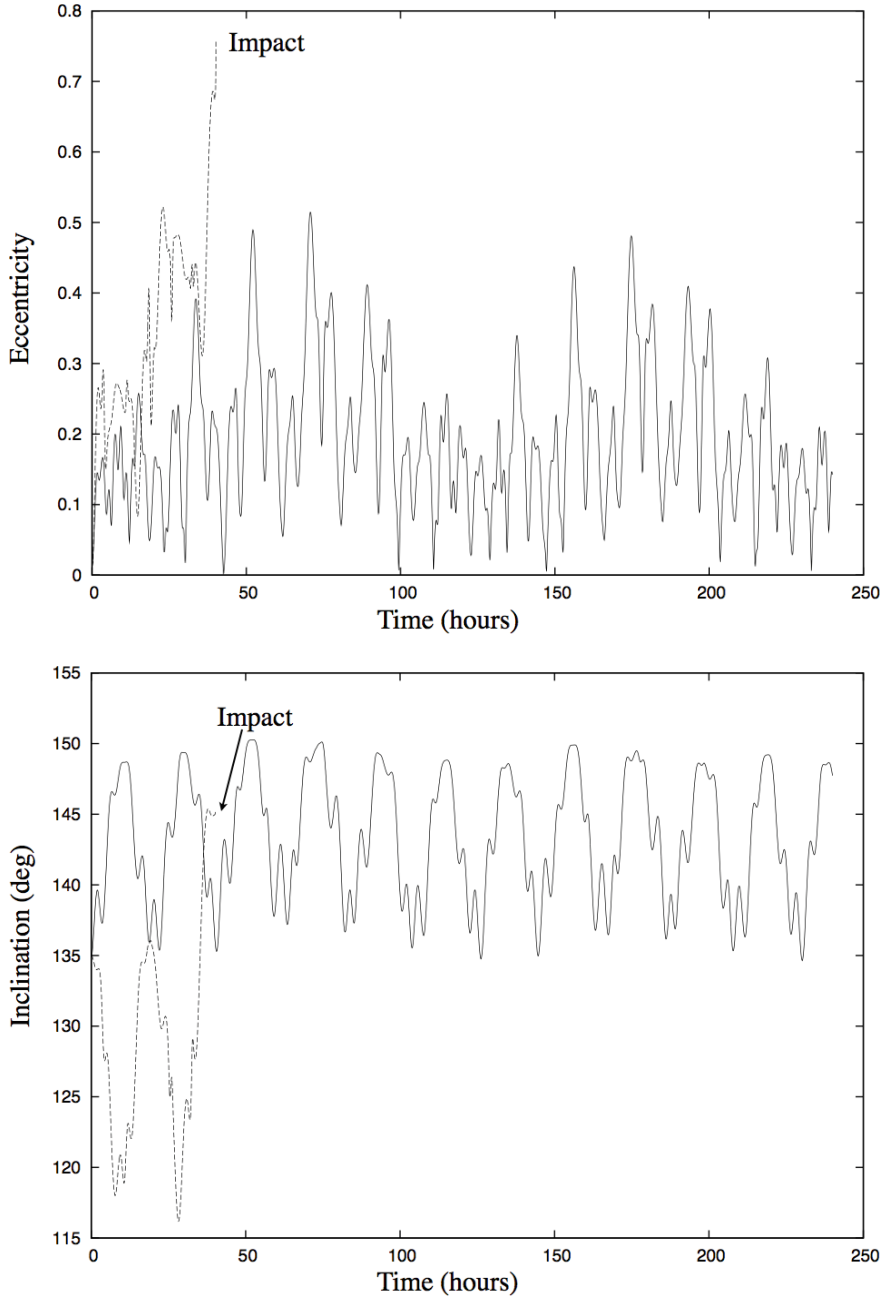


Fig. 18.3 Initially 135° inclination, circular orbits about Deimos with semi-major axis of 15 km. Top: Eccentricity as a function of time. Bottom: Inclination as a function of time.

18.3.3.1 Direct Orbits

First consider a family of direct periodic orbits close to Deimos, with members shown in Fig. 18.4. This particular family is interesting as on either end of the family, the members of these orbits impact onto Deimos. From their morphology, they can be identified with the g family of periodic orbits described by Hénon in [67]. Despite their similarity, the current family has significantly different stability characteristics from the family described by Hénon. In the Hill problem this family becomes stable at lower values of radius, scaled to our problem at a distance of $\sim$10 km, and at this point the family has a bifurcation into the g' family, which is elliptic with a their line of apses aligned with the x-axis. About Deimos, however, this g-like family is unstable for all of its members and does not intersect with any other family. Furthermore, scaling the g' family to Deimos indicates that most of its members would have their periapsis beneath the body, and thus do not form a viable periodic orbit family in this situation. This serves as an excellent reminder of how the scaled size of the planetary satellite in the Hill problem can significantly modify conclusions and shape viable orbital strategies for planetary satellites.

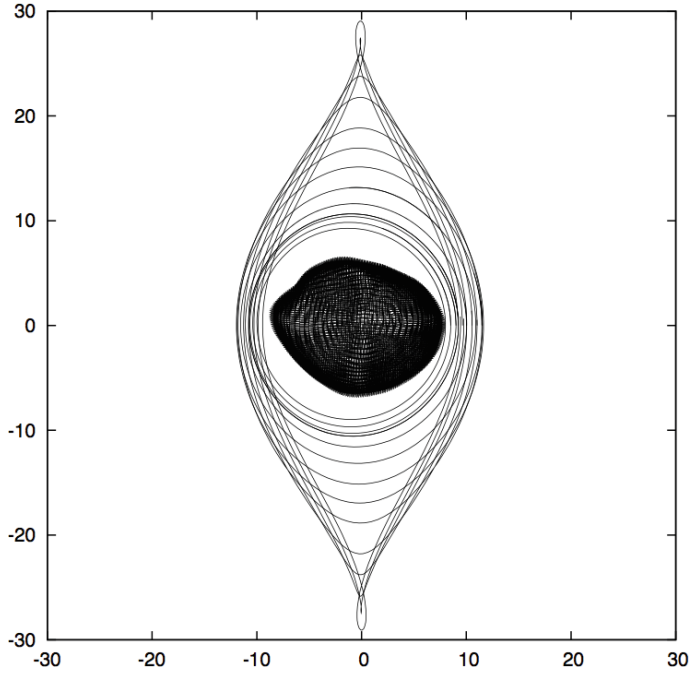


Fig. 18.4 Members of the direct family of periodic orbits about Deimos.

18.3.3.2 Retrograde Periodic Orbits

To end this chapter we focus on retrograde periodic orbits. Per the previous discussion on general trajectories, we expect that retrograde orbits may be stable and could serve as safe and viable candidate orbits for missions to planetary satellites. This is definitely the case. We find an entire family of stable periodic orbits, nominally called the *f* family by Hénon, that exist about Deimos. When far from Deimos, these orbits degenerate to periodic solutions of the Clohessy–Wiltshire equations given above in Eqs. 18.4 to 18.6, and are identified as the larger orbits in Fig. 18.5 which describe a 2:1 ellipse centered on Deimos. These all travel retrograde in the rotating frame, and retrograde relative to Deimos, but they are still direct orbits about Mars. As these orbits come closer to Deimos, the mass of the body begins to affect the dynamics and they become more circular. The entire family is stable, down to the lowest orbit radii shown in Fig. 18.5. At these close orbits, however, the mass distribution of Deimos starts to significantly affect the dynamics. Note that they become tilted out of the plane and actually orbit at a radius closer than the maximum radius of Deimos (Fig. 18.6). The stability of this family across all orbital distances is contrasted with the stability of retrograde orbits about Eros, studied in Chapter 7. There it was seen that destabilizing resonances occurred at some close orbit distances. For the current problem the strength of the Mars tidal accelerations apparently prevents these resonant instabilities from occurring.

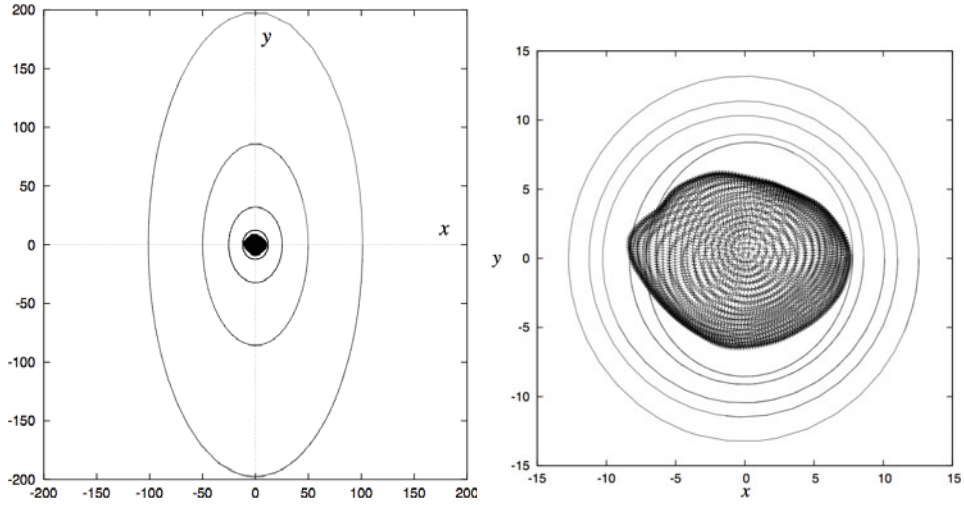


Fig. 18.5 Members of the retrograde family of periodic orbits about Deimos viewed from the z-axis.

As a final point of discussion, we also show plots of the orbit period and x-axis crossing speed of the retrograde periodic orbit family, parameterized by its x-axis crossing value. Figure 18.7 shows the period of the family members, and we note that as the orbit becomes large the period approaches that of Deimos about Mars. Figure 18.8 shows the speeds of the family members at their closest approach to

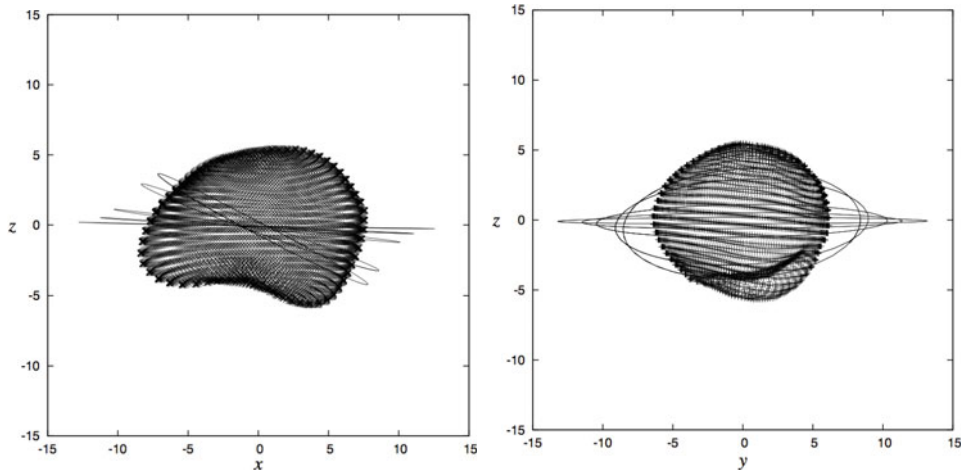


Fig. 18.6 Members of the retrograde family of periodic orbits about Deimos viewed from the y-axis (left) and the x-axis (right).

Deimos. We note that the speed is non-monotonic as the crossing comes closer to Deimos.

Now consider the periodic orbits when far from Deimos. As mentioned previously, as the distance from the body increases, the relative strength of its gravitation attraction becomes insignificant as compared to the tidal effects. This enables the

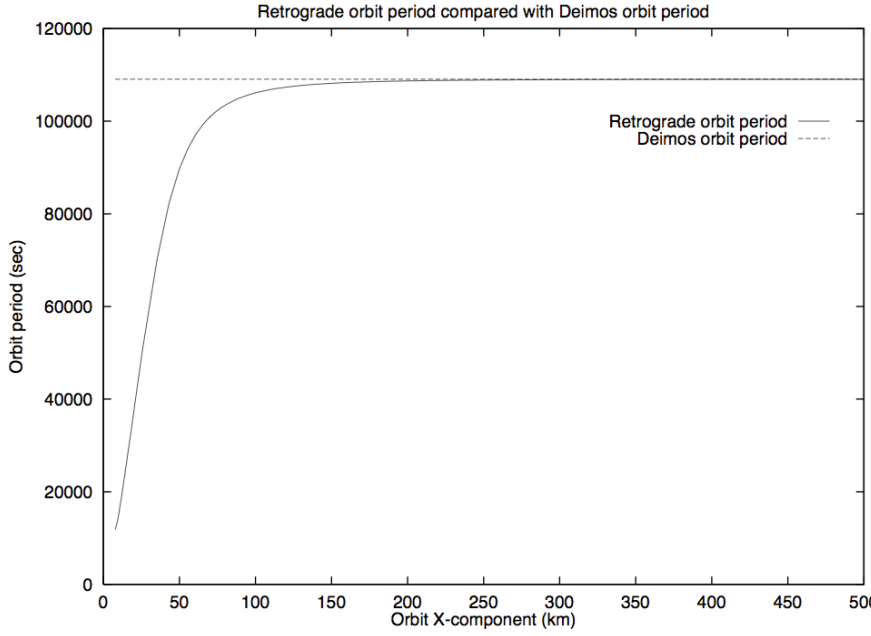


Fig. 18.7 Retrograde periodic orbit periods compared to Deimos's orbit period.

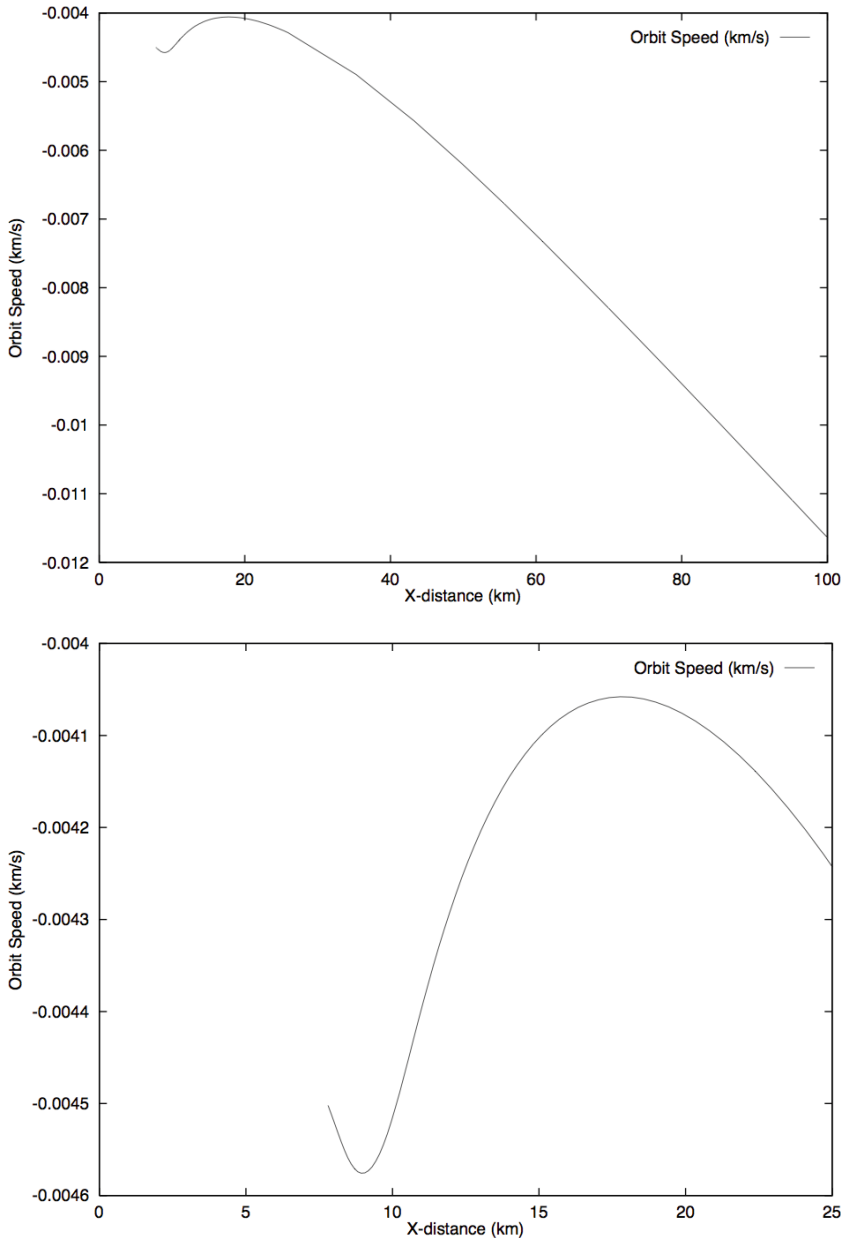


Fig. 18.8 Retrograde periodic orbit speeds at closest approach to Deimos. Top: Shows the entire family. Bottom: Shows details for the closest orbits.

Clohessy–Wiltshire equations to be used to analyze the motion. We set $z = 0$ and only consider planar motion in the following.

Equation 18.5 corresponds to an integral of motion for this system, and can be immediately integrated once to find

$$\dot{y} + 2\varpi x = D \tag{18.9}$$

Substituting this into Eqn. 18.4 to eliminate $\dot{y}$ yields the second-order differential equation

$$\ddot{x} + \varpi^2 x = 2\varpi D \tag{18.10}$$

This is just a simple harmonic oscillator with a constant forcing term. Evaluating its solution yields

$$x(t) = A\cos(\varpi t) + B\sin(\varpi t) + \frac{2}{\varpi}D \tag{18.11}$$

We note that $x(t)$ is completely periodic with a period of $2\pi/\varpi$, equal to the orbit period of Deimos about Mars. Substituting this solution back into Eqn. 18.9 yields a first-order differential equation

$$\dot{y} = -2A\varpi\cos(\varpi t) - 2B\varpi\sin(\varpi t) - 3D \tag{18.12}$$

For the motion of the particle to be periodic, it is clear that the constant D must equal zero. Otherwise, if it is non-zero there will be a secular drift associated with the orbit and the y component will change secularly, taking the particle away from Deimos. Enforcing this condition then provides the necessary condition for the orbit to be periodic

$$\dot{y}_o = -2\varpi x_o \tag{18.13}$$

where the condition can be evaluated at any point of the motion, but is usually computed at the x-axis crossing. Figure 18.8 shows the correct dependence predicted by this formula as x_o grows large. With this proscription, the corresponding orbit is then periodic with a period equal to the Deimos orbit period, which in turn is consistent with Fig. 18.7.

Part IV
Appendices

A. Two-Body Orbit Relations

The following is a collection of useful relationships relating position, velocity and time within an orbit for the two-body problem, using both true anomaly and eccentric anomaly. The main application of these relationships is to averaging, and many of these are used in the following appendix. Due to this only results relevant to elliptic orbits are shown, although most of these results will have analogues for parabolic and hyperbolic orbits.

Mean, Eccentric and True Anomaly

Following are the basic relations between mean, M, eccentric, E, and true, f anomaly. Note the usual symbolic definition of the semi-major axis, a, and the eccentricity, e.

$$M = E - e \sin E \tag{A.1}$$

$$\tan \frac{1}{2} E = \sqrt{\frac{1-e}{1+e}} \tan \frac{1}{2} f \tag{A.2}$$

$$\cos E = \frac{e + \cos f}{1 + e \cos f} \tag{A.3}$$

$$\sin E = \frac{\sqrt{1-e^2} \sin f}{1 + e \cos f} \tag{A.4}$$

$$\cos f = \frac{\cos E - e}{1 - e \cos E} \tag{A.5}$$

$$\sin f = \frac{\sqrt{1-e^2} \sin E}{1 - e \cos E} \tag{A.6}$$

D.J. Scheeres, *Orbital Motion in Strongly Perturbed Environments*, Springer Praxis Books,
DOI 10.1007/978-3-642-03256-1, © Springer-Verlag Berlin Heidelberg 2012

Next are the associated differential relationships between these quantities

$$dM = (1 - e\cos E)\, dE \tag{A.7}$$

$$dM = \frac{(1-e^2)^{3/2}}{(1+e\cos f)^2}\, df \tag{A.8}$$

$$dE = \frac{\sqrt{1-e^2}}{1+e\cos f}\, df \tag{A.9}$$

$$df = \frac{\sqrt{1-e^2}}{1-e\cos E}\, dE \tag{A.10}$$

Scalar Kinematic Quantities

Following are the main scalar kinematic quantities of interest, these being the radius r, speed v, and flight path angle γ. Each are expressed both in terms of true and eccentric anomaly.

$$r = \begin{cases} \dfrac{a(1-e^2)}{1+e\cos f} \\ \\ a(1-e\cos E) \end{cases} \tag{A.11}$$

$$v = \begin{cases} \sqrt{\dfrac{\mu}{a(1-e^2)}\,(1+2e\cos f+e^2)} \\ \\ \sqrt{\dfrac{\mu}{a}\left(\dfrac{1+e\cos E}{1-e\cos E}\right)} \end{cases} \tag{A.12}$$

$$\cos\gamma = \begin{cases} \dfrac{1+e\cos f}{\sqrt{1+2e\cos f+e^2}} \\ \\ \sqrt{\dfrac{1-e^2}{1-e^2\cos^2 E}} \end{cases} \tag{A.13}$$

$$\sin\gamma = \begin{cases} \dfrac{e\sin f}{\sqrt{1+2e\cos f+e^2}} \\ \\ \dfrac{e\sin E}{\sqrt{1-e^2\sin^2 E}} \end{cases} \tag{A.14}$$

$$\tan\gamma = \begin{cases} \dfrac{e\sin f}{1+e\cos f} \\ \\ \dfrac{e}{\sqrt{1-e^2}}\sin E \end{cases} \tag{A.15}$$

Vector Kinematic Quantities

Next the main vector kinematic quantities of interest are expressed, these primarily being the position and velocity vectors as well as the defining orientation integrals of the two-body problem – the eccentricity and angular momentum vectors. First the fundamental vectors are presented, specified using the classical orbit elements relative to an inertial frame.

$$\hat{\boldsymbol{H}} = \sin\Omega \sin i \hat{\boldsymbol{x}} - \cos\Omega \sin i \hat{\boldsymbol{y}} + \cos i \hat{\boldsymbol{z}} \tag{A.16}$$

$$\begin{aligned} \hat{\boldsymbol{e}} &= (\cos\omega\cos\Omega - \cos i \sin\omega \sin\Omega)\, \hat{\boldsymbol{x}} \\ &\quad + (\cos\omega \sin\Omega + \cos i \sin\omega \cos\Omega)\, \hat{\boldsymbol{y}} \\ &\quad + \sin\omega \sin i \hat{\boldsymbol{z}} \end{aligned} \tag{A.17}$$

$$\hat{\boldsymbol{e}}_\perp = \hat{\boldsymbol{h}} \times \hat{\boldsymbol{e}} \tag{A.18}$$

$$\begin{aligned} &= -(\sin\omega\cos\Omega + \cos i \cos\omega \sin\Omega)\, \hat{\boldsymbol{x}} \\ &\quad - (\sin\omega \sin\Omega - \cos i \cos\omega \cos\Omega)\, \hat{\boldsymbol{y}} \\ &\quad + \cos\omega \sin i \hat{\boldsymbol{z}} \end{aligned} \tag{A.19}$$

These reference directions are all conserved in the two-body problem and can be used to specify the position and velocity vectors.

$$\boldsymbol{r} = \begin{cases} r\left[\cos f \hat{\boldsymbol{e}} + \sin f \hat{\boldsymbol{e}}_\perp\right] \\ a\left[(\cos E - e)\, \hat{\boldsymbol{e}} + \sqrt{1-e^2} \sin E \hat{\boldsymbol{e}}_\perp\right] \end{cases} \tag{A.20}$$

$$\boldsymbol{v} = \begin{cases} \sqrt{\frac{\mu}{a(1-e^2)}}\left[-\sin f \hat{\boldsymbol{e}} + (e + \cos f)\, \hat{\boldsymbol{e}}_\perp\right] \\ \frac{\sqrt{\mu a}}{r}\left[\sin E \hat{\boldsymbol{e}} + \sqrt{1-e^2} \cos E \hat{\boldsymbol{e}}_\perp\right] \end{cases} \tag{A.21}$$

B. Fourier Series Expansions of Radius Functions

Following are a simple set of results that are of use in carrying out averaging, the expansion of quantities of the form r^n and $1/r^n$ in terms of Fourier series coefficients, taken from [152].

$$(1 + e\cos\tau)^n = \sum_{m=0}^{\infty} a_m^n \cos(m\tau) \tag{B.1}$$

$$\frac{1}{(1 + e\cos\tau)^n} = \sum_{m=0}^{\infty} b_m^n \cos(m\tau) \tag{B.2}$$

where the coefficients are defined as follows:

$$a_0^n = c_0^n \tag{B.3}$$

$$a_k^n = 2\left(\frac{e}{2}\right)^k c_k^n \tag{B.4}$$

$$b_0^n = \frac{\sqrt{1-e^2}}{(1-e^2)^n} f_0^n \tag{B.5}$$

$$b_k^n = (-1)^k 2\left(\frac{e}{2}\right)^k \frac{\sqrt{1-e^2}}{(1-e^2)^n} f_k^n \tag{B.6}$$

The coefficients c_k^n and f_k^n have the general definitions:

$$c_k^n = \begin{cases} \displaystyle\sum_{l=0}^{[(n-k)/2]} \frac{n!}{l!(l+k)!(n-k-2l)!} \left(\frac{e}{2}\right)^{2l} & n \geq k \\ 0 & n < k \end{cases} \tag{B.7}$$

$$f_k^{n+1} = \begin{cases} \dfrac{(n-k)!(n+k)!}{n!^2} \displaystyle\sum_{l=0}^{[(n-k)/2]} \frac{n!}{l!(l+k)!(n-k-2l)!} \left(\frac{e}{2}\right)^{2l} & n+1 > k \\ \dfrac{n-k}{n}(1-e^2)f_k^n + 2f_{k-1}^{n+1} & n+1 \leq k \end{cases} \tag{B.8}$$

$$f_k^1 = \left(\frac{2}{1+\sqrt{1-e^2}}\right)^k \tag{B.9}$$

where $[a]$ denotes the integer part of a. Note that to compute f_k^{n+1} for $n+1 \leq k$ the recursion relation stated above must be solved.

C. Averaging Results

In orbital dynamics averaging is defined as an operator that computes the time average of a quantity over one orbit period while keeping all other orbit elements constant. The periodicity implies that averaging is only taken for elliptic orbits, and for consistency the averaging is performed over mean anomaly. Thus given a quantity $f(\mathbf{æ}, M)$, defined as a function of mean anomaly M in addition to other orbit elements represented as $\mathbf{æ}$, the average is defined as

$$\bar{f}(\mathbf{æ}) = \frac{1}{2\pi} \int_0^{2\pi} f(\mathbf{æ}, M)\, dM \tag{C.1}$$

where $\bar{f}(\mathbf{æ})$ is independent of mean anomaly but still a function of the other orbit elements. Although the average is defined with respect to mean anomaly, it is often more convenient to compute averages using the true or eccentric anomaly, using the differential relationships given in Appendix A.

First consider powers of the radius r^n for $n \geq 0$.

$$\overline{r^n} = \frac{1}{2\pi} \int_0^{2\pi} r^n\, dM \tag{C.2}$$

$$= \frac{a^n}{2\pi} \int_0^{2\pi} (1 - e\cos E)^{n+1}\, dE \tag{C.3}$$

Using the Fourier series expansion in Eq. B.1 this equals

$$\overline{r^n} = a^n c_0^{n+1}(e) \tag{C.4}$$

$$= a^n \sum_{l=0}^{\left[\frac{n+1}{2}\right]} \frac{(n+1)!}{l!l!(n+1-2l)!} \left(\frac{e}{2}\right)^{2l} \tag{C.5}$$

where $[n + 1/2]$ denotes the integer part of the expression.

Next, consider $1/r^n$ for $n \geq 2$.

$$\overline{\frac{1}{r^n}} = \frac{(1-e^2)^{3/2}}{2\pi p^n} \int_0^{2\pi} (1+e\cos f)^{n-2}\, df \tag{C.6}$$

Using the Fourier series expansion in Eq. B.1 again yields

$$\overline{\frac{1}{r^n}} = \frac{(1-e^2)^{3/2}}{p^n} c_0^{n-2}(e) \tag{C.7}$$

$$= \frac{(1-e^2)^{3/2}}{p^n} \sum_{l=0}^{\left[\frac{n-2}{2}\right]} \frac{(n-2)!}{l!l!(n-2-2l)!} \left(\frac{e}{2}\right)^{2l} \tag{C.8}$$

Finally, for the special case of $1/r$ the average is

$$\overline{\frac{1}{r}} = \frac{1}{a} \tag{C.9}$$

A few specific examples from the above general formula are

$$\overline{r} = a\left(1+\frac{1}{2}e^2\right) \tag{C.10}$$

$$\overline{\frac{1}{r^2}} = \frac{1}{a^2\sqrt{1-e^2}} \tag{C.11}$$

$$\overline{\frac{1}{r^3}} = \frac{1}{a^3(1-e^2)^{3/2}} \tag{C.12}$$

More complex averaging needs to occur when gravitational fields are considered. For the most general results it becomes necessary to average inverse powers of radius times the dot product of $\hat{\boldsymbol{r}}$ with an arbitrary unit vector $\hat{\boldsymbol{p}}$. One can always choose the canonical $\hat{\boldsymbol{z}}$ axis to be aligned with the given unit vector, with the result that $\hat{\boldsymbol{r}} \cdot \hat{\boldsymbol{p}} = \sin i \sin(\omega + f)$. If averaging over the true anomaly eliminates the argument of periapsis, resulting in $\sin i$ terms in the final average, these can be re-expressed using the unit vector and the unitized angular moment as $\sin i = \sqrt{1-(\hat{\boldsymbol{p}} \cdot \hat{\boldsymbol{H}})^2}$. In other cases when the argument of periapsis remains in the averaged expression, it can be related to the dot product of the eccentricity vector with the unit vector. As the average is ultimately expressed in terms of vectors again, the final result is independent of any particular coordinate system. Applying this approach yields the following

$$\overline{\left(\frac{(\hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{r}})^2}{r^3}\right)} = \frac{1}{2a^3(1-e^2)^{3/2}}\left[1-(\hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{H}})^2\right] \tag{C.13}$$

$$\overline{\left(\frac{(\hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{r}})}{r^4}\right)} = \frac{1}{a^4(1-e^2)^{5/2}}\boldsymbol{e}\cdot\hat{\boldsymbol{p}} \tag{C.14}$$

$$\overline{\left(\frac{(\hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{r}})^3}{r^4}\right)} = \frac{3}{4a^4(1-e^2)^{5/2}}\boldsymbol{e}\cdot\hat{\boldsymbol{p}}\left[1-(\hat{\boldsymbol{p}}\cdot\hat{\boldsymbol{H}})^2\right] \tag{C.15}$$

Finally, consider the direct averaging of vectors and their combinations. Here the explicit description of the radius and velocity vectors in terms of orbit elements found in Appendix A are used. Carrying out the averaging operations yields

$$\overline{\left(\frac{\hat{\boldsymbol{r}}\hat{\boldsymbol{r}}}{r^3}\right)} = \frac{1}{2a^3(1-e^2)^{3/2}}\left[\hat{\boldsymbol{e}}\hat{\boldsymbol{e}}+\hat{\boldsymbol{e}}_\perp\hat{\boldsymbol{e}}_\perp\right] \tag{C.16}$$

$$= \frac{1}{2a^3(1-e^2)^{3/2}}\left[\boldsymbol{U}-\hat{\boldsymbol{H}}\hat{\boldsymbol{H}}\right] \tag{C.17}$$

where the second equation was simplified using the identity $\boldsymbol{U}=\hat{\boldsymbol{e}}\hat{\boldsymbol{e}}+\hat{\boldsymbol{e}}_\perp\hat{\boldsymbol{e}}_\perp+\hat{\boldsymbol{H}}\hat{\boldsymbol{H}}$. Next, averaging the position and velocity vectors and some of their dyad products yields

$$\overline{\boldsymbol{r}} = -\frac{3}{2}a\boldsymbol{e} \tag{C.18}$$

$$\overline{\boldsymbol{v}} = \boldsymbol{0} \tag{C.19}$$

$$\overline{\boldsymbol{r}\boldsymbol{r}} = \frac{1}{2}a^2\left[(1+4e^2)\hat{\boldsymbol{e}}\hat{\boldsymbol{e}}+(1-e^2)\hat{\boldsymbol{e}}_\perp\hat{\boldsymbol{e}}_\perp\right] \tag{C.20}$$

$$\overline{\boldsymbol{r}\boldsymbol{v}} = -\frac{1}{2}\tilde{\boldsymbol{H}} \tag{C.21}$$

where $\boldsymbol{H}$ denotes the angular momentum vector. The final averaging result for $\boldsymbol{r}\boldsymbol{v}$ is difficult, and so it described below.

The dyad product of these two vectors can be reduced to the following dyadic

$$\begin{aligned}\boldsymbol{r}\boldsymbol{v} = H\,[&\cos f(\tan\gamma\cos f-\sin f)\hat{\boldsymbol{e}}\hat{\boldsymbol{e}}\\ &+\sin f(\tan\gamma\sin f+\cos f)\hat{\boldsymbol{e}}_\perp\hat{\boldsymbol{e}}_\perp\\ &+\cos f(\tan\gamma\sin f+\cos f)\hat{\boldsymbol{e}}\hat{\boldsymbol{e}}_\perp\\ &+\sin f(\tan\gamma\cos f-\sin f)\hat{\boldsymbol{e}}_\perp\hat{\boldsymbol{e}}]\end{aligned} \tag{C.22}$$

Noting that $\tan\gamma$ and $\sin(mf)$ are both odd functions in true anomaly the factors of $\hat{\boldsymbol{e}}\hat{\boldsymbol{e}}$ and $\hat{\boldsymbol{e}}_\perp\hat{\boldsymbol{e}}_\perp$ will ultimately be odd in true anomaly, and hence will average to zero. The same is not true of the other factors. Applying trigonometric identities these factors can be rewritten as

$$\frac{1}{2}H\left[(\hat{\boldsymbol{e}}\hat{\boldsymbol{e}}_\perp-\hat{\boldsymbol{e}}_\perp\hat{\boldsymbol{e}})+\frac{\cos(2f)+e\cos f}{1+e\cos f}(\hat{\boldsymbol{e}}\hat{\boldsymbol{e}}_\perp+\hat{\boldsymbol{e}}_\perp\hat{\boldsymbol{e}})\right] \tag{C.23}$$

The first term can be identified with the dyadic $-\frac{1}{2}\tilde{\boldsymbol{H}}$, and thus is a constant. The second term can be averaged over true anomaly and, using the Fourier series expansions, be shown to equal

$$\frac{H}{2}\frac{e^2}{1-e^2}\left[\frac{1}{2}f_2^3 - f_1^3\right](\hat{\boldsymbol{e}}\hat{\boldsymbol{e}}_\perp + \hat{\boldsymbol{e}}_\perp\hat{\boldsymbol{e}}) \tag{C.24}$$

where the f functions are defined in Eq. B.9. Computation of the factors yields $f_2^3 = 3!$ and $f_1^3 = 3!/2$, and hence the term in Eq. C.24 is identically zero, proving the result.

D. Canonical Transformations

The theory of canonical transformations lies at the heart of Hamiltonian dynamics. They serve as an equivalence principle between solutions and coordinate changes in these systems, and in essence make entire classes of Hamiltonian dynamical systems of similar order equivalent. It is important to note that there remains fundamental differences between Hamiltonian systems that are integrable and those that are non-integrable, a topic that goes beyond the current text.

Symplectic Matrix Properties

Define the symplectic unit matrix $J_{2n} \in \mathbf{R}^{2n\times 2n}$ as

$$J_{2n} = \begin{bmatrix} 0_n & I_n \\ -I_n & 0_n \end{bmatrix} \tag{D.1}$$

where I_n and 0_n are $n \times n$ identity and zero matrices, respectively. The $2n$ subscript is suppressed, as the dimension will usually be obvious. This matrix is skew-symmetric, $J^T = -J$. Also note that $JJ = -I$ and that it is orthonormal, or, $JJ^T = J^T J = I$.

A real matrix $M \in \mathbf{R}^{2n\times 2n}$ is defined as a symplectic matrix if it satisfies the equation

$$J = M^T JM \tag{D.2}$$

Based on this definition a series of equivalent statements and identities for symplectic matrices can be recounted. The following list is compiled from [132].

- J is a symplectic matrix.
- I is a symplectic matrix.
- If M is symplectic, then so is M^T, $-M$, and M^{-1}.
- If M_1 and M_2 are symplectic, then $M_1 M_2$ is symplectic.
- The determinant of a symplectic matrix always equals 1, hence an inverse always exists.

- The inverse of a symplectic matrix M is $M^{-1} = -JM^TJ$.
- From the previous statements it can be proven that the symplectic matrices of a fixed size form a group.
- Assume a matrix M has the following form:

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \tag{D.3}$$

  where each of the sub-matrices is an $n \times n$ matrix. Then M is a symplectic matrix if and only if the following identities hold: (i) A^TC and B^TD are symmetric, (ii) $D^TA - B^TC = I$, the identity.
- If M is a 2×2 matrix, then it is a symplectic matrix if and only if its determinant equals unity.
- The characteristic equation of a symplectic matrix, $\|\lambda I - M\| = 0$, is a symmetric polynomial in λ.
- The eigenvalues of a symplectic matrix come in inverse pairs, i.e., if λ is an eigenvalue (possibly complex), then so is λ^{-1}. As M is a real matrix, the complex conjugates of the eigenvalues must also be eigenvalues. Thus if λ is an eigenvalue, then so must λ^{-1}, $\bar{\lambda}$, and $\bar{\lambda}^{-1}$. Note that in some cases $\lambda = \bar{\lambda}$ or $\bar{\lambda} = \lambda^{-1}$. In these cases the eigenvalues only need come in pairs.
- If $\boldsymbol{u}$ is a right eigenvector of a symplectic matrix with eigenvalue λ, then $\boldsymbol{v} = -J\boldsymbol{u}$ is a left eigenvector with eigenvalue λ^{-1}.

Tests for Canonical Transformations

There are a number of fundamental tests for whether a transformation is canonical, and thus whether it will transform the Hamiltonian equations of motion back into a similar form. We list a few common tests in the following. Assume an initial Hamiltonian system denoted as $\boldsymbol{x} = [\boldsymbol{q}, \boldsymbol{p}]$, and a candidate transformation $\boldsymbol{X}(\boldsymbol{x}) = [\boldsymbol{Q}(\boldsymbol{x}), \boldsymbol{P}(\boldsymbol{x})]$ that will be tested for whether it is canonical.

- The transformation is canonical if the Jacobian of the transformation, $M = \partial \boldsymbol{X}/\partial \boldsymbol{x}$, is a symplectic matrix.
- The Lagrange brackets are defined for a transformation $\boldsymbol{Q}(u, v) \in \mathbf{R}^n$ and $\boldsymbol{P}(u, v) \in \mathbf{R}^n$ by the operator:

$$[u, v] = \sum_{i=1}^{n} \left(\frac{\partial Q_i}{\partial u} \frac{\partial P_i}{\partial v} - \frac{\partial P_i}{\partial u} \frac{\partial Q_i}{\partial v} \right) \tag{D.4}$$

  A transformation $\boldsymbol{X}(\boldsymbol{x})$ is canonical if and only if the following Lagrange bracket values hold

$$[q_j, q_k] = 0 \tag{D.5}$$
$$[p_j, p_k] = 0 \tag{D.6}$$
$$[q_j, p_k] = \delta_{jk} \tag{D.7}$$

where δ_{jk} is the Kronecker delta function. Note that Lagrange brackets are formally only a function of the original variables. Recall that the Lagrange brackets were also used in the derivation of the Lagrange planetary equations.

- The Poisson brackets can be considered to be the inverse of the Lagrange brackets, and are defined for a transformation $u(\boldsymbol{q}, \boldsymbol{p})$ and $v(\boldsymbol{q}, \boldsymbol{p})$ by the operator:

$$\{u, v\} = \sum_{i=1}^{n} \left(\frac{\partial u}{\partial q_i} \frac{\partial v}{\partial p_i} - \frac{\partial u}{\partial p_i} \frac{\partial v}{\partial q_i} \right) \tag{D.8}$$

A transformation $\boldsymbol{X}(\boldsymbol{x})$ is canonical if and only if the following Poisson bracket values hold

$$\{Q_j, Q_k\} = 0 \tag{D.9}$$
$$\{P_j, P_k\} = 0 \tag{D.10}$$
$$\{Q_j, P_k\} = \delta_{jk} \tag{D.11}$$

Note that the Poisson brackets are formally only a function of the new variables. Recall that the Poisson brackets were also used in the definition of the time derivative of a scalar quantity for a Hamiltonian dynamical system.

- Another important test of whether a proposed transformation is canonical is through the definition of the generating functions. Specifically, given a transformation of the form given above, it is a canonical transformation if and only if the following differential forms are exact after replacement of $\boldsymbol{Q}$ and $\boldsymbol{P}$ as functions of $\boldsymbol{q}$ and $\boldsymbol{p}$.

$$\boldsymbol{P} \cdot d\boldsymbol{Q} - \boldsymbol{p} \cdot d\boldsymbol{q} \tag{D.12}$$
$$\boldsymbol{Q} \cdot d\boldsymbol{P} + \boldsymbol{p} \cdot d\boldsymbol{q} \tag{D.13}$$
$$\boldsymbol{P} \cdot d\boldsymbol{Q} + \boldsymbol{q} \cdot d\boldsymbol{p} \tag{D.14}$$
$$\boldsymbol{Q} \cdot d\boldsymbol{P} - \boldsymbol{q} \cdot d\boldsymbol{p} \tag{D.15}$$

These conditions can all be rigorously related to each other (see [132, 56]). If such a condition is exact, then it equals the differential of a scalar function, called a generating function.

As a simple application of the above tests, consider a Hamiltonian system defined by coordinates $\boldsymbol{q}$ and momenta $\boldsymbol{p}$ and the proposed transformation $\boldsymbol{P} = -\boldsymbol{q}$ and $\boldsymbol{Q} = \boldsymbol{p}$, which swaps the roles of the coordinates and momenta.

The Jacobian of this transformation can be easily shown to be J, the symplectic unit matrix. As this is a symplectic matrix, the transformation is seen to be canonical.

For the Lagrange and Poisson bracket tests, it suffices to only consider one of the elements, say $Q_1 = p_1$ and $P_1 = -q_1$. The first two conditions of each test are trivially satisfied, and the third yields

$$[q_1, p_1] = \frac{\partial p_1}{\partial q_1}\frac{\partial(-q_1)}{\partial p_1} - \frac{\partial(-q_1)}{\partial q_1}\frac{\partial p_1}{\partial p_1} = 1 \tag{D.16}$$

$$\{Q_1, P_1\} = \frac{\partial Q_1}{\partial(-P_1)}\frac{\partial P_1}{\partial Q_1} - \frac{\partial Q_1}{\partial Q_1}\frac{\partial P_1}{\partial(-P_1)} = 1 \tag{D.17}$$

Since the identity transformation is obviously canonical, the above result also shows that just swapping one coordinate-momentum pair is also a canonical transformation.

For the exactness tests, we find the following:

$$\boldsymbol{P}\cdot d\boldsymbol{Q} - \boldsymbol{p}\cdot d\boldsymbol{q} = -d(\boldsymbol{q}\cdot\boldsymbol{p}) \tag{D.18}$$

$$\boldsymbol{Q}\cdot d\boldsymbol{P} + \boldsymbol{p}\cdot d\boldsymbol{q} = 0 \tag{D.19}$$

$$\boldsymbol{P}\cdot d\boldsymbol{Q} + \boldsymbol{q}\cdot d\boldsymbol{p} = 0 \tag{D.20}$$

$$\boldsymbol{Q}\cdot d\boldsymbol{P} - \boldsymbol{q}\cdot d\boldsymbol{p} = -d(\boldsymbol{q}\cdot\boldsymbol{p}) \tag{D.21}$$

all of which are exact differentials and thus canonical.

E. Legendre Polynomials and Associated Functions

The associated Legendre functions are defined by the simple rule

$$P_{lm}(x) = (1 - x^2)^{m/2} \frac{d^m}{dx^m}(P_{l0}(x)) \tag{E.1}$$

The $P_{l0}(x)$ are the Legendre polynomials and are defined as

$$P_{l0} = \frac{1}{2^l l!} \frac{d^l}{dx^l}(x^2 - 1)^l \tag{E.2}$$

These functions and polynomials arise in various expansions of interest, both for the perturbation of a third body and in the spherical harmonic gravity field expansion.

A list of these polynomials and functions up to degree and order 4 are given below:

$$\begin{aligned}
P_{00}(x) &= 1 \\
P_{10}(x) &= x \\
P_{11}(x) &= \sqrt{1 - x^2} \\
P_{20}(x) &= \frac{1}{2}(3x^2 - 1) \\
P_{21}(x) &= 3x\sqrt{1 - x^2} \\
P_{22}(x) &= 3(1 - x^2) \\
P_{30}(x) &= \frac{1}{2}(5x^3 - 3x) \\
P_{31}(x) &= \frac{3}{2}(5x^2 - 1)\sqrt{1 - x^2} \\
P_{32}(x) &= 15x(1 - x^2) \\
P_{33}(x) &= 15(1 - x^2)\sqrt{1 - x^2}
\end{aligned}$$

$$
\begin{aligned}
P_{40}(x) &= \frac{1}{8}(35x^4 - 30x^2 + 3) \\
P_{41}(x) &= \frac{5}{2}(7x^3 - 3x)\sqrt{1 - x^2} \\
P_{42}(x) &= \frac{15}{2}(7x^2 - 1)(1 - x^2) \\
P_{43}(x) &= 105x(1 - x^2)\sqrt{1 - x^2} \\
P_{44}(x) &= 105(1 - x^2)^2
\end{aligned}
$$

F. Elliptic Functions and Integrals

The following sections briefly cover some of the essential qualities and definitions of elliptic functions and integrals. There is a wealth of material pertaining to these quantities not discussed here as it is not needed in the current analysis. See References [34] and [24] for complete discussions of these functions and integrals.

Elliptic Integrals

There are three types of Jacobi elliptic integrals. They are called elliptic integrals of the 1st, 2nd and 3rd kind. The elliptic integral of the 1st kind can be represented as:

$$F(x,k) = \int_0^x \frac{dy}{\sqrt{(1-y^2)(1-k^2y^2)}} \tag{F.1}$$

or its equivalent (assuming $x = \sin\phi$)

$$F(\phi,k) = \int_0^\phi \frac{d\theta}{\sqrt{1-k^2\sin^2\theta}} \tag{F.2}$$

where $k^2 < 1$. This elliptic integral is called complete if $x = 1$ or $\phi = \pi/2$, and is then denoted as:

$$K(k) = F(1,k) \tag{F.3}$$

$$= F(\pi/2,k) \tag{F.4}$$

$$= \pi/2 \sum_{n=0}^{\infty} \left(\frac{(2n)!}{2^{2n}(n!)^2} \right)^2 k^{2n} \tag{F.5}$$

The elliptic integral of the 2nd kind can be represented as:

$$E(x,k) = \int_0^x \sqrt{\frac{(1-k^2y^2)}{(1-y^2)}}\, dy \tag{F.6}$$

or its equivalent (assuming $x = \sin\phi$)

$$E(\phi,k) = \int_0^\phi \sqrt{1-k^2\sin^2\theta}\, d\theta \tag{F.7}$$

where $k^2 < 1$. This elliptic integral is called complete if $x = 1$ or $\phi = \pi/2$, and is then denoted as:

$$E(k) = E(1,k) \tag{F.8}$$
$$= E(\pi/2,k) \tag{F.9}$$
$$= \pi/2 \sum_{n=0}^{\infty} \left(\frac{(2n)!}{2^{2n}(n!)^2}\right)^2 \frac{k^{2n}}{2n-1} \tag{F.10}$$

The elliptic integral of the 3rd kind is can be represented as:

$$\Pi(x,n,k) = \int_0^x \frac{dy}{(1+ny^2)\sqrt{(1-y^2)(1-k^2y^2)}} \tag{F.11}$$

or its equivalent (assuming $x = \sin\phi$)

$$\Pi(\phi,n,k) = \int_0^\phi \frac{d\theta}{(1+n\sin^2\theta)\sqrt{1-k^2\sin^2\theta}} \tag{F.12}$$

where $k^2 < 1$, and there are no conditions on n. Again the integral is called complete if $x = 1$ or $\phi = \pi/2$, although there are no simple expansions as there are for the complete elliptic integrals of the 1st and 2nd kind.

The above integrals will be briefly reconsidered in the following section.

Elliptic Functions

The elliptic functions may be defined as the inverse of the elliptic integral of the first kind. Consider the elliptic integral of the 1st kind $F(x_1,k)$. Then the elliptic function denoted by $\mathrm{sn}(\tau,k) = \mathrm{sn}(\tau)$ is defined by the relation:

$$x_1 = \mathrm{sn}[F(x_1,k)] \tag{F.13}$$

or, more suggestively:

$$F(\mathrm{sn}(\tau),k) = \int_0^{\mathrm{sn}(\tau)} \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}} \tag{F.14}$$

The function $\mathrm{sn}(\tau)$ is periodic with a real period $4K(k)$, so that $\mathrm{sn}\left(\tau + 4K(k)\right) = \mathrm{sn}(\tau)$. Also note that $\mathrm{sn}(\tau, k = 0) = \sin(\tau)$, $\mathrm{sn}(\tau, k = 1) = \tanh(\tau)$, and $|\mathrm{sn}(\tau)| \leq 1$.

There are three other related elliptic functions, denoted as $\mathrm{cn}(\tau)$, $\mathrm{dn}(\tau)$ and $\mathrm{tn}(\tau)$. These are easily defined given $\mathrm{sn}(\tau)$ as:

$$\mathrm{cn}(\tau) = \pm\sqrt{1 - \mathrm{sn}^2(\tau)} \tag{F.15}$$

$$\mathrm{dn}(\tau) = \sqrt{1 - k^2\mathrm{sn}^2(\tau)} \tag{F.16}$$

$$\mathrm{tn}(\tau) = \frac{\mathrm{sn}(\tau)}{\mathrm{cn}(\tau)} \tag{F.17}$$

Note that the proper sign of cn must be chosen, and that $0 \leq \mathrm{dn} \leq 1$.

It is instructive to rewrite the elliptic integrals in terms of elliptic functions. The elliptic integral of the 1st kind is then simply:

$$F(\tau) = \int_0^\tau d\tau \tag{F.18}$$

which appears to be, and is, trivial.

The elliptic integral of the 2nd kind is then:

$$E(\tau) = \int_0^\tau \mathrm{dn}^2(\tau)\, d\tau \tag{F.19}$$

and the elliptic integral of the 3rd kind is:

$$\Pi(\tau, n) = \int_0^\tau \frac{d\tau}{1 + n\mathrm{sn}^2(\tau)} \tag{F.20}$$

Note that the dependence of the integrals on the k parameter has been suppressed.

In the above definition of the elliptic integrals, there is no a priori bound on the value of τ, and indeed τ may increase arbitrarily. Note that due to the periodicity of the elliptic functions, the integrands will always cycle through the same values, although the total integral will continue to increase with τ. This property of the elliptic integrals is not evident in the usual definitions given in the previous section, where the parameter x can never be greater than 1. Due to this property, the elliptic integrals must be carefully evaluated to ensure that the total integral is being computed, and not just the principal value.

References

1. S. Abe, T. Mukai, N. Hirata, O.S. Barnouin-Jha, A.F. Cheng, H. Demura, R.W. Gaskell, T. Hashimoto, K. Hiraoka, T. Honda, T. Kubota, M. Matsuoka, T. Mizuno, R. Nakamura, D.J. Scheeres, and M. Yoshikawa. Mass and local topography measurements of Itokawa by Hayabusa. *Science*, 312:1344–1349, June 2006.
2. E.L. Akim, E.G. Ruzskiy, V.A. Shishov, V.A. Stepaniants, and A.G. Tuchin. Ballistics, navigation and motion control of the S/C on stages of the Phobos surface approaching and landing. In *18th International Symposium on Space Flight Dynamics*, volume 548, page 461, 2004.
3. P. Antreasian, C. Helfrich, J. Giorgini, J. Miller, W. Owen, B. Williams, D. Yeomans, D. Dunham, R. Farquhar, J. McAdams, et al. Preliminary considerations for NEAR's low-altitude passes and landing operations at 433 Eros. In *Astrodynamics Specialist Conference*. American Institute of Aeronautics and Astronautics, 1998. Paper AIAA 1998-4397.
4. V. I. Arnol'd. *Mathematical Methods of Classical Mechanics*. Springer, 2nd edition, September 1989.
5. V.I. Arnold and A. Avez. *Ergodic Problems of Classical Mechanics*. Benjamin, 1968.
6. V.I. Arnold, V.V. Kozlov, and A.I. Neishtadt. *Mathematical Aspects of Classical and Celestial Mechanics*. Springer, 2006.
7. E. Asphaug. Critical crater diameter and asteroid impact seismology. *Meteoritics and Planetary Science*, 43:1075–1084, September 2008.
8. G. Balmino. Gravitational potential harmonics from the shape of an homogeneous body. *Celestial Mechanics and Dynamical Astronomy*, 60:331–364, November 1994.
9. P. Bartczak and S. Breiter. Double material segment as the model of irregular bodies. *Celestial Mechanics and Dynamical Astronomy*, 86:131–141, June 2003.
10. J.E. Bellerose. *The Restricted Full Three Body Problem: Applications to Binary Asteroid Exploration*. PhD thesis, The University of Michigan, 2008.
11. M.J.S. Belton, C.R. Chapman, P.C. Thomas, M.E. Davies, R. Greenberg, K. Klaasen, D. Byrnes, L. D'Amario, S. Synnott, T.V. Johnson, A. McEwen, W.J. Merline, D.R. Davis, J.-M. Petit, A. Storrs, J. Veverka, and B. Zellner. Bulk density of asteroid 243 Ida from the orbit of its satellite Dactyl. *Nature*, 374:785–788, April 1995.
12. M.J.S. Belton, N.H. Samarasinha, Y.R. Fernández, and K.J. Meech. The excited spin state of Comet 2P/Encke. *Icarus*, 175:181–193, May 2005.
13. D. Boccaletti and G. Pucacco. *Theory of Orbits: Volume 2: Perturbative and Geometrical Methods*. Springer-Verlag Berlin Heidelberg New York. Also Astronomy and Astrophysics Library, 1999.
14. W.F. Bottke, Jr., A. Cellino, P. Paolicchi, and R.P. Binzel, editors. *Asteroids III*. Univ. Arizona Press, 2002.
15. D.T. Britt, D.K. Yeomans, K. Housen, and G. Consolmagno. Asteroid density, porosity, and structure. In *Asteroids III*, pages 485–500. Univ. Arizona Press, 2002.

16. S.B. Broschart and D.J. Scheeres. Control of hovering spacecraft near small bodies: Application to Asteroid 25143 Itokawa. *Journal of Guidance, Control and Dynamics*, 28(2):343–354, 2005.
17. S.B. Broschart and D.J. Scheeres. Boundedness of spacecraft hovering under dead-band control in time-invariant systems. *Journal of Guidance, Control and Dynamics*, 30(2):601–610, 2007.
18. R. Broucke. Periodic collision orbits in the elliptic restricted three-body problem. *Celestial Mechanics*, 3:461–477, July 1971.
19. D. Brouwer and G.M. Clemence. *Methods of Celestial Mechanics*. New York: Academic Press, 1961.
20. J.A. Burns and V.S. Safronov. Asteroid nutation angles. *Monthly Notices of the Royal Astronomical Society*, 165:403–+, 1973.
21. S.M. Byram. *The Effects of Outgassing Jets on the Rotation of a Comet Nucleus and on the Trajectory of an Orbiting Spacecraft*. PhD thesis, The University of Michigan, 2008.
22. S.M. Byram and D.J. Scheeres. Stability of Sun-synchronous orbits in the vicinity of a comet. *Journal of Guidance, Control and Dynamics*, 32(5):1550–1559, 2009.
23. S.M. Byram, D.J. Scheeres, and M.R. Combi. Models for the comet dynamical environment. *Journal of Guidance, Control and Dynamics*, 30(5):1445–1454, 2007.
24. P.F. Byrd and M.D. Friedman. *Handbook of Elliptic Integrals for Engineers and Physicists*. Springer-Verlag, 1954.
25. S.R. Chesley, S.J. Ostro, D. Vokrouhlický, D. Čapek, J.D. Giorgini, M.C. Nolan, J.-L. Margot, A.A. Hine, L.A.M. Benner, and A.B. Chamberlin. Direct detection of the Yarkovsky effect by radar ranging to Asteroid 6489 Golevka. *Science*, 302:1739–1742, December 2003.
26. P. Colwell. *Solving Kepler's Equation over Three Centuries*. Richmond, VA: Willmann-Bell, 1993.
27. H. Dankowicz. Some special orbits in the two-body problem with radiation pressure. *Celestial Mechanics and Dynamical Astronomy*, 58:353–370, April 1994.
28. H. Dankowicz. The two-body problem with radiation pressure in a rotating reference frame. *Celestial Mechanics and Dynamical Astronomy*, 61:287–313, March 1995.
29. H. Dankowicz. Escape of particles orbiting asteroids in the presence of radiation pressure through separatrix splitting. *Celestial Mechanics and Dynamical Astronomy*, 67:63–85, January 1997.
30. H. Dankowicz. Slow diffusion and effective stability of dust particles orbiting asteroids. *Celestial Mechanics and Dynamical Astronomy*, 84:1–25, September 2002.
31. A. Deprit and J. Henrard. Construction of orbits asymptotic to a periodic orbit. *Astronomical Journal*, 74:308–+, 1969.
32. A.R. Dobrovolskis and J.A. Burns. Life near the Roche limit: Behavior of ejecta from satellites close to planets. *Icarus*, 42(3):422–441, 1980.
33. M. Efroimsky and P. Goldreich. Gauge symmetry of the N-body problem in the Hamilton-Jacobi approach. *Journal of Mathematical Physics*, 44:5958–5977, December 2003.
34. A. Erdelyi, editor. *Higher Transcendental Functions, Vol. II*. McGraw-Hill, 1953.
35. E.G. Fahnestock. *The Full Two Body Problem: Simulation, Analysis, and Application to the Dynamics, Characteristics, and Evolution of Binary Asteroid Systems*. PhD thesis, The University of Michigan, 2008.
36. E.G. Fahnestock and D.J. Scheeres. Simulation of the full two rigid body problem using polyhedral mutual potential and potential derivatives approach. *Celestial Mechanics and Dynamical Astronomy*, 96:317–339, November 2006.
37. E.G. Fahnestock and D.J. Scheeres. Dynamical characterization and stabilization of large gravity tractor designs. *Journal of Guidance, Control and Dynamics*, 31(3):501–521, 2008.
38. E.G. Fahnestock and D.J. Scheeres. Simulation and analysis of the dynamics of binary near-Earth Asteroid (66391) 1999 KW4. *Icarus*, 194:410–435, April 2008.
39. P. Farinella et al. Triaxial equilibrium ellipsoids among the asteroids? *Icarus*, 46(1):114–123, 1981.

40. R.W. Farquhar and A.A. Kamel. Quasi-periodic orbits about the translunar libration point. *Celestial Mechanics*, 7:458–473, June 1973.
41. M.C. Festou, H.U. Keller, and H.A. Weaver, editors. *Comets II*. Univ. Arizona Press, 2004.
42. Cl. Froeschlé. The Lyapunov characteristic exponents-applications to celestial mechanics. *Celestial Mechanics and Dynamical Astronomy*, 34:95–115, 1984. 10.1007/BF01235793.
43. R. Garmier, J.-P. Barriot, A.S. Konopliv, and D.K. Yeomans. Modeling of the Eros gravity field as an ellipsoidal harmonic expansion from the NEAR Doppler tracking data. *Geophysical Research Letters*, 29(8):1231, April 2002.
44. R. Gaskell, O. Barnouin-Jha, D. Scheeres, T. Mukai, N. Hirata, S. Abe, J. Saito, M. Ishiguro, T. Kubota, T. Hashimoto, et al. Landmark navigation studies and target characterization in the Hayabusa encounter with Itokawa. In *Astrodynamics Specialist Conference*. American Institute of Aeronautics and Astronautics, 2006. Paper AIAA 2006-6660.
45. R. Gaskell, J. Saito, M. Ishiguro, T. Kubota, T. Hashimoto, N. Hirata, S. Abe, O. Barnouin-Jha, and D. Scheeres. Gaskell Itokawa Shape Model V1.0. *NASA Planetary Data System*, 92, September 2008.
46. R.W. Gaskell. Automated landmark identification for spacecraft navigation. In *Advances in the Astronautical Sciences*, volume 109, 2002. Paper AAS 01-422.
47. R.W. Gaskell. Gaskell Eros Shape Model V1.0. *NASA Planetary Data System*, 96, September 2008.
48. T. Gehrels, editor. *Asteroids*. Univ. Arizona Press, 1979.
49. B. Gladman, P. Michel, and C. Froeschlé. The near-Earth object population. *Icarus*, 146:176–189, July 2000.
50. P. Goldreich and R. Sari. Tidal evolution of rubble piles. *Astrophysics Journal*, 691:54–60, January 2009.
51. G. Gómez, J. Llibre, R. Martínez, and C. Simó. *Dynamics and Mission Design Near Libration Points Volume I: Fundamentals: The Case of Collinear Libration Points*. World Scientific, 2001.
52. K.R. Grazier, W.I. Newman, W.M. Kaula, and J.M. Hyman. Dynamical evolution of planetesimals in the outer solar system. I. The Jupiter/Saturn zone. *Icarus*, 140:341–352, August 1999.
53. K.R. Grazier, W.I. Newman, F. Varadi, W.M. Kaula, and J.M. Hyman. Dynamical evolution of planetesimals in the outer solar system. II. The Saturn/Uranus and Uranus/Neptune zones. *Icarus*, 140:353–368, August 1999.
54. R. Greenberg, M.C. Nolan, W.F. Bottke, Jr., R.A. Kolvoord, and J. Veverka. Collisional history of Gaspra. *Icarus*, 107:84–+, January 1994.
55. D.T. Greenwood. *Principles of Dynamics, 2nd edition.* Prentice-Hall (Englewood Cliffs, NJ), 1988.
56. D.T. Greenwood. *Classical Dynamics*. Dover, 1997.
57. V. Guibout and D.J. Scheeres. Stability of surface motion on a rotating ellipsoid. *Celestial Mechanics and Dynamical Astronomy*, 87:263–290, November 2003.
58. E. Gustafson and D.J. Scheeres. Dynamically relevant local coordinates for halo orbits. In *Astrodynamics Specialist Conference*. American Institute of Aeronautics and Astronautics, 2008. Paper AIAA 2008-6432.
59. E.D. Gustafson and D.J. Scheeres. Optimal timing of control law updates for unstable systems with continuous control. *Journal of Guidance, Control and Dynamics*, 32(3):878–887, 2009.
60. E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*. Springer, 2006.
61. D.P. Hamilton and A.V. Krivov. Dynamics of distant moons of asteroids. *Icarus*, 128(1):241–249, 1997.
62. J.K. Harmon, M.C. Nolan, J.D. Giorgini, and E.S. Howell. Radar observations of 8P/Tuttle: A contact-binary comet. *Icarus*, 207:499–502, May 2010.

63. A.W. Harris. Tumbling asteroids. *Icarus*, 107(1):209–211, 1994.
64. T. Hashimoto, T. Kubota, S. Sawai, J. Kawaguchi, and M. Uo. Final autonomous descent based on target marker tracking. In *Astrodynamics Specialist Conference*. American Institute of Aeronautics and Astronautics, 2006. Paper AIAA 2006-6538.
65. M. Hénon. Exploration numérique du problème restreint. I. Masses égales; orbites périodiques. *Annales d'Astrophysique*, 28:499–+, February 1965.
66. M. Hénon. Exploration numérique du problème restreint. II. Masses égales, stabilité des orbites périodiques. *Annales d'Astrophysique*, 28:992–+, February 1965.
67. M. Hénon. Numerical exploration of the restricted problem. V. Hill's case: Periodic orbits and their stability. *Academie des Science Paris Comptes Rendus Serie B Sciences Physiques*, 268:223–238, 1969.
68. M. Hénon. Numerical exploration of the restricted problem. VI. Hill's case: Non-periodic orbits. *Astronomy and Astrophysics*, 9:24–36, November 1970.
69. K.A. Holsapple. Spin limits of Solar System bodies: From the small fast-rotators to 2003 EL61. *Icarus*, 187:500–509, April 2007.
70. K.C. Howell. Three-dimensional, periodic, halo orbits. *Celestial Mechanics*, 32:73–+, January 1984.
71. W. Hu. *Orbital Motion in Uniformly Rotating Second Degree and Order Gravity Fields*. PhD thesis, The University of Michigan, 2002.
72. W. Hu and D.J. Scheeres. Spacecraft motion about slowly rotating asteroids. *Journal of Guidance, Control, and Dynamics*, 25(4):765–775, 2002.
73. W. Hu and D.J. Scheeres. Numerical determination of stability regions for orbital motion in uniformly rotating second degree and order gravity fields. *Planetary and Space Science*, 52:685–692, July 2004.
74. W.-D. Hu and D.J. Scheeres. Periodic orbits in rotating second degree and order gravity fields. *Chinese Journal of Astronomy and Astrophysics*, 8:108–118, February 2008.
75. R.S. Hudson and S.J. Ostro. Shape of Asteroid 4769 Castalia (1989 PB) from inversion of radar images. *Science*, 263:940–943, February 1994.
76. R.S. Hudson and S.J. Ostro. Shape and non-principal axis spin state of Asteroid 4179 Toutatis. *Science*, 270:84–86, October 1995.
77. R.S. Hudson, S.J. Ostro, and D.J. Scheeres. High-resolution model of Asteroid 4179 Toutatis. *Icarus*, 161:346–355, February 2003.
78. M. Kaasalainen and J. Torppa. Optimization methods for asteroid lightcurve inversion. I. Shape determination. *Icarus*, 153:24–36, September 2001.
79. M. Kaasalainen, J. Torppa, and K. Muinonen. Optimization methods for asteroid lightcurve inversion. II. The complete inverse problem. *Icarus*, 153:37–51, September 2001.
80. W.M. Kaula. *Theory of Satellite Geodesy. Applications of Satellites to Geodesy*. Waltham, MA: Blaisdell, 1966.
81. J. Kawaguchi, S. Aida, and H. Morita. Hayabusa, detailed guidance and navigation operations during descents and touchdowns. In *Astrodynamics Specialist Conference*. American Institute of Aeronautics and Astronautics, 2006. Paper AIAA 2006-6536.
82. T. Kominato, M. Matsuoka, M. Uo, T. Hashimoto, and J. Kawaguchi. Optical hybrid navigation and station keeping around Itokawa. In *Astrodynamics Specialist Conference*. American Institute of Aeronautics and Astronautics, 2006. Paper AIAA 2006-6535.
83. A.S. Konopliv, J.K. Miller, W.M. Owen, D.K. Yeomans, J.D. Giorgini, R. Garmier, and J.-P. Barriot. A global solution for the gravity field, rotation, landmarks, and ephemeris of Eros. *Icarus*, 160:289–299, December 2002.
84. T. Kubota, M. Otsuki, T. Hashimoto, N. Bando, H. Yano, M. Uo, K. Shirakawa, and J. Kawaguchi. Touchdown dynamics for sampling in Hayabusa mission. In *Astrodynamics Specialist Conference*. American Institute of Aeronautics and Astronautics, 2006. Paper AIAA 2006-6539.
85. T. Kubota and T. Yoshimitsu. Intelligent rover with advanced mobility for minor body surface exploration. In H. Ayanna and E.W. Tunstel, editors, *Intelligence for Space Robotics*, chapter 6. San Antonio, TX: TSI, 2006.

86. M. Lara, A. Deprit, and A. Elipe. Numerical continuation of families of frozen orbits in the zonal problem of artificial satellite theory. *Celestial Mechanics and Dynamical Astronomy*, 62:167–181, June 1995.
87. M. Lara, J.F. Palacián, and R.P. Russell. Mission design through averaging of perturbed Keplerian systems: the paradigm of an Enceladus orbiter. *Celestial Mechanics and Dynamical Astronomy*, 108:1–22, September 2010.
88. M. Lara and R.P. Russell. Computation of a science orbit about Europa. *Journal of Guidance, Control, and Dynamics*, 30(1):259–263, 2007.
89. M. Lara, R.P. Russell, and B. Villac. Classification of the distant stability regions at Europa. *Journal of Guidance, Control, and Dynamics*, 30(2):409–418, 2007.
90. M. Lara, J.F. San-Juan, and S. Ferrer. Secular motion around synchronously orbiting planetary satellites. *Chaos*, 15:043101, 2005.
91. M. Lara and D.J. Scheeres. Stability bounds for three-dimensional motion close to asteroids. *Journal of the Astronautical Sciences*, 50(4):389–409, 2002.
92. J. Laskar. Accurate methods in general planetary theory. *Astronomy and Astrophysics*, 144:133–146, March 1985.
93. P. Lee. Dust levitation on asteroids. *Icarus*, 124:181–194, November 1996.
94. W.D. MacMillan. *Dynamics of Rigid Bodies*. McGraw-Hill, 1936.
95. W.D. Macmillan. *The Theory of the Potential*. Dover, 1958.
96. C. Magri, S.J. Ostro, D.J. Scheeres, M.C. Nolan, J.D. Giorgini, L.A.M. Benner, and J.-L. Margot. Radar observations and a physical model of Asteroid 1580 Betulia. *Icarus*, 186:152–177, January 2007.
97. C. Marchal. *The Three-Body Problem*. Studies in Astronautics, Studies in Aeronautics, 4. Amsterdam: Elsevier, 1990.
98. M.Y. Marov, V.S. Avduevsky, E.L. Akim, T.M. Eneev, R.S. Kremnev, S.D. Kulikov, K.M. Pichkhadze, G.A. Popov, and G.N. Rogovsky. Phobos-Grunt: Russian sample return mission. *Advances in Space Research*, 33(12):2276–2280, 2004.
99. B.G. Marsden, Z. Sekanina, and D.K. Yeomans. Comets and nongravitational forces. V. *Astronomical Journal*, 78:211–+, March 1973.
100. J.M. Maruskin, D.J. Scheeres, and A.M. Bloch. Dynamics of symplectic subvolumes. *SIAM Journal on Applied Dynamical Systems*, 8:180–201, 2009.
101. M. Maruya, H. Ohyama, M. Uo, N. Muranaka, H. Morita, T. Kubota, T. Hashimoto, J. Saito, and J. Kawaguchi. Navigation shape and surface topography model of Itokawa. In *Astrodynamics Specialist Conference*. American Institute of Aeronautics and Astronautics, 2006. Paper AIAA 2006-6659.
102. M.S. Matthews, R.P. Binzel, and T. Gehrels, editors. *Asteroids II*. Univ. Arizona Press, 1989.
103. C.R. McInnes. *Solar Sailing. Technology, Dynamics and Mission Applications*. Springer, London (UK), 1999.
104. J. McMahon and D.J. Scheeres. New radiation pressure force model for navigation. *Journal of Guidance, Control and Dynamics*, 33(5):1418–1428, 2010.
105. W.J. Merline, L.M. Close, C. Dumas, C.R. Chapman, F. Roddier, F. Menard, D.C. Slater, G. Duvert, C. Shelton, and T. Morgan. Discovery of a moon orbiting the asteroid 45 Eugenia. *Nature*, 401:565–+, October 1999.
106. F. Mignard and M. Hénon. About an unsuspected integrable problem. *Celestial Mechanics*, 33:239–250, July 1984.
107. J.K. Miller, A.S. Konopliv, P.G. Antreasian, J.J. Bordi, S. Chesley, C.E. Helfrich, W.M. Owen, T.C. Wang, B.G. Williams, D.K. Yeomans, and D.J. Scheeres. Determination of shape, gravity, and rotational state of Asteroid 433 Eros. *Icarus*, 155:3–17, January 2002.
108. H. Miyamoto, H. Yano, D.J. Scheeres, S. Abe, O. Barnouin-Jha, A.F. Cheng, H. Demura, R.W. Gaskell, N. Hirata, M. Ishiguro, T. Michikami, A.M. Nakamura, R. Nakamura, J. Saito, and S. Sasaki. Regolith migration and sorting on Asteroid Itokawa. *Science*, 316:1011–, May 2007.
109. J.M. Mondelo. *Contribution to the Study of Fourier Methods for QuasiPeriodic Functions and the Vicinity of the Collinear Libration Points*. PhD thesis, Departament de Matemàtica Aplicada i Anàlisi Universitat de Barcelona, 2001.

110. O. Montenbruck and E. Gill. *Satellite Orbits: Models, Methods, Applications*. Springer, 2000.
111. F.R. Moulton. *Differential Equations*. The Macmillan Company, 1930.
112. F.R. Moulton. *An Introduction to Celestial Mechanics*. New York, The Macmillan Company; London, Macmillan & Co., 1935.
113. K. Muinonen. Introducing the Gaussian shape hypothesis for asteroids and comets. *Astronomy and Astrophysics*, 332:1087–1098, 1998.
114. P.M. Muller and W.L. Sjogren. Mascons: Lunar mass concentrations. *Science*, 161:680–684, August 1968.
115. E. Mysen, A.V. Rodionov, and J.-F. Crifo. An analysis of outgassing pressure forces on the Rosetta orbiter using realistic 3D+t coma simulations. *Astronomy & Astrophysics*, 512:A69+, March 2010.
116. M. Nakamiya, D.J. Scheeres, H. Yamakawa, and M. Yoshikawa. Analysis of capture trajectories into periodic orbits about libration points. *Journal of Guidance, Control and Dynamics*, 31(5):1344–1351, 2008.
117. NASA Planetary Data System – Asteroid/Dust Archive, August 2011. http://sbn.psi.edu/pds/archive/shape.html.
118. Z.P. Olikara and K.C. Howell. Computation of quasi-periodic invariant tori in the restricted three-body problem. In *Spaceflight Mechanics 2010; Proceedings of the AAS/AIAA Space Flight Mechanics Meeting*. American Astronautical Society, 2010. Paper AAS 2010-120.
119. V. I. Oseledets. Multiplicative ergodic theorem: Characteristic Lyapunov exponents of dynamical systems. *Trudy MMO*, 19:179210, 1968.
120. S.J. Ostro, R.S. Hudson, L.A.M. Benner, J.D. Giorgini, C. Magri, J.L. Margot, and M.C. Nolan. Asteroid radar astronomy. In *Asteroids III*, pages 151–168. Univ. Arizona Press, 2002.
121. S.J. Ostro, J.-L. Margot, L.A.M. Benner, J.D. Giorgini, D.J. Scheeres, E.G. Fahnestock, S.B. Broschart, J. Bellerose, M.C. Nolan, C. Magri, P. Pravec, P. Scheirich, R. Rose, R.F. Jurgens, E.M. De Jong, and S. Suzuki. Radar imaging of binary Near-Earth Asteroid (66391) 1999 KW4. *Science*, 314:1276–1280, November 2006.
122. M. Paetzold, T. Andert, B. Häusler, S. Tellmann, J.D. Anderson, S.W. Asmar, J. Barriot, and M.K. Bird. The mass and density of (21) Lutetia from radio tracking during the Rosetta flyby. In *Bulletin of the American Astronomical Society: AAS/Division for Planetary Sciences Meeting Abstracts #42*, volume 42, page 1044, 2010.
123. R.S. Park. *Nonlinear Trajectory Navigation*. PhD thesis, The University of Michigan, 2006.
124. R.S. Park and D.J. Scheeres. Nonlinear mapping of Gaussian state uncertainties: theory and applications to spacecraft control and navigation. *Journal of Guidance, Control and Dynamics*, 29(6):1367–1375, 2006.
125. M.E. Paskowitz. *Orbit Design and Control of Planetary Satellite Orbiters in the Hill 3-Body Problem*. PhD thesis, The University of Michigan, 2006.
126. M.E. Paskowitz and D.J. Scheeres. Design of science orbits about planetary satellites: Application to Europa. *Journal of Guidance, Control and Dynamics*, 29(5):1147–1158, 2006.
127. M.E. Paskowitz and D.J. Scheeres. Robust capture and transfer trajectories for planetary satellite orbiters. *Journal of Guidance, Control, and Dynamics*, 29(2):342–353, 2006.
128. M. Paskowitz Possner and D.J. Scheeres. Control of science orbits about planetary satellites. *Journal of Guidance, Control and Dynamics*, 32(1):223–231, 2009.
129. O. Peñagaricano-Muñoa. *A Perturbation Theory for Hamilton's Principal Function: Applications to Boundary Value Problems*. PhD thesis, The University of Michigan, 2010.
130. O. Peñagaricano Muñoa and D.J. Scheeres. A perturbation theory. *Acta Astronautica*, 67(1-2):27–37, 2010.
131. H. Poincaré. *New Methods of Celestial Mechanics*. History of Modern Physics and Astronomy, New York: American Institute of Physics (AIP), 1993.
132. H. Pollard. *Celestial Mechanics*. The Carus Mathematical Monographs, Providence: Mathematical Association of America, 1976.
133. P. Pravec and A.W. Harris. Fast and slow rotation of asteroids. *Icarus*, 148:12–20, November 2000.

134. P. Pravec, A.W. Harris, P. Scheirich, P. Kušnirák, L. Šarounová, C.W. Hergenrother, S. Mottola, M.D. Hicks, G. Masi, Y.N. Krugly, V.G. Shevchenko, M.C. Nolan, E.S. Howell, M. Kaasalainen, A. Galád, P. Brown, D.R. Degraff, J.V. Lambert, W.R. Cooney, and S. Foglia. Tumbling asteroids. *Icarus*, 173:108–131, January 2005.
135. P. Pravec, P. Scheirich, P. Kušnirák, L. Šarounová, S. Mottola, G. Hahn, P. Brown, G. Esquerdo, N. Kaiser, Z. Krzeminski, D.P. Pray, B.D. Warner, A.W. Harris, M.C. Nolan, E.S. Howell, L.A.M. Benner, J.-L. Margot, A. Galád, W. Holliday, M.D. Hicks, Y.N. Krugly, D. Tholen, R. Whiteley, F. Marchis, D.R. Degraff, A. Grauer, S. Larson, F.P. Velichko, W.R. Cooney, R. Stephens, J. Zhu, K. Kirsch, R. Dyvig, L. Snyder, V. Reddy, S. Moore, Š. Gajdoš, J. Világi, G. Masi, D. Higgins, G. Funkhouser, B. Knight, S. Slivan, R. Behrend, M. Grenon, G. Burki, R. Roy, C. Demeautis, D. Matter, N. Waelchli, Y. Revaz, A. Klotz, M. Rieugné, P. Thierry, V. Cotrez, L. Brunetto, and G. Kober. Photometric survey of binary near-Earth asteroids. *Icarus*, 181:63–93, March 2006.
136. W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical recipes in FORTRAN. The Art of Scientific Computing*, 2nd edition Cambridge: University Press, 1992.
137. C.A. Renault and D.J. Scheeres. Statistical analysis of control maneuvers in an unstable orbital environments. *Journal of Guidance, Control and Dynamics*, 26(5):758–769, 2003.
138. J.E. Richardson, H.J. Melosh, C.M. Lisse, and B. Carcich. A ballistics analysis of the deep impact ejecta plume: Determining Comet Tempel 1's gravity, mass, and density. *Icarus*, 191(2):176–209, 2007.
139. K. Richter and H.U. Keller. On the stability of dust particle orbits around cometary nuclei. *Icarus*, 114:355–371, April 1995.
140. L. Rios-Reyes and D.J. Scheeres. Generalized model for solar sails. *Journal of Spacecraft and Rockets*, 42(1):182–185, 2005.
141. M.S. Robinson, P.C. Thomas, J. Veverka, S. Murchie, and B. Carcich. The nature of ponded deposits on Eros. *Nature*, 413:396–400, September 2001.
142. A.V. Rodionov, J.-F. Crifo, K. Szegő, J. Lagerros, and M. Fulle. An advanced physical model of cometary activity. *Planetary and Space Science*, 50:983–1024, August 2002.
143. S.D. Ross and D.J. Scheeres. Multiple gravity assists, capture, and escape in the restricted three-body problem. *SIAM Journal on Applied Dynamical Systems*, 6(3):576–596, 2007.
144. D.P. Rubincam. Yarkovsky thermal drag on small asteroids and Mars–Earth delivery. *Journal of Geophysical Research*, 103:1725–1732, January 1998.
145. D.P. Rubincam. Radiative spin-up and spin-down of small asteroids. *Icarus*, 148:2–11, November 2000.
146. R.P. Russell. Global search for planar and three-dimensional periodic orbits near Europa. *Journal of the Astronautical Sciences*, 54(2):199–226, 2006.
147. R.P. Russell and A.T. Brinckerhoff. Circulating eccentric orbits around planetary moons. *Journal of Guidance, Control, and Dynamics*, 32(2):423–435, 2009.
148. R.P. Russell and T. Lam. Designing ephemeris capture trajectories at Europa using unstable periodic orbits. *Journal of Guidance, Control, and Dynamics*, 30(2):482–491, 2007.
149. N.H. Samarasinha and M.J.S. Belton. Long-term evolution of rotational stress and nongravitational effects for Halley-like cometary nuclei. *Icarus*, 116:340–358, August 1995.
150. N.H. Samarasinha, B.E.A. Mueller, M.J.S. Belton, and L. Jorda. Rotation of Cometary Nuclei. In *Comets II*, pages 281–299. Univ. Arizona Press, 2004.
151. S. Sawai, D.J. Scheeres, and S. Broschart. Control of hovering spacecraft using altimetry. *Journal of Guidance, Control and Dynamics*, 25(4):786–795, 2002.
152. D.J. Scheeres. *On Symmetric Central Configurations with application to satellite Motion about Rings*. PhD thesis, The University of Michigan, 1992.
153. D.J. Scheeres. Dynamics about uniformly rotating triaxial ellipsoids: Applications to asteroids. *Icarus*, 110:225–238, August 1994.
154. D.J. Scheeres. Analysis of orbital motion around 433 Eros. *The Journal of the Astronautical Sciences*, 43:427–52, 1995.

155. D.J. Scheeres. The restricted Hill four-body problem with applications to the Earth–Moon–Sun system. *Celestial Mechanics and Dynamical Astronomy*, 70:75–98, February 1998.
156. D.J. Scheeres. Satellite dynamics about small bodies: Averaged solar radiation pressure effects. *The Journal of the Astronautical Sciences*, 47:25–46, 1999.
157. D.J. Scheeres. Stability of hovering orbits around small bodies. In *Spaceflight Mechanics 1999, Part II; Advances in the Astronautical Sciences Series*, volume 102, pages 855–875. Univelt, San Diego, 1999. AAS Paper 99-159.
158. D.J. Scheeres. The effect of C22 on orbit energy and angular momentum. *Celestial Mechanics and Dynamical Astronomy*, 73:339–348, January 1999.
159. D.J. Scheeres. Changes in rotational angular momentum due to gravitational interactions between two finite bodies. *Celestial Mechanics and Dynamical Astronomy*, 81:39–44, 2001.
160. D.J. Scheeres. Relative equilibria for general gravity fields in the sphere-restricted full 2-body problem. *Celestial Mechanics and Dynamical Astronomy*, 94:317–349, March 2006.
161. D.J. Scheeres. Orbit mechanics about small asteroids. In *2009 AAS/AIAA Space Flight Mechanics Meeting, Savannah, Georgia*, February 2009. Paper AAS 09-220.
162. D.J. Scheeres and J. Bellerose. The restricted Hill full 4-body problem: Application to spacecraft motion about binary asteroids. *Dynamical Systems: An International Journal*, 20(1):23–44, 2005.
163. D.J. Scheeres, S. Bhargava, and A. Enzian. A navigation model of the continuous outgassing field around a comet. *Interplanetary Network Progress Report*, 142:1–19, April 2000.
164. D.J. Scheeres, D.W. Dunham, R.W. Farquhar, C.E. Helfrich, J.V. McAdams, and W.M. Owen. Mission design and navigation of NEAR's encounter with Asteroid 253 Mathilde. In *Spaceflight Mechanics 1998; Proceedings of the AAS/AIAA Space Flight Mechanics Meeting, Monterey, CA*, pages 1157–1173, 1998.
165. D.J. Scheeres, E.G. Fahnestock, S.J. Ostro, J.-L. Margot, L.A.M. Benner, S.B. Broschart, J. Bellerose, J.D. Giorgini, M.C. Nolan, C. Magri, P. Pravec, P. Scheirich, R. Rose, R.F. Jurgens, E.M. De Jong, and S. Suzuki. Dynamical configuration of Binary Near-Earth Asteroid (66391) 1999 KW4. *Science*, 314:1280–1283, November 2006.
166. D.J. Scheeres, R. Gaskell, S. Abe, O. Barnouin-Jha, T. Hashimoto, J. Kawaguchi, T. Kubota, J. Saito, M. Yoshikawa, N. Hirata, T. Mukai, M. Ishiguro, T. Kominato, K. Shirakawa, and M. Uo. The actual dynamical environment about Itokawa. In *Astrodynamics Specialist Conference*. American Institute of Aeronautics and Astronautics, 2006. Paper AIAA 2006-6661.
167. D.J. Scheeres, M.D. Guman, and B. Villac. Stability analysis of planetary satellite orbiters: Application to the Europa Orbiter. *Journal of Guidance, Control and Dynamics*, 24(4):778–787, 2001.
168. D.J. Scheeres, D. Han, and Y. Hou. Influence of unstable manifolds on orbit uncertainty. *Journal of Guidance, Control and Dynamics*, 24(3):573–585, 2001.
169. D.J. Scheeres, C.M. Hartzell, P. Sánchez, and M. Swift. Scaling forces to asteroid surfaces: The role of cohesion. *Icarus*, 210:968–984, December 2010.
170. D.J. Scheeres, B. Khushalani, and R.A. Werner. Estimating asteroid density distributions from shape and gravity information. *Planetary and Space Science*, 48:965–971, August 2000.
171. D.J. Scheeres and F. Marzari. Spacecraft dynamics far from a comet. *The Journal of the Astronautical Sciences*, 50(1):35–52, 2002.
172. D.J. Scheeres, F. Marzari, L. Tomasella, and V. Vanzani. ROSETTA mission: satellite orbits around a cometary nucleus. *Planetary and Space Science*, 46:649–671, February 1998.
173. D.J. Scheeres, S.J. Ostro, R.S. Hudson, E.M. Dejong, and S. Suzuki. Dynamics of orbits close to Asteroid 4179 Toutatis. *Icarus*, 132:53–79, March 1998.
174. D.J. Scheeres, S.J. Ostro, R.S. Hudson, and R.A. Werner. Orbits close to Asteroid 4769 Castalia. *Icarus*, 121:67–87, May 1996.
175. D.J. Scheeres, B.G. Williams, and J.K. Miller. Evaluation of the dynamic environment of an asteroid: Applications to 433 Eros. *Journal of Guidance, Control and Dynamics*, 23:466–475, 2000.

176. V.A. Shishov. Determination of spacecraft and Phobos parameters of motion in the Phobos-Grunt project. *Solar System Research*, 42(4):319–328, 2008.
177. N.C. Shupe. Orbit options for an Orion-class spacecraft mission to a near-Earth object. Master's thesis, The University of Colorado, 2010.
178. Solar System Dynamics Group, Jet Propulsion Laboratory, July 2011. http://neo.jpl.nasa.gov/neo/.
179. V.G. Szebehely. *Theory of Orbits*. Academic Press, 1967.
180. Y. Takahashi and D.J. Scheeres. Small body post-rendezvous characterization via slow hyperbolic flybys. *Journal of Guidance, Control and Dynamics*, in press, 2011.
181. P.C. Thomas. Gravity, tides, and topography on small satellites and asteroids – Application to surface features of the Martian satellites. *Icarus*, 105:326–+, October 1993.
182. P.C. Thomas and M.S. Robinson. Seismic resurfacing by a single impact on the asteroid 433 Eros. *Nature*, 436:366–369, July 2005.
183. P.C. Thomas, J. Veverka, M.J.S. Belton, A. Hidy, M.F. A'Hearn, T.L. Farnham, O. Groussin, J.-Y. Li, L.A. McFadden, J. Sunshine, D. Wellnitz, C. Lisse, P. Schultz, K.J. Meech, and W.A. Delamere. The shape, topography, and geology of Tempel 1 from Deep Impact observations. *Icarus*, 187(1):4–15, 2007.
184. Y. Tsuda and D.J. Scheeres. Computation and applications of an orbital dynamics symplectic state transition matrix. *Journal of Guidance, Control and Dynamics*, 32(4):1111–1123, 2009.
185. J. Veverka, P.C. Thomas, J.F. Bell, III, M. Bell, B. Carcich, B. Clark, A. Harch, J. Joseph, P. Martin, M. Robinson, S. Murchie, N. Izenberg, E. Hawkins, J. Warren, R. Farquhar, A. Cheng, D. Dunham, C. Chapman, W.J. Merline, L. McFadden, D. Wellnitz, M. Malin, W.M. Owen, Jr., J.K. Miller, B.G. Williams, and D.K. Yeomans. Imaging of asteroid 433 Eros during NEAR's flyby reconnaissance. *Science*, 285:562–564, July 1999.
186. B. Villac and D.J. Scheeres. Escaping trajectories in the Hill three body problem and applications. *Journal of Guidance, Control and Dynamics*, 26(2):224–232, 2003.
187. B.F. Villac. *Dynamics in the Hill Problem with Applications to Spacecraft Maneuvers*. PhD thesis, The University of Michigan, 2003.
188. B.F. Villac and D.J. Scheeres. On the concept of periapsis in Hill's problem. *Celestial Mechanics and Dynamical Astronomy*, 90:165–178, September 2004.
189. S.J. Weidenschilling. How fast can an asteroid spin? *Icarus*, 46(1):124–126, 1981.
190. R.A. Werner. The gravitational potential of a homogeneous polyhedron or don't cut corners. *Celestial Mechanics and Dynamical Astronomy*, 59:253–278, July 1994.
191. R.A. Werner. Spherical harmonic coefficients for the potential of a constant-density polyhedron. *Computers & Geosciences*, 23(10):1071–1077, 1997.
192. R.A. Werner. Evaluating descent and ascent trajectories near non-spherical bodies. Technical report, Jet Propulsion Laboratory, 2010. Report Number: NPO-46697.
193. R.A. Werner and D.J. Scheeres. Exterior gravitation of a polyhedron derived and compared with harmonic and mascon gravitation representations of Asteroid 4769 Castalia. *Celestial Mechanics and Dynamical Astronomy*, 65:313–344, 1997.
194. R.A. Werner and D.J. Scheeres. Mutual potential of homogeneous polyhedra. *Celestial Mechanics and Dynamical Astronomy*, 91:337–349, March 2005.
195. E.T. Whittaker. *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, 4th edition. Cambridge, 1988.
196. W.E. Wiesel. Stable orbits about the martian moons. *Journal of Guidance Control Dynamics*, 16:434–440, 1993.
197. L.L. Wilkening, editor. *Comets*. Univ. Arizona Press, 1982.
198. D.K. Yeomans. *Comets. A Chronological History of Observation, Science, Myth, and Folklore*. New York: Wiley, 1991.
199. D.K. Yeomans, J.-P. Barriot, D.W. Dunham, R.W. Farquhar, J.D. Giorgini, C.E. Helfrich, A.S. Konopliv, J.V. McAdams, J.K. Miller, W.M. Owen, Jr., D.J. Scheeres, S.P. Synnott, and B.G. Williams. Estimating the mass of Asteroid 253 Mathilde from tracking data during the NEAR flyby. *Science*, 278:2106–+, December 1997.

200. M. Yoshikawa, H. Ikeda, H. Yano, J. Saito, T. Kubota, T. Hashimoto, A. Fujiwara, J. Kawaguchi, T. Kominato, M. Matsuoka, K. Shirakawa, T. Ohnishi, S. Abe, T. Mukai, R. Gaskell, and D. Scheeres. Astrodynamics science about Itokawa, gravity and ephemeris. In *Astrodynamics Specialist Conference*. American Institute of Aeronautics and Astronautics, 2006. Paper AIAA 2006-6658.
201. T. Yoshimitsu, J. Kawaguchi, T. Hashimoto, T. Kubota, M. Uo, H. Morita, and K. Shirakawa. Hayabusa-final autonomous descent and landing based on target marker tracking. In *Astrodynamics Specialist Conference*. American Institute of Aeronautics and Astronautics, 2006. Paper AIAA 2006-6537.
202. M. Ziebart. Generalized analytical solar radiation pressure modeling algorithm for spacecraft of complex shape. *Journal of Spacecraft and Rockets*, 41(5):840–848, 2004.

Index